VOLUME 1

Cientometria com *Inteligência Artificial*

Prompts para analisar, mapear e interpretar a ciência

Eduardo Amadeu Dutra Moresi

unesco chair

Cátedra UNESCO de Juventude, Educação e Sociedade

Volume 1

Cientometria com *Inteligência Artificial*

Prompts para analisar, mapear e interpretar a ciência

Eduardo Amadeu Dutra Moresi

Brasília, DF

2026

“The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.”

unesco chair

Cátedra UNESCO de Juventude, Educação e Sociedade

É proibida a reprodução total ou parcial desta publicação, por quaisquer meios, sem autorização prévia, por escrito, da Cátedra Unesco de Juventude, Educação e Sociedade.

A exatidão das informações, conceitos e opiniões é de exclusiva responsabilidade dos autores, os quais também se responsabilizam pelas imagens utilizadas.

Coleção Juventude, Educação e Sociedade

Comitê Editorial:

Geraldo Caliman (Coordenador), Célio da Cunha, Gilvan Charles Cerqueira de Araújo, Jenerton Arlan Schütz, Marta Helena de Freitas, Renato de Oliveira Brito.

Conselho Editorial Consultivo:

Esther Martinéz (Portugal), Azucena Ochoa Cervantes (México), Cristina Costa Lobo (Portugal), Marilia Costa Morosini (Brasil), Paulo César Nodari (Brasil).

Capa:

Laura Fugita.

Diagramação:

Laura Fugita e Júlia de Aquino.

Design Instrucional:

Isolda Gusmão.

Equipe técnica de apoio:

Isolda Gusmão, Fabiano Carvalho, Weldes Oliveira, Laura Fugita, Júlia de Aquino, e Thiago Rybalowsky.

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Moresi, Eduardo Amadeu Dutra
Cientometria com inteligência artificial [livro eletrônico] : prompts para analisar, mapear e interpretar a ciência / Eduardo Amadeu Dutra Moresi. -- Brasília : Ed. do Autor, 2026. -- (Cientometria com inteligência artificial ; 1)
PDF

Bibliografia.
ISBN 978-65-02-27154-4

1. Bibliometria 2. Bibliografia - Metodos gráficos 3. ChatGPT 4. Ciência da informação 5. Dados - Análise 6. Inteligência artificial 7. Indicadores de ciência 8. Pesquisa científica I. Título II. Série.

26-382755.1 CDD-025

Índices para catálogo sistemático:

1. Bibliometria e cientometria 025

Camila Aparecida Rodrigues - Bibliotecária CRB - SP-010133/O

Eduardo Amadeu Dutra Moresi

Possui graduação em Engenharia Eletrônica pelo Instituto Militar de Engenharia - IME - Rio de Janeiro (1989), mestrado em Engenharia Elétrica pela Universidade de Brasília - UnB (1994) e doutorado em Ciência da Informação pela UnB (2001). É professor da Universidade Católica de Brasília – UCB, desde 1997. É autor de diversas publicações científicas nacionais e internacionais. Desde 2014 coordena o Programa Apple Developer Academy da UCB. Desde 2023, coordena o Projeto TIC em Trilhas da UCB. Atua como docente e pesquisador nos Programas de Mestrado e Doutorado em Educação e de Mestrado Profissional em Governança, Tecnologia e Inovação.

Pode ser contatado pelos e-mails: moresi@p.ucb.br e eduardo.moresi@gmail.com

ORCID: https://orcid.org/0000-0001-6058-3883

LATTES: http://lattes.cnpq.br/1088020888142000

Prefácio

A ciência contemporânea é produzida em uma escala sem precedentes. O crescimento do número de publicações, a diversidade de bases bibliográficas, a expansão das redes de colaboração e a multiplicação de indicadores tornaram mais complexa a tarefa de compreender como o conhecimento é organizado, difundido e transformado. Nesse contexto, a bibliometria e a cientometria assumem papel estratégico ao permitir a análise da produção científica, das citações, das relações de colaboração, das estruturas intelectuais, das frentes de pesquisa e das mudanças temáticas ao longo do tempo. Entretanto, contar documentos, ordenar autores ou gerar mapas não é suficiente. A análise da ciência exige perguntas bem formuladas, dados adequadamente tratados, métodos coerentes e interpretações proporcionais às evidências disponíveis.

É nesse cenário que se insere a obra "Cientometria com Inteligência Artificial: *Prompts* para Analisar, Mapear e Interpretar a Ciência". O livro apresenta uma abordagem integrada para o uso da inteligência artificial generativa em estudos bibliométricos e cientométricos, acompanhando o pesquisador desde a formulação da pergunta até a produção de relatórios, mapas, indicadores e diagnósticos científicos assistidos por IA. Sua proposta não é substituir o pesquisador, o bibliotecário, o analista de ciência ou o especialista de domínio, mas ampliar sua capacidade de organizar informações, explorar dados, comparar alternativas, revisar interpretações e comunicar resultados.

A incorporação da inteligência artificial aos estudos da ciência oferece possibilidades importantes. Modelos de linguagem podem apoiar a delimitação de temas, a expansão de vocabulários, a elaboração de estratégias de busca, a normalização de termos, a classificação de documentos, a análise de resumos, a nomeação preliminar de *clusters*, a interpretação de redes e a redação de relatórios. Essas funções podem tornar a pesquisa mais ágil, sistemática e escalável. Contudo, a fluência das respostas geradas não garante sua validade. A IA pode produzir referências inexistentes, interpretações excessivas, classificações instáveis e explicações que não encontram respaldo nos dados. Por essa razão, sua utilização somente se torna cientificamente legítima quando acompanhada de curadoria humana, validação, rastreabilidade, transparência e prudência interpretativa.

Um dos principais méritos desta obra está em atribuir à engenharia de *prompts* um estatuto metodológico. O *prompt* deixa de ser entendido como simples comando e passa a ser concebido como instrumento de pesquisa. Um *prompt* cientométrico robusto explicita o papel atribuído ao modelo, a tarefa solicitada, o contexto da investigação, os dados disponíveis, a unidade de análise, os critérios de decisão, as restrições, o formato esperado da resposta e os procedimentos de validação. Quando documentado, testado e versionado, ele pode integrar o protocolo do estudo de maneira semelhante a uma estratégia de busca, a um script de análise, a uma matriz de codificação ou a uma regra de normalização.

A engenharia de *prompts* pode apoiar análises bibliométricas e cientométricas mais aprofundadas porque favorece a decomposição de problemas complexos em etapas verificáveis. Em vez de solicitar genericamente a interpretação de uma rede, o pesquisador pode pedir que a IA descreva sua estrutura, diferencie frequência de centralidade, identifique nós dominantes e periféricos, proponha rótulos alternativos para os *clusters*, explicite incertezas e indique quais documentos precisam ser consultados para validar cada interpretação. Em análises de cocitação, os *prompts* podem ajudar a distinguir influência intelectual de simples proximidade documental. No acoplamento bibliográfico, podem apoiar a identificação de

frentes contemporâneas de pesquisa. Na coocorrência de palavras-chave, podem auxiliar a diferenciar temas substantivos, conceitos, métodos, tecnologias e contextos de aplicação.

Esse aprofundamento também ocorre quando a IA é utilizada como interlocutora crítica. *Prompts* específicos podem solicitar explicações alternativas para um padrão, examinar a coerência entre perguntas e indicadores, identificar possíveis vieses, revisar afirmações causais e apontar quais evidências ainda são necessárias. Cadeias de *prompts* permitem organizar o estudo em etapas sucessivas: formular a pergunta, construir a busca, revisar o *corpus*, tratar os dados, selecionar indicadores, interpretar redes, analisar tendências e produzir a narrativa final. Cada saída deve ser validada antes de alimentar a etapa seguinte, preservando a memória metodológica da pesquisa.

Ao mesmo tempo, o livro reafirma que indicadores, redes e mapas são construções dependentes de decisões humanas. Bases de dados apresentam diferentes coberturas; estratégias de busca tornam certos documentos visíveis e outros ausentes; algoritmos, limiares e técnicas de normalização influenciam a configuração das redes; e rótulos temáticos dependem da leitura dos documentos representativos. A inteligência artificial pode auxiliar nessas decisões, mas não deve transformá-las em operações opacas. O pesquisador continua responsável pela qualidade do *corpus*, pela adequação das métricas, pela interpretação dos resultados e pelas conclusões apresentadas.

A obra oferece, assim, um percurso abrangente que integra fundamentos de cientometria, inteligência artificial, engenharia de *prompts*, formulação de perguntas, estratégias de busca, coleta e limpeza de dados, indicadores, redes, mapas temáticos, análise temporal, classificação da literatura, visualização, interpretação, escrita, ética e auditoria. Reúne ainda protocolos aplicáveis a diferentes objetos e uma biblioteca de *prompts* para formular, buscar, tratar, analisar, interpretar, redigir e revisar estudos.

Mais do que ensinar a obter respostas de uma máquina, este livro convida o leitor a formular perguntas melhores, tornar explícitos seus pressupostos, registrar suas escolhas e reconhecer os limites das evidências. Sua mensagem central é clara: a inteligência artificial deve ampliar a capacidade analítica do pesquisador, não substituir sua responsabilidade científica. Quando orientada por *prompts* bem construídos, dados confiáveis e validação humana, pode contribuir para uma cientometria mais profunda, crítica, transparente e metodologicamente responsável.

Brasília, agosto de 2026

Prof. Dr. Hércules Antônio do Prado

Universidade Católica de Brasília

Sumário

Capítulo 1 - Resumo executivo do livro 18

1.1 Síntese executiva 18
1.2 Problema central e tese do livro 19
1.3 Fundamentos conceituais: da bibliometria à cientometria aumentada por IA 20
1.4 Fluxo metodológico proposto pela obra 21
1.5 Prompts como instrumentos metodológicos 22
1.6 Indicadores, redes, mapas e análise temporal 23
1.7 Leitura, sumarização, classificação e visualização assistidas por IA 23
1.8 Interpretação, escrita e comunicação de evidências 24
1.9 Questões epistemológicas, sociológicas e éticas 25
1.10 Aplicações estratégicas e produtos do livro 25
1.11 Principais entregáveis conceituais e operacionais 26
1.12 Conclusão executiva 26

Capítulo 2 - Introdução: por que Cientometria com Inteligência Artificial? 28

2.1 Introdução 28
2.2 O que muda quando a análise da ciência passa a ser mediada por IA 29
2.3 Da bibliometria tradicional à cientometria aumentada por prompts 32
2.4 Prompts como instrumentos de pesquisa, não apenas comandos 33
2.5 O papel humano: curadoria, validação, interpretação e responsabilidade 35
2.6 Limites da IA generativa em estudos científicos 37
2.7 Prompts do capítulo 39
2.8 Quadro de aprofundamento: princípios operacionais para o uso de IA em cientometria 41
2.9 Exemplo aplicado: da ideia vaga ao problema cientométrico 43
2.10 Discussão conceitual: cientometria, IA e a nova materialidade dos métodos 44
2.11 Síntese do capítulo 45
2.12 Considerações finais 55

Capítulo 3 - Fundamentos de Cientometria 57

3.1 Introdução 57
3.2 Cientometria, bibliometria, informetria, webometria e altmetria 58
3.3 Produção científica, impacto, colaboração, circulação e visibilidade 62
3.4 Unidade de análise: artigo, autor, periódico, instituição, país, área, conceito 65
3.5 Indicadores de produtividade, impacto e colaboração 69
3.6 A diferença entre medir, mapear, avaliar e interpretar a ciência 71

3.7 Prompts do capítulo 74
3.8 Considerações finais 76

Capítulo 4 - Inteligência Artificial aplicada à Cientometria 78

4.1 Introdução 78
4.2 O que a IA pode fazer em um estudo cientométrico 80
4.3 O que a IA não deve fazer sem validação humana 86
4.4 IA generativa, modelos de linguagem e análise textual 88
4.5 Classificação, sumarização, extração de entidades e apoio interpretativo 91
4.6 Exemplos aplicados de uso controlado da IA 93
4.7 Riscos: alucinação, viés, falsa autoridade, erros de citação e perda de rastreabilidade 94
4.8 Prompts do capítulo 105
4.9 Considerações finais 107
Apêndice A.4 - Sugestão para registrar uso de IA 108

Capítulo 5 - Engenharia de Prompts para Estudos Cientométricos 109

5.1 Introdução 109
5.2 Anatomia de um bom prompt científico 111
5.3 Contexto, papel, tarefa, dados, restrições e formato de saída 115
5.4 Prompts exploratórios, analíticos, críticos e redacionais 118
5.5 Cadeias de prompts: da pergunta ao relatório 122
5.6 Como documentar prompts em protocolos reprodutíveis 125
5.7 Temperatura, variação, consistência e controle de respostas 128
5.8 Prompts do capítulo 131
5.9 Discussão integrada: engenharia de prompts como competência cientométrica 135
5.10 Conclusão 138
Apêndice A.5 - Matriz sintética para desenho de prompts cientométricos 141

Capítulo 6 - Formulação da pergunta cientométrica 142

6.1 Introdução 142
6.2 Perguntas sobre produção científica 144
6.3 Perguntas sobre impacto e influência 146
6.4 Perguntas sobre colaboração e redes 149
6.5 Perguntas sobre emergência de temas 152
6.6 Perguntas sobre políticas científicas e avaliação institucional 154
6.7 Transformando objetivos vagos em perguntas operacionalizáveis 157
6.8 Prompts do capítulo 160
6.9 Exemplo aplicado 164
6.10 Diretrizes práticas para formulação de perguntas cientométricas com IA 165

6.11 Considerações finais 166
Apêndice A.6 - Checklist de formulação da pergunta cientométrica 167

Capítulo 7 - Estratégias de busca com apoio de IA 175

7.1 Introdução 175
7.2 Palavras-chave, descritores e termos emergentes 176
7.3 Estratégias booleanas para bases bibliográficas 179
7.4 Tradução terminológica entre áreas do conhecimento 184
7.5 Busca sensível versus busca específica 186
7.6 Teste, refinamento e documentação da estratégia de busca 189
7.7 Como evitar que a IA invente termos, autores ou bases inexistentes 191
7.8 Prompts do capítulo 193
7.9 Quadro metodológico: da ideia inicial à estratégia documentada 196
7.10 Exemplo aplicado: busca sobre prompts em estudos cientométricos 197
7.11 Critérios de qualidade para estratégias de busca assistidas por IA 197
7.12 Considerações finais 198
Apêndice A.7.1 - Checklist operacional para busca com IA 198
Apêndice A.7.2 - Exemplos de decisões de busca e justificativas metodológicas 199

Capítulo 8 - Coleta de dados bibliográficos 202

8.1 Introdução 202
8.2 Fontes de dados: bases comerciais, abertas e institucionais 203
8.3 Metadados essenciais: título, resumo, palavras-chave, autores, afiliações, referências e citações 207
8.4 Exportação de registros e formatos de arquivo 210
8.5 Uso de APIs em estudos cientométricos 213
8.6 Coleta manual, semiautomática e automatizada 216
8.7 Registro do processo de coleta para reprodutibilidade 219
8.8 Prompts do capítulo 221
8.9 Exemplo aplicado: planejamento de coleta para um estudo sobre IA generativa na educação superior 225
8.10 Considerações finais 225
Apêndice A.8 - Perguntas de auditoria da coleta 230

Capítulo 9 - Pré-processamento e limpeza de dados com prompts 233

9.1 Introdução 233
9.2 Padronização de nomes de autores 234
9.3 Desambiguação de instituições 235
9.4 Tratamento de duplicatas 236

9.5 Normalização de palavras-chave 237
9.6 Agrupamento de sinônimos e variantes terminológicas 237
9.7 Verificação humana e amostragem de qualidade 238
9.8 A limpeza como parte do desenho de pesquisa 239
9.9 Modelos de prompts para pré-processamento e limpeza 239
9.10 Estudo de caso ilustrativo: limpeza de uma base sobre inteligência artificial na educação 241
9.11 Matriz de riscos do pré-processamento assistido por IA 242
9.12 Protocolo reprodutível de limpeza com prompts 243
9.13 Checklist comentado para equipes de pesquisa 244
9.14 Considerações finais 245

Capítulo 10 - Indicadores cientométricos 247

10.1 Introdução 247
10.2 Indicadores de produção 248
10.3 Indicadores de citação e impacto 250
10.4 Indicadores de colaboração 253
10.5 Indicadores de interdisciplinaridade 255
10.6 Indicadores de internacionalização 256
10.7 Indicadores alternativos e atenção social 258
10.8 Limitações éticas e técnicas do uso de métricas 260
10.9 Prompts aplicados ao uso de indicadores cientométricos 261
10.10 Matriz de decisão para escolha de indicadores 264
10.11 Aprofundamento metodológico sobre indicadores e prompts 265
10.12 Complementos operacionais para indicadores em relatórios assistidos por IA 268
10.13 Biblioteca comentada de usos de IA para indicadores 271
10.14 Estudo de caso ilustrativo: indicadores para IA em saúde pública 273
10.15 Nota metodológica final sobre prudência interpretativa 273
10.16 Checklist de qualidade para relatórios de indicadores 274
10.17 Considerações finais 274

Capítulo 11 - Redes científicas e análise relacional 276

11.1 Introdução 276
11.2 Redes de coautoria 277
11.3 Redes de cocitação 280
11.4 Redes de acoplamento bibliográfico 282
11.5 Redes de coocorrência de palavras-chave 283
11.6 Redes institucionais e geográficas 286
11.7 Centralidade, modularidade, clusters e comunidades 288

11.8 Como usar IA para interpretar redes sem inventar causalidades 290
11.9 Cadeia metodológica para análise relacional assistida por IA 292
11.10 Prompts do capítulo 295
11.11 Estudo de caso ilustrativo: interpretando uma rede sem causalidades 297
11.12 Matriz de riscos interpretati*vos* em redes científicas 299
11.13 Redação de resultados de redes com apoio de IA 301
11.14 Comparação entre tipos de rede e triangulação interpretativa 301
11.15 Reprodutibilidade, documentação e governança dos prompts em redes científicas 303
11.16 Validação humana de clusters, rótulos e narrativas geradas por IA 303
11.17 Boas práticas para figuras, legendas e comunicação visual de redes 304
11.18 Integração entre análise quantitativa de redes e leitura qualitativa assistida por IA 305
11.19 Conclusão 306

Capítulo 12 - Mapas temáticos e estrutura conceitual da ciência 310

12.1 Introdução 310
12.2 Temas motores, básicos, emergentes e declinantes 312
12.3 Frentes de pesquisa 314
12.4 Escolas teóricas e comunidades epistêmicas 318
12.5 Mapas conceituais a partir de títulos, resumos e palavras-chave 320
12.6 IA para rotular clusters temáticos 323
12.7 Validação dos rótulos por especialistas 324
12.8 Prompts do capítulo 326
12.9 Protocolo recomendado para mapas temáticos assistidos por IA 328
12.10 Estudo de caso ilustrativo: IA, avaliação da pesquisa e ciência aberta 330
12.11 Notas metodológicas complementares 331
12.12 Integração entre mapas temáticos, indicadores e interpretação qualitativa 332
12.13 Erros comuns na leitura de mapas temáticos e estratégias de mitigação 336
12.14 Do mapa ao relatório: redação científica assistida por prompts 337
12.15 Aplicações em política científica, gestão de pesquisa e organização do conhecimento 340
12.16 Modelos avançados de prompts para análise temática reprodutível 342
12.17 Reprodutibilidade, versionamento e preservação das decisões interpretativas 343
12.18 Considerações finais 344

Capítulo 13 - Análise temporal e detecção de tendências 346

13.1 Introdução 346
13.2 Crescimento da literatura 347
13.3 Ciclo de vida de temas científicos 349
13.4 Detecção de temas emergentes 351
13.5 Mudanças terminológicas ao longo do tempo 353

13.6 Trajetórias de autores, instituições e países....354
13.7 Cenários prospectivos com cautela metodológica 356
13.8 O papel da inteligência artificial na análise temporal 358
13.9 Diferenciando tendência real, flutuação e artefato de busca 359
13.10 Prompts do capítulo 360
13.11 Estudo de caso ilustrativo: a evolução de um tema cientométrico hipotético 362
13.12 Checklist metodológico para análises temporais assistidas por IA....363
13.13 Protocolo resumido de reprodutibilidade para o capítulo 364
13.14 Técnicas quantitativas e híbridas para análise temporal 365
13.15 Visualização longitudinal e comunicação de tendências....367
13.16 Qualidade dos dados, bases bibliográficas e vieses temporais 368
13.17 Validação humana, amostragem e triangulação interpretativa....370
13.18 Redação de resultados temporais com apoio de prompts....371
13.19 Governança, ética e responsabilidade na detecção de tendências 372
13.20 Considerações finais....373

Capítulo 14 - IA para leitura, sumarização e classificação de literatura....375

14.1 Introdução....375
14.2 Sumarização de grandes conjuntos de resumos 376
14.3 Classificação temática assistida....379
14.4 Extração de métodos, objetos e populações estudadas 381
14.5 Identificação de lacunas de pesquisa....383
14.6 Comparação entre revisões sistemáticas e estudos cientométricos 385
14.7 Controle de qualidade em classificações feitas por IA 387
14.8 Cadeia de prompts para leitura, sumarização e classificação....390
14.9 Estudo de caso ilustrativo 392
14.10 Notas metodológicas 393
14.11 Roteiro para documentação reprodutível....394
14.12 Síntese conclusiva 395
Apêndice A.14 - Notas complementares....396

Capítulo 15 - Visualização científica assistida por IA 398

15.1 Introdução....398
15.2 Gráficos de produção e impacto....400
15.3 Mapas de rede....403
15.4 Diagramas de Sankey, mapas temáticos e linhas do tempo 405
15.5 Dashboards para gestão científica 409
15.6 Como pedir à IA recomendações de visualização....412
15.7 Como evitar visualizações bonitas, mas metodologicamente frágeis 414

15.8 Prompts do capítulo....418
15.9 Estudo de caso ilustrativo: visualização de uma área emergente....420
15.10 Checklist de qualidade e recomendações finais....420
15.11 Protocolo operacional para visualização científica assistida por IA....422
15.12 Boas práticas por público e meio de comunicação....423
15.13 Ética, governança e responsabilidade nas visualizações assistidas por IA....424
15.14 Exemplo aplicado de sequência de prompts para um relatório visual....426
15.15 Conclusão....427

Capítulo 16 - Interpretação cientométrica com IA 429

16.1 Introdução....429
16.2 Da descrição dos dados à explicação prudente....430
16.3 Como interpretar indicadores em contexto....434
16.4 Como formular hipóteses sem exagerar conclusões....434
16.5 Triangulação entre métricas, literatura e conhecimento especializado....435
16.6 O papel da IA como “interlocutor crítico”....436
16.7 Redação interpretativa com transparência metodológica....437
16.8 Prompts do capítulo: interpretação, auditoria e redação....437
16.9 Matriz prática de auditoria interpretativa....439
16.10 Estudo de caso ilustrativo....440
16.11 Checklist de redação interpretativa....441
16.12 Conclusão....442

Capítulo 17 - Escrita de relatórios, artigos e pareceres cientométricos... 443

17.1 Introdução....443
17.2 Estrutura de um relatório cientométrico....444
17.3 Escrita da metodologia....446
17.4 Escrita dos resultados....448
17.5 Escrita da discussão....450
17.6 Produção de sumários executivos....452
17.7 Comunicação para gestores, pesquisadores e formuladores de políticas....453
17.8 Prompts do capítulo....455
17.9 Protocolos de escrita assistida por IA....456
17.10 Estudo de caso ilustrativo....459
17.11 Checklist de qualidade para escrita cientométrica....460
17.12 Escrita de pareceres cientométricos....460
17.13 Erros frequentes na escrita cientométrica....461
17.14 Governança editorial e versões do documento....462
17.15 Ética da comunicação, responsabilização e usos posteriores....463

17.16 Referências, citações e integridade bibliográfica 465
17.17 Ampliação analítica: notas complementares para redação responsável 466
17.18 Considerações finais 468

Capítulo 18 - Questões epistemológicas, sociológicas e éticas 469

18.1 Introdução 469
18.2 O que os indicadores mostram e o que escondem 470
18.3 Ciência publicada versus ciência em andamento 472
18.4 Citações como reconhecimento, retórica, convenção ou estratégia 474
18.5 Viés linguístico, geográfico, disciplinar e de base de dados 475
18.6 A falsa neutralidade dos mapas e rankings 476
18.7 IA, opacidade e responsabilidade analítica 478
18.8 Transparência no uso de prompts 479
18.9 Síntese integradora: da crítica das métricas à governança dos prompts 480
18.10 Prompts do capítulo 482
18.11 Estudo de caso ilustrativo: avaliando a ética de um mapa cientométrico sobre IA na saúde 485
18.12 Checklist de qualidade epistemológica, sociológica e ética 485
18.13 Orientações para redação ética de resultados cientométricos 486
18.14 Protocolo de auditoria crítica de indicadores e mapas 487
18.15 Justiça cognitiva, diversidade epistêmica e invisibilidade bibliométrica 490
18.16 Matriz decisória para conclusões cientométricas assistidas por IA 491
18.17 Governança institucional do uso de métricas e IA 493
18.18 Critérios mínimos para publicação de estudos cientométricos com IA 494
18.19 Conclusão 495
Apêndice A.18 - Modelo de registro ético do uso de IA em estudo cientométrico 495

Capítulo 19 - Aplicações da cientometria com inteligência artificial 497

19.1 Introdução 497
19.2 Diagnóstico de áreas de pesquisa 498
19.3 Identificação de temas emergentes 499
19.4 Apoio à política científica 500
19.5 Avaliação de programas de pós-graduação 500
19.6 Estratégias de colaboração 501
19.7 Inteligência competitiva acadêmica 502
19.8 Prospecção tecnológica 502
19.9 Organização do conhecimento e curadoria de acervos 503
19.10 Matriz operacional das aplicações 503
19.11 Integração entre aplicações: da evidência à decisão responsável 504

19.12 Prompts do capítulo 505
19.13 Checklist de qualidade para aplicações com IA 507
19.14 Estudo de caso ilustrativo: saúde digital em uma universidade pública 507
19.15 Riscos transversais e estratégias de mitigação 508
19.16 Considerações finais 508
Apêndice A.19 - Protocolo operacional para aplicações da cientometria com IA 509

Capítulo 20 - Protocolos completos de estudo 511

20.1 Introdução: protocolos como infraestrutura metodológica da cientometria com IA 511
20.2 Arquitetura comum dos protocolos 512
20.3 Protocolo para análise de uma área científica 513
20.4 Protocolo para análise de autores e grupos de pesquisa 516
20.5 Protocolo para análise institucional 519
20.6 Protocolo para análise de temas emergentes 521
20.7 Protocolo para análise de colaboração internacional 524
20.8 Protocolo para relatório de política científica 526
20.9 Protocolos comparados: quando usar cada um 529
20.10 Documentação reprodutível dos protocolos 529
20.11 Qualidade, validação e triangulação 530
20.12 Matriz de prompts transversais 531
20.13 Estudo de caso ilustrativo: da pergunta ao produto final 532
20.14 Erros frequentes na aplicação de protocolos 533
20.15 Checklist final para protocolos completos 534
20.16 Considerações finais 535
Apêndice A.20 - Operacionalização, governança e anexos dos protocolos 535

Capítulo 21 - Biblioteca de prompts cientométricos 546

21.1 Prompts para formular perguntas 547
21.2 Prompts para busca bibliográfica 552
21.3 Prompts para limpeza de dados 555
21.4 Prompts para indicadores 559
21.5 Prompts para redes 562
21.6 Prompts para mapas temáticos 565
21.7 Prompts para interpretação 569
21.8 Prompts para redação 572
21.9 Prompts para revisão crítica 575
21.10 Prompts para auditoria e reprodutibilidade 579
21.11 Modelo integrado de cadeia de prompts 582
21.12 Ficha padrão de documentação de prompt 583

21.13 Critérios de qualidade e matriz de riscos 584
21.14 Estudo de caso integrador: da demanda institucional ao relatório reprodutível 585
21.15 Adaptação dos prompts por área, idioma e maturidade do campo 586
21.16 Banco de meta-prompts para calibrar a biblioteca 587
21.17 Boas práticas para equipes de pesquisa 589
21.18 Integração com ferramentas cientométricas 590
21.19 Limites da biblioteca e agenda de atualização 590
21.20 Considerações finais 591

Capítulo 22 - Conclusão: cientometria com inteligência artificial como prática metodológica responsável 593

22.1 Tese de encerramento 593
22.2 Síntese do percurso metodológico da obra 593
22.3 Engenharia de prompts como infraestrutura metodológica 595
22.4 IA como interlocutor crítico, não como autoridade epistêmica 596
22.5 Limites e riscos persistentes 596
22.6 Perspectivas futuras 597
22.7 Princípios para uma cientometria aumentada por IA 598
22.8 Implicações para pesquisa, gestão e política científica 598
22.9 Agenda de pesquisa 599
22.10 Considerações finais e fechamento prospectivo 600

Referências Bibliográficas 601

Capítulo 1

Resumo executivo do livro

1.1 Síntese executiva

Este livro propõe uma abordagem integrada para a cientometria em contexto de inteligência artificial generativa. Seu tema central é a transformação do trabalho cientométrico por meio de *prompts* documentados, verificáveis e metodologicamente orientados. A IA não é apresentada como substituta do pesquisador, do bibliotecário, do gestor de pesquisa ou do especialista em avaliação científica, mas como uma camada adicional de mediação analítica. Essa camada pode acelerar a formulação de perguntas, a preparação de estratégias de busca, a limpeza de metadados, a seleção de indicadores, a interpretação de redes, a redação de relatórios e a comunicação de evidências. Tal aceleração, contudo, somente se torna cientificamente legítima quando acompanhada de curadoria humana, validação, rastreabilidade, transparência e prudência interpretativa.

A obra contrapõe-se a duas simplificações recorrentes: reduzir a cientometria a rankings, contagens e gráficos de impacto; e presumir que a fluência da IA generativa equivale à capacidade de interpretar automaticamente a ciência. Indicadores, mapas e narrativas cientométricas são construções metodológicas situadas, dependentes da escolha das bases, da delimitação temporal, dos critérios de busca, da deduplicação, da normalização de nomes, das métricas, dos algoritmos, da visualização e da interpretação. Modelos de linguagem podem, ainda, produzir afirmações factualmente frágeis ou referências inexistentes quando não há controle metodológico (MAYNEZ *et al.*, 2020; JI *et al.*, 2023).

O fio condutor do livro é a ideia de que *prompts* podem ser tratados como instrumentos de pesquisa. Um *prompt* científico não é apenas uma pergunta enviada a um sistema conversacional; é uma instrução analítica que explicita papel, contexto, tarefa, dados disponíveis, restrições, critérios de saída, nível de incerteza e forma de validação. Quando documentado, versionado e testado, integra o protocolo de pesquisa; quando improvisado, pode produzir ruído, viés, falsa autoridade e perda de rastreabilidade.

A contribuição executiva da obra é oferecer uma arquitetura completa para a análise da ciência com apoio de IA: fundamentos conceituais; desenho de perguntas e estratégias de busca; coleta e limpeza de dados; indicadores, redes e mapas temáticos; análise temporal; leitura assistida e visualização; interpretação e escrita; ética; aplicações; protocolos completos; e biblioteca de *prompts*. O livro dirige-se a universidades, agências de fomento, programas de pós-graduação, bibliotecas acadêmicas, núcleos de inteligência científica, grupos de pesquisa e órgãos de formulação de políticas públicas.

Resultado esperado. Ao final da obra, o leitor deverá ser capaz de formular problemas cientométricos mais precisos, construir buscas robustas, selecionar indicadores coerentes, interpretar mapas sem transformar associação em causalidade, auditar respostas de IA, redigir relatórios transparentes e comunicar evidências a públicos especializados e não especializados. A proposta não é automatizar a cientometria, mas torná-la mais reflexiva, documentada, escalável e criticamente mediada por IA.

Figura 1 - Ciclo da Cientometria Aumentada

Fonte: Imagem gerada por IA, ChatGPT (2026).

1.2 Problema central e tese do livro

A produção científica contemporânea cresce em escala, diversidade e complexidade. A multiplicação de artigos, *preprints*, anais de conferências, livros, dados de pesquisa, repositórios, citações abertas, perfis institucionais e indicadores alternativos amplia as possibilidades de estudar a ciência, mas torna sua interpretação mais exigente. A cientometria permite observar padrões de publicação, impacto, colaboração, circulação, especialização temática, emergência de áreas e mudanças na estrutura intelectual do conhecimento, dando continuidade à tradição de estudar redes e relações documentais da ciência (GARFIELD, 1955; PRICE, 1965).

Nesse cenário, a IA generativa é uma tecnologia ambivalente. Modelos de linguagem podem sintetizar resumos, sugerir vocabulários de busca, transformar objetivos vagos em perguntas operacionalizáveis, classificar documentos, extrair entidades, comparar métricas, revisar limitações e produzir relatórios preliminares. Os mesmos modelos, porém, podem inventar referências, afirmar conclusões sem evidência, ignorar vieses de cobertura e confundir correlação com causalidade (MAYNEZ *et al.*, 2020; JI *et al.*, 2023). A tese do livro é que a utilidade da IA na cientometria depende menos da fluência da resposta e mais da qualidade do protocolo que orienta seu uso.

A partir desses princípios, os *prompts* são apresentados como dispositivos para explicitar o método, e não como atalhos para substituí-lo. Um bom *prompt* obriga o analista a declarar o objetivo, os dados de entrada, o formato de saída, as inferências permitidas, os limites e a forma de verificação. Essa lógica permite padronizar procedimentos, comparar versões, auditar decisões e reduzir ambiguidades.

Questão central. Como usar IA para ampliar a capacidade de analisar, mapear e interpretar a ciência sem perder rigor, responsabilidade e transparência? A resposta proposta distribui o trabalho em camadas sucessivas - pergunta, busca, coleta, limpeza, análise, visualização, interpretação, comunicação e auditoria - nas quais a IA atua como assistente metodológico, nunca como instância final de verdade.

Quadro 1.1 - Princípios estruturantes da cientometria com IA

Princípio	Exigência metodológica	Implicação prática
Rastreabilidade	Toda decisão relevante deve poder ser reconstruída, da pergunta ao relatório final.	Registrar bases, *strings*, datas, critérios, dados, ferramentas, parâmetros, *prompts* e decisões humanas.

Princípio	Exigência metodológica	Implicação prática
Proporcionalidade interpretativa	A força da conclusão deve ser compatível com a qualidade do dado e o desenho do estudo.	Não converter crescimento em consolidação, coautoria em colaboração substantiva ou centralidade em liderança.
Validação situada	Resultados devem ser avaliados à luz da literatura, do conhecimento do domínio, das limitações e da finalidade do estudo.	Triangular métricas, conteúdo textual, especialistas e contexto institucional.

Fonte: Elaborado pelo autor (2026).

1.3 Fundamentos conceituais: da bibliometria à cientometria aumentada por IA

Bibliometria, cientometria, informetria, webometria e altmetria são campos inter-relacionados voltados à análise quantitativa da informação, da comunicação científica e dos rastros de circulação do conhecimento. A bibliometria concentra-se em documentos, publicações, citações, periódicos e padrões de comunicação registrada; a cientometria amplia o foco para a ciência como sistema social, cognitivo e institucional. A consolidação das análises de citações, redes de documentos e mapeamento científico é sustentada por contribuições clássicas e contemporâneas (GARFIELD, 1955; PRICE, 1965; ARIA; CUCCURULLO, 2017).

A cientometria aumentada por IA não substitui esses fundamentos; ela os reconfigura operacionalmente. A IA pode traduzir problemas substantivos em perguntas mensuráveis, identificar termos relacionados, sugerir fontes, classificar resumos, apoiar a rotulagem de *clusters*, redigir descrições de mapas e revisar inconsistências. A evolução das arquiteturas de transformadores e dos modelos de linguagem ampliou essas capacidades de processamento textual (VASWANI *et al.*, 2017; DEVLIN *et al.*, 2019; BROWN *et al.*, 2020). Sem domínio de conceitos como unidade de análise, cobertura, citação, colaboração, impacto, normalização, rede, densidade, centralidade e visibilidade, contudo, as respostas podem ser plausíveis e metodologicamente frágeis.

Um ponto essencial é diferenciar medir, mapear, avaliar e interpretar. Medir significa produzir indicadores segundo regras definidas; mapear, representar relações entre autores, instituições, documentos, conceitos ou referências; avaliar, atribuir valor com base em critérios explícitos; e interpretar, situar resultados em contexto histórico, institucional, disciplinar e epistemológico. Indicadores quantitativos devem apoiar, e não substituir, o julgamento qualitativo e contextualizado (HICKS *et al.*, 2015).

A escolha da unidade de análise é igualmente decisiva. Artigos, autores, periódicos, instituições, países, áreas, conceitos, referências, citações, palavras-chave, patentes, projetos e programas de pós-graduação produzem perguntas e indicadores diferentes. A IA pode auxiliar na construção de matrizes entre pergunta, unidade, fonte, indicador e visualização, mas a decisão final precisa ser teoricamente e metodologicamente justificada.

Esses fundamentos apontam para uma cientometria que combina análise quantitativa, interpretação relacional e consciência crítica. A IA contribui para escala, sistematização e comunicação; sua utilidade, entretanto, depende do acoplamento a conceitos cientométricos bem definidos, evitando tanto o tecnicismo vazio quanto o entusiasmo acrítico.

1.4 Fluxo metodológico proposto pela obra

O livro organiza a cientometria com IA como um fluxo de trabalho no qual cada etapa depende das decisões anteriores e condiciona as posteriores. A estrutura não é uma sequência mecânica: cada passagem exige registro, validação e possibilidade de retorno quando surgem ruídos, lacunas ou inconsistências.

Quadro 1.2 - Fluxo metodológico da cientometria assistida por IA

Etapa	Foco	Apoio possível da IA	Controle indispensável
1. Pergunta	Delimitar problema, escopo e unidade de análise.	Gerar alternativas, decompor objetivos e explicitar limitações.	Revisar viabilidade, relevância e coerência entre pergunta e método.
2. Busca	Construir vocabulário, blocos conceituais e estratégia por base.	Expandir termos, traduzir variantes e adaptar sintaxes.	Testar em bases reais e documentar data, campos, filtros e resultados (PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021).
3. Coleta	Selecionar fontes e preservar dados brutos.	Planejar metadados, checklists e pseudocódigo de APIs.	Reconhecer diferenças de cobertura e qualidade entre fontes (PRIEM; PIWOWAR; ORR, 2022; VISSER; VAN ECK; WALTMAN, 2021; MARTÍN-MARTÍN *et al.*, 2021).
4. Limpeza	Deduplicar e normalizar autores, instituições, termos e referências.	Sugerir agrupamentos, equivalências e casos ambíguos.	Registrar dicionários de normalização e validar decisões humanas.
5. Análise	Calcular indicadores, redes, mapas e séries temporais.	Apoiar escolha de métricas, leitura preliminar e rotulagem.	Documentar parâmetros e usar ferramentas apropriadas (CHEN, 2006; COBO *et al.*, 2011; VAN ECK; WALTMAN, 2010).
6. Interpretação	Separar descrição, inferência, hipótese e recomendação.	Atuar como interlocutor crítico e propor explicações alternativas.	Triangular resultados com literatura, conteúdo e especialistas.
7. Comunicação e auditoria	Produzir artigos, relatórios, pareceres e briefings.	Adaptar linguagem, revisar limitações e checar consistência.	Preservar incertezas, responsabilidade autoral e rastreabilidade.

Fonte: Elaborado pelo autor (2026).

A busca bibliográfica é tratada como decisão epistemológica, não como simples operação técnica. Uma *string* excessivamente sensível pode incorporar ruído, enquanto uma estratégia excessivamente específica pode excluir literatura relevante. Da mesma forma, não existe base universalmente superior: há fontes mais adequadas a determinados problemas, campos e objetivos, com coberturas e metadados distintos (VISSER; VAN ECK; WALTMAN, 2021; MARTÍN-MARTÍN *et al.*, 2021).

O fluxo culmina em protocolos completos e em uma biblioteca de *prompts*. Cada estudo pode ser documentado como encadeamento de decisões verificáveis, de modo que a IA deixe

de ser uma ferramenta informal e passe a integrar um sistema de produção de evidências com rastreabilidade, auditoria e responsabilidade.

1.5 Prompts como instrumentos metodológicos

Um dos aportes centrais da obra é redefinir o *prompt* como artefato metodológico. Seu valor não reside apenas no resultado gerado, mas na capacidade de tornar explícitas as condições de produção desse resultado. Um *prompt* cientométrico deve indicar papel, contexto, tarefa, dados disponíveis, restrições, formato de saída, nível de incerteza, regras de validação e proibições, como não inventar dados, autores, bases, DOI ou relações causais.

A engenharia de *prompts* envolve desenho de pesquisa. Ao solicitar indicadores, o analista deve informar pergunta, unidade de análise, fonte, período, campo e finalidade. Ao pedir interpretação de uma rede, deve exigir separação entre descrição estrutural, possíveis explicações, hipóteses, evidências necessárias e riscos. Ao solicitar redação, deve proibir extrapolação e exigir correspondência entre afirmação e dado.

O livro também introduz cadeias de *prompts*: sequências em que uma saída validada alimenta a etapa seguinte. Essa abordagem aproxima a interação com IA de um protocolo computacional e interpretativo. Técnicas como *few-shot prompting*, *chain-of-thought* e *self-consistency* demonstram que a formulação das instruções altera o comportamento dos modelos, reforçando a necessidade de teste e documentação (BROWN *et al.*, 2020; WEI *et al.*, 2022; WANG *et al.*, 2023; SCHULHOFF *et al.*, 2024).

Quadro 1.3 - Famílias funcionais de *prompts* cientométricos

Família	Finalidade	Exemplos de uso	Risco principal
Exploratórios	Abrir possibilidades e delimitar o problema.	Formular perguntas, levantar termos e mapear perspectivas.	Ampliar excessivamente o escopo ou sugerir termos frágeis.
Analíticos	Organizar, comparar e interpretar dados.	Classificar documentos, extrair entidades e comparar métricas.	Tratar classificação probabilística como resultado definitivo.
Críticos	Examinar fragilidades, vieses e força das conclusões.	Revisar causalidade, limitações e coerência entre dado e argumento.	Produzir críticas genéricas ou descontextualizadas.
Redacionais	Transformar evidências validadas em texto.	Escrever metodologia, resultados, discussão e sumários executivos.	Preencher lacunas com procedimentos ou resultados inexistentes.
Auditoria	Verificar rastreabilidade e reprodutibilidade.	Checar protocolos, referências, parâmetros e decisões humanas.	Confundir documentação reconstruída com registro contemporâneo.

Fonte: Elaborado pelo autor (2026).

Os *prompts* devem ser versionados. Mudanças no papel atribuído ao modelo, nos exemplos, no formato de saída, no grau de confiança solicitado ou nas restrições podem modificar a resposta. Por isso, recomenda-se registrar versão, modelo, data, parâmetros quando disponíveis, dados de entrada, saída, revisão humana e decisão final. Em contextos institucionais, essa documentação é essencial para auditoria, replicação e governança.

Ao tratar *prompts* como instrumentos metodológicos, a obra desloca a questão genérica "a IA pode fazer cientometria?" para uma pergunta operacional: em quais etapas, sob quais controles, com quais dados, limites e formas de validação a IA pode apoiar a análise cientométrica?

1.6 Indicadores, redes, mapas e análise temporal

A análise cientométrica combina múltiplas famílias de evidências. Indicadores de produção informam volume, crescimento e distribuição; indicadores de citação e impacto observam reconhecimento e circulação; indicadores de colaboração representam coautorias e redes institucionais; indicadores de interdisciplinaridade examinam diversidade e integração; e indicadores alternativos captam formas de atenção social. Nenhuma dessas medidas equivale, isoladamente, à qualidade científica. O índice h, por exemplo, foi concebido para combinar produtividade e impacto de citação, mas permanece dependente de carreira, campo e base de dados (HIRSCH, 2005).

A obra recomenda explicitar o que cada métrica mede, o que não mede, de quais dados depende, quais vieses carrega e que interpretações indevidas pode estimular. Essa orientação está alinhada às métricas responsáveis, segundo as quais a avaliação quantitativa deve apoiar o julgamento qualitativo, e não substituí-lo (HICKS *et al.*, 2015). A IA pode ajudar a selecionar indicadores coerentes e explicar limitações, mas deve reforçar a crítica metodológica.

No campo das redes científicas, são examinadas coautoria, cocitação, acoplamento bibliográfico, coocorrência de palavras-chave e relações institucionais ou geográficas. A cocitação permite observar proximidades intelectuais construídas pela literatura posterior, enquanto o acoplamento identifica documentos que compartilham referências (SMALL, 1973; KESSLER, 1963). Algoritmos de mapeamento e *cluster*ing permitem representar essas estruturas, mas seus resultados dependem de normalização, limiares e parâmetros (WALTMAN; VAN ECK; NOYONS, 2010).

Mapas temáticos permitem distinguir temas motores, básicos, emergentes e declinantes, frentes de pesquisa e comunidades epistêmicas. A IA é útil na nomeação preliminar de *clusters*, na comparação de rótulos e na elaboração de narrativas, desde que os rótulos sejam validados por documentos representativos e especialistas. Um *cluster* não é uma essência temática natural: é resultado de dados, métodos e interpretação.

A análise temporal complementa indicadores e mapas ao examinar crescimento, ciclo de vida, emergência, mudanças terminológicas e trajetórias. A IA pode apoiar a leitura de séries, mas é necessário controlar artefatos de busca, alterações de indexação, variação terminológica e atraso de citações. Cenários prospectivos devem ser formulados como hipóteses, não como previsões determinísticas.

1.7 Leitura, sumarização, classificação e visualização assistidas por IA

A escala da literatura torna inviável, em muitos estudos, a leitura manual integral de todos os documentos. A IA pode resumir conjuntos de resumos, classificar documentos, extrair métodos, objetos e populações, identificar lacunas e comparar grupos. Essa capacidade ajuda a conectar padrões quantitativos a conteúdos substantivos, mas sumarização automática não equivale a revisão sistemática. Estudos sobre factualidade mostram que resumos gerados podem introduzir informações não sustentadas pelo texto-fonte (MAYNEZ *et al.*, 2020).

A classificação temática assistida por IA deve ser iterativa: definir categorias, testar uma amostra, revisar ambiguidades e recalibrar classes excessivamente amplas ou estreitas. O

modelo pode indicar grau de confiança e justificar cada classificação com trechos do resumo; casos duvidosos devem ser encaminhados à revisão humana. O desenvolvimento de modelos como BERT ampliou a capacidade de classificação contextual, sem eliminar a necessidade de validação (DEVLIN *et al.*, 2019).

A visualização científica é simultaneamente etapa analítica e comunicacional. Gráficos, mapas de rede, diagramas de Sankey, mapas temáticos, linhas do tempo e *dashboards* organizam evidências e influenciam interpretações. Ferramentas como VOSviewer, CiteSpace e ambientes integrados de análise ampliaram a capacidade de representar estruturas bibliométricas (CHEN, 2006; VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017).

Visualizações devem manter coerência entre pergunta e forma gráfica; clareza de unidades e métricas; explicitação de limiares; distinção entre tamanho, cor, proximidade e força de relação; indicação da fonte; registro de parâmetros; e legenda que não ultrapasse o que o gráfico permite afirmar. A IA pode atuar como revisora de legibilidade, desde que receba dados e contexto suficientes.

A integração entre sumarização, classificação e visualização amplia a capacidade analítica, mas exige triangulação constante. Métricas, conteúdo textual, conhecimento especializado e literatura substantiva devem ser confrontados antes de qualquer conclusão.

1.8 Interpretação, escrita e comunicação de evidências

A interpretação cientométrica exige distinguir operações que frequentemente são confundidas. Descrever é apresentar padrões observados; interpretar é atribuir sentido a esses padrões; explicar é propor mecanismos; e recomendar é orientar ação. Um relatório pode registrar crescimento de produção ou centralidade em uma rede, mas não deve afirmar, sem evidências adicionais, que uma política causou o crescimento ou que centralidade equivale a liderança substantiva.

A IA pode desempenhar o papel de interlocutor crítico, revisando conclusões, identificando extrapolações, propondo explicações alternativas e perguntando que evidências adicionais seriam necessárias. Essa função não transforma o modelo em parecerista especializado: a qualidade da crítica depende do *prompt*, dos dados fornecidos e da capacidade do usuário de avaliar a resposta.

A escrita de relatórios, artigos e pareceres requer transparência metodológica. A metodologia deve informar pergunta, fontes, data de coleta, estratégia de busca, critérios, exportação, limpeza, indicadores, ferramentas, parâmetros, uso de IA e validação. Resultados devem apresentar padrões sem exagerar inferências; a discussão deve relacionar achados à literatura e às limitações; e o sumário executivo deve traduzir evidências sem apagar incertezas.

A comunicação precisa ser adaptada ao público. Pesquisadores demandam detalhes metodológicos; gestores, implicações e riscos; formuladores de políticas, evidências claras, limites de generalização e recomendações condicionais. *Prompts* específicos podem adaptar a linguagem, mas não devem simplificar indevidamente o conteúdo.

A responsabilidade autoral permanece humana. Textos gerados por modelos podem conter erros, omissões e referências fabricadas (JI *et al.*, 2023). A IA deve apoiar redação, revisão e organização, e não operar como fonte final de informação. Autores respondem pelo conteúdo, pelas evidências, pelas citações e pelas conclusões.

1.9 Questões epistemológicas, sociológicas e éticas

A cientometria não é neutra. Indicadores mostram determinados aspectos da ciência e ocultam outros; bases privilegiam idiomas, áreas, tipos documentais, periódicos e regiões; citações podem expressar reconhecimento, uso, crítica ou convenção; mapas e rankings produzem aparência de objetividade apesar de dependerem de decisões contestáveis. Diretrizes de métricas responsáveis procuram tornar essas escolhas visíveis e limitar usos avaliativos inadequados (HICKS *et al.*, 2015).

A cientometria baseada em publicações observa a parcela registrada, indexada e estabilizada da atividade científica. Não capta plenamente pesquisas em andamento, conhecimento tácito, trabalho técnico, formação de recursos humanos, impacto local, inovação pedagógica, extensão, colaboração informal ou produção em línguas menos indexadas. Comparações entre bases demonstram diferenças relevantes de cobertura, com consequências para avaliações institucionais e disciplinares (VISSER; VAN ECK; WALTMAN, 2021; MARTÍN-MARTÍN *et al.*, 2021).

A falsa neutralidade de mapas e rankings é particularmente crítica. Um ranking reorganiza reputações e incentivos; um mapa dá forma visual a relações condicionadas por algoritmos e parâmetros. Todo produto deve apresentar método, fonte, limitações e advertências de uso. A IA pode auxiliar a revisar implicações éticas, mas não deve suavizar conflitos ou incertezas.

A opacidade dos modelos de linguagem constitui problema metodológico. Respostas podem refletir vieses linguísticos, culturais, disciplinares e temporais, reproduzindo hierarquias de visibilidade e invisibilizando perspectivas periféricas. A literatura sobre alucinação e técnicas de *prompting* reforça a necessidade de validação, testes de robustez e registro do uso do modelo (JI *et al.*, 2023; SCHULHOFF *et al.*, 2024).

A transparência no uso de *prompts* é, portanto, exigência ética e metodológica. O leitor deve saber em que etapas a IA foi usada, quais dados recebeu, que instruções orientaram tarefas relevantes, como as respostas foram verificadas e quais decisões humanas prevaleceram. Essa documentação protege a confiabilidade sem transformar a pesquisa em burocracia.

1.10 Aplicações estratégicas e produtos do livro

A obra apresenta aplicações da cientometria com IA em diagnóstico de áreas, detecção de temas emergentes, avaliação de programas, planejamento institucional, estratégias de colaboração, inteligência acadêmica, prospecção tecnológica, organização do conhecimento e política científica. Em cada aplicação, resultados quantitativos são tratados como evidências parciais que precisam ser contextualizadas.

Na avaliação de programas de pós-graduação, a abordagem pode apoiar diagnósticos de produção, colaboração, inserção internacional, diversidade temática e alinhamento a áreas de concentração, sem substituir avaliação qualitativa. Em estratégias de colaboração, ajuda a identificar parceiros, pontes institucionais, lacunas de rede e complementaridades. Na prospecção tecnológica, articula literatura, patentes, temas emergentes e sinais de aplicação; na curadoria, apoia classificação de coleções e serviços de informação.

Um produto central é o conjunto de protocolos completos para análise de áreas científicas, autores e grupos, instituições, temas emergentes, colaboração internacional e relatórios de política científica. Cada protocolo integra objetivo, pergunta, fontes, busca, indicadores, *prompts*, ferramentas, validação, produto final e limitações.

Outro produto é a biblioteca de *prompts* cientométricos, organizada para formular perguntas, buscar literatura, limpar dados, selecionar indicadores, interpretar redes, nomear mapas temáticos, redigir resultados, revisar conclusões e auditar reprodutibilidade. A biblioteca é um repertório adaptável ao tema, à fonte, ao público, à finalidade e às limitações do estudo.

O impacto esperado é contribuir para uma competência emergente: combinar cientometria, inteligência artificial e responsabilidade metodológica. Essa competência é relevante para pesquisadores, gestores, bibliotecários, avaliadores, analistas de ciência e formuladores de política científica que trabalham com grandes volumes de informação e decisões baseadas em evidências.

1.11 Principais entregáveis conceituais e operacionais

A obra reúne contribuições conceituais e instrumentos operacionais que podem ser utilizados de forma integrada ou adaptados a necessidades específicas.

Quadro 1.4 - Entregáveis centrais da obra

Entregável	Descrição
Visão integrada	Articulação entre cientometria, bibliometria, análise de redes, mapas temáticos, indicadores e avaliação responsável.
IA metodologicamente orientada	Uso de *prompts* documentados, com versões, entradas, saídas, decisões humanas e critérios de validação.
Fluxo completo de trabalho	Pergunta, busca, coleta, limpeza, indicadores, redes, mapas, análise temporal, interpretação, redação e auditoria.
Protocolos aplicáveis	Modelos para áreas científicas, autores, grupos, instituições, temas emergentes, colaboração internacional e política científica.
Biblioteca de prompts	Instrumentos para formular perguntas, buscar, limpar, analisar, interpretar, redigir e auditar.
Arcabouço ético	Distinção entre medição, avaliação e interpretação, com prevenção de rankings acríticos, causalidade indevida e uso opaco de IA.

Fonte: Elaborado pelo autor (2026).

1.12 Conclusão executiva

O livro apresenta uma proposta contemporânea para estudar a ciência em ambiente de abundância informacional e mediação algorítmica. A obra reconhece o potencial da IA para ampliar escala, velocidade e sistematicidade, mas recusa a ideia de que esse potencial dispense método. Quanto maior a capacidade tecnológica, maior deve ser o compromisso com documentação, validação e responsabilidade.

Em perspectiva executiva, o livro responde a uma necessidade institucional concreta. Universidades, agências, programas de pós-graduação, bibliotecas e grupos de pesquisa precisam analisar grandes volumes de produção para decidir sobre prioridades, colaboração, avaliação, internacionalização e investimento. Ao mesmo tempo, precisam evitar que métricas e IA sejam usadas de forma simplificadora ou injusta. A obra constrói uma ponte entre capacidade técnica e responsabilidade analítica.

A mensagem central é que a cientometria com IA deve ser menos uma prática de automação e mais uma prática de inteligência metodológica. A IA amplia perguntas, organiza evidências e acelera sínteses; o pesquisador define sentido, valida resultados, reconhece limites e assume responsabilidade. O *prompt*, escrito, versionado, testado e auditado, torna-se o elo entre intenção científica e operação algorítmica. Essa abordagem permite produzir mapas e indicadores sem reduzir a ciência a números, comunicar evidências sem transformar tendências em certezas e construir diagnósticos com transparência, rigor e compromisso ético.

Capítulo 2

Introdução: por que Cientometria com Inteligência Artificial?

2.1 Introdução

Este capítulo introduz a problemática central da trilha "Cientometria com Inteligência Artificial: *Prompts* para Analisar, Mapear e Interpretar a Ciência". O objetivo é estabelecer as bases conceituais, metodológicas e críticas para compreender por que a análise da ciência passa por uma inflexão relevante quando modelos de inteligência artificial generativa, especialmente modelos de linguagem de grande escala, são incorporados aos fluxos de trabalho da cientometria. A questão não é meramente instrumental. Não se trata apenas de substituir uma ferramenta por outra, nem de acelerar operações já conhecidas. A mediação por IA altera a forma como perguntas são formuladas, como escopos são delimitados, como categorias são construídas, como evidências são sumarizadas e como interpretações são comunicadas. Assim, a entrada da IA no campo exige uma reflexão simultaneamente técnica, epistemológica e ética.

A cientometria pode ser compreendida como o estudo quantitativo, relacional e interpretativo da ciência, de suas publicações, de suas citações, de seus atores, de suas instituições, de seus temas e de seus fluxos de comunicação. Ela compartilha fronteiras com a bibliometria, a informetria, a webometria, a altmetria e a ciência da ciência. Desde os trabalhos clássicos de Garfield, Price, Merton, Small e outros autores, consolidou-se a ideia de que a comunicação científica deixa rastros mensuráveis e que tais rastros podem ser analisados para compreender a dinâmica da produção do conhecimento. A ciência, entretanto, não é redutível aos seus rastros documentais. Publicações, citações, palavras-chave, resumos, redes de coautoria e métricas de impacto são representações parciais de processos sociais e cognitivos mais amplos. A cientometria rigorosa, portanto, exige sempre a articulação entre mensuração, contextualização e interpretação (GARFIELD, 1955; PRICE, 1965; MERTON, 1968; SMALL, 1973; WHITE; GRIFFITH, 1981; FORTUNATO *et al.*, 2018).

A inteligência artificial, quando aplicada à cientometria, intensifica essa tensão entre representação e interpretação. Por um lado, amplia a capacidade de manipular grandes volumes de textos, metadados e descrições de redes. Modelos de linguagem podem apoiar a formulação de estratégias de busca, sugerir categorias temáticas, resumir conjuntos de resumos, rotular *clusters*, identificar possíveis lacunas de pesquisa, redigir descrições de indicadores, propor hipóteses interpretativas e auxiliar na comunicação dos resultados. Por outro lado, esses modelos podem produzir respostas plausíveis, mas factualmente incorretas; podem reforçar vieses existentes nos dados; podem ocultar incertezas; podem confundir inferência com evidência; e podem criar uma aparência de autoridade discursiva que não corresponde à robustez metodológica do estudo. A IA, portanto, deve ser incorporada como instrumento assistivo, auditável e subordinado ao desenho de pesquisa, não como instância autônoma de validação científica (ARIA; CUCCURULLO, 2017; CHEN, 2006; VAN ECK; WALTMAN, 2010; BENDER *et al.*, 2021; JI *et al.*, 2023; FARQUHAR *et al.*, 2024).

A proposta deste capítulo é defender que os *prompts* devem ser tratados como instrumentos de pesquisa. Em uso comum, *prompt* é frequentemente entendido como uma instrução dada a um sistema de IA. Em um contexto científico, porém, essa definição é

insuficiente. Um *prompt* pode funcionar como um protocolo micro analítico: explicita o papel atribuído ao modelo, delimita a tarefa, define os dados de entrada, estabelece restrições, indica o formato de saída, exige separação entre evidência e inferência e obriga o sistema a declarar incertezas. Quando registrado, versionado e avaliado, o *prompt* se aproxima de um procedimento metodológico. Ele não garante validade por si só, mas pode aumentar a transparência do processo de análise assistida por IA (CHEN *et al.*, 2025).

Este texto segue uma estrutura inspirada em sumários de obras de mapeamento científico, nas quais a introdução é seguida de definições, fluxos de trabalho, tipos de mapas, representação temporal, questões epistemológicas e aplicações. A diferença é que aqui a introdução concentra-se no deslocamento produzido pela IA e na necessidade de transformar a interação com modelos de linguagem em prática metodológica controlada. Desse modo, o capítulo prepara o leitor para os capítulos seguintes: fundamentos, engenharia de *prompts*, formulação de perguntas, estratégias de busca, coleta, pré-processamento, indicadores, redes, mapas temáticos, análise temporal, visualização, interpretação, redação, ética e aplicações (COBO *et al.*, 2011; ARIA; CUCCURULLO, 2017).

Ao longo do capítulo, a expressão "Cientometria com IA" designa o conjunto de práticas em que técnicas de inteligência artificial, especialmente modelos de linguagem, são usadas para apoiar etapas de um estudo cientométrico. A expressão não implica que toda análise se torne automatizada, nem que a IA produza resultados cientométricos sem intervenção humana. Pelo contrário, a noção de "com IA" enfatiza uma relação de coprodução controlada: a máquina auxilia tarefas de organização, síntese e inferência preliminar, enquanto o pesquisador define a pergunta, escolhe a fonte, audita os dados, interpreta os resultados e assume responsabilidade pelas conclusões.

Figura 2 - O fator humano e os limites da tecnologia

Fonte: Imagem gerada por IA, ChatGPT (2026).

2.2 O que muda quando a análise da ciência passa a ser mediada por IA

A primeira mudança diz respeito à escala operacional da análise. Estudos cientométricos tradicionais frequentemente envolvem grandes bases de dados, mas o processamento interpretativo de textos associados a esses dados permanecia limitado. Títulos, resumos

e palavras-chave podiam ser submetidos a técnicas de mineração de texto, análise de coocorrência, modelagem de tópicos ou classificação supervisionada. Entretanto, a leitura interpretativa, a elaboração de rótulos semânticos e a produção de narrativas analíticas demandavam tempo considerável e dependiam de equipes especializadas. A IA generativa amplia a possibilidade de interagir semanticamente com grandes conjuntos documentais, tornando mais acessível a exploração preliminar de campos científicos complexos (DONTHU *et al.*, 2021; ZUPIC; ČATER, 2015).

A segunda mudança está na passagem de interfaces predominantemente técnicas para interfaces conversacionais. Ferramentas clássicas de cientometria, como softwares de mapeamento bibliométrico, ambientes estatísticos e pacotes de programação, exigem conhecimento específico sobre formatos de dados, parâmetros de rede, normalização, algoritmos de *cluster*ização e visualização. Com modelos de linguagem, o pesquisador pode formular tarefas em linguagem natural: "identifique os temas recorrentes", "proponha rótulos para estes *clusters*", "explique a diferença entre acoplamento bibliográfico e cocitação para um relatório executivo", "indique riscos metodológicos desta estratégia de busca". Essa conversacionalidade reduz barreiras de entrada, mas também pode criar uma falsa sensação de simplicidade. A linguagem natural é mais acessível, porém menos precisa se não for cuidadosamente estruturada (COBO *et al.*, 2011; CHEN, 2006; VAN ECK; WALTMAN, 2010).

A terceira mudança refere-se ao deslocamento do trabalho analítico. Parte do esforço que antes se concentrava na execução manual de tarefas passa a ser dedicado à formulação de instruções, à avaliação de saídas e à documentação das decisões. Em vez de apenas "fazer" uma classificação temática, o pesquisador precisa decidir como pedir a classificação, quais critérios explicitar, quais exemplos fornecer, como lidar com ambiguidades, como medir consistência e como revisar casos duvidosos. A competência central deixa de ser somente o domínio de ferramentas e passa a incluir a capacidade de desenhar interações confiáveis com sistemas probabilísticos.

A quarta mudança diz respeito à temporalidade da pesquisa. Modelos de IA podem acelerar ciclos iterativos de exploração: uma pergunta inicial pode ser refinada, convertida em *string* de busca, testada contra um conjunto de resumos, revista em função de ruídos e reorganizada em poucos minutos. O mesmo vale para a redação de sínteses preliminares, a preparação de quadros comparativos e a adaptação de resultados para públicos distintos. Essa aceleração, contudo, não deve ser confundida com encurtamento do método científico. A velocidade da iteração aumenta a necessidade de controle. Quanto mais rapidamente uma hipótese é produzida, mais importante se torna distinguir hipótese, evidência e conclusão (CHEN *et al.*, 2025).

A quinta mudança recai sobre a interpretação dos mapas e indicadores. A cientometria sempre exigiu cautela interpretativa: uma rede de coautoria não equivale automaticamente a colaboração intelectual intensa; uma citação não expressa necessariamente concordância; um *cluster* não representa uma comunidade homogênea; um aumento de publicações pode refletir crescimento real do campo, mudança terminológica, expansão de bases de dados ou efeito de políticas de indexação. A IA pode ajudar a explicitar essas alternativas, mas também pode narrar relações espúrias de modo convincente. Por isso, a interpretação assistida por IA deve ser construída sob um princípio de prudência: o modelo pode sugerir explicações, mas o pesquisador deve avaliá-las contra o desenho do estudo, os dados, a literatura substantiva e o conhecimento do domínio (BORNMANN; DANIEL, 2008; SMALL, 1973; WALTMAN; VAN ECK; NOYONS, 2010).

A sexta mudança é a ampliação da camada textual do relatório cientométrico. Estudos cientométricos produzem tabelas, redes, gráficos, mapas e indicadores, mas também precisam

de narrativas interpretativas. A IA generativa é particularmente eficaz na produção textual, o que pode favorecer relatórios mais claros e sistemáticos. Entretanto, a fluência textual é um risco quando mascara incertezas. Um parágrafo bem escrito pode fazer uma interpretação frágil parecer sólida. Portanto, a mediação por IA exige uma ética da linguagem: os textos gerados ou revisados devem explicitar limites, pressupostos, escolhas metodológicas e graus de confiança (LUND; WANG, 2023; LIEBRENZ *et al.*, 2023).

A sétima mudança envolve a rastreabilidade. Em métodos tradicionais, scripts, planilhas, parâmetros e versões de software podem ser preservados para fins de reprodutibilidade. Na IA conversacional, parte do procedimento ocorre em interações textuais que podem se perder se não forem documentadas. O *prompt* usado para classificar resumos, a versão do modelo, a data da execução, os dados fornecidos, os critérios de revisão e as alterações humanas precisam ser registrados. Caso contrário, a análise torna-se difícil de auditar. A cientometria com IA não deve apenas produzir resultados; deve produzir também memória metodológica.

A oitava mudança é o aumento da dependência de decisões de curadoria. Modelos de linguagem dependem do contexto fornecido e podem responder de modo diferente conforme a seleção dos dados. Em uma análise de resumos, por exemplo, o modelo pode ser induzido a privilegiar certos termos se a amostra estiver desequilibrada. Em uma interpretação de rede, pode superestimar *clusters* maiores ou atribuir significado excessivo a rótulos automáticos. O pesquisador precisa curar o material de entrada, definir critérios de amostragem, separar dados brutos de dados derivados e estabelecer padrões de validação. A IA não elimina o trabalho de curadoria; ela o torna mais visível e mais decisivo (BENDER *et al.*, 2021; JI *et al.*, 2023; FARQUHAR *et al.*, 2024).

A nona mudança diz respeito à democratização relativa da análise. Pesquisadores sem formação avançada em programação podem usar IA para compreender comandos, gerar scripts, interpretar erros, sugerir fluxos de trabalho e produzir descrições iniciais. Esse potencial é importante para ampliar o acesso à cientometria, especialmente em instituições com menor infraestrutura analítica. Contudo, democratização não deve significar banalização. A facilidade de gerar uma resposta não equivale à validade da resposta. A formação cientométrica continua indispensável para compreender bases de dados, indicadores, normalização, vieses disciplinares e limites de inferência.

A décima mudança é o surgimento de novas vulnerabilidades. Além dos erros clássicos de cobertura de bases, duplicidade, desambiguação e normalização, entram em cena riscos específicos da IA: alucinação de referências, invenção de relações causais, simplificação de controvérsias, viés de confirmação induzido pelo *prompt*, instabilidade de saídas, dependência de modelos proprietários e dificuldade de explicar o processo interno de geração. A cientometria com IA deve, portanto, ampliar seu repertório de controle de qualidade. Não basta verificar os dados; é preciso verificar também a interação com o modelo.

Essas mudanças sugerem que a IA não é uma camada neutra adicionada a um processo preexistente. Ela interfere no modo como o pesquisador pergunta, seleciona, agrupa, interpreta e comunica. Essa interferência pode ser produtiva quando explicitada, controlada e validada. Pode ser problemática quando invisibilizada. O desafio é construir uma cientometria aumentada por IA que preserve os fundamentos da boa pesquisa: clareza de pergunta, adequação de método, transparência de procedimentos, proporcionalidade das conclusões e abertura à crítica.

2.3 Da bibliometria tradicional à cientometria aumentada por prompts

A bibliometria tradicional consolidou um repertório metodológico baseado na contagem e na análise estatística de documentos científicos. Entre seus objetos clássicos estão a produtividade de autores, periódicos, instituições e países; a frequência de citações; os padrões de coautoria; a dispersão de publicações; a obsolescência da literatura; e a estrutura de campos por meio de cocitação, acoplamento bibliográfico ou coocorrência de termos. Ao longo do tempo, esse repertório expandiu-se para uma cientometria mais ampla, interessada não apenas em contar documentos, mas em compreender a organização intelectual, social e temporal da ciência (GARFIELD, 1955; PRICE, 1965; FORTUNATO *et al.*, 2018; DONTHU *et al.*, 2021; ZUPIC; ČATER, 2015).

A evolução da bibliometria para a cientometria envolveu uma mudança de ênfase: do indicador isolado para a estrutura relacional; da contagem para o mapa; do retrato estático para a dinâmica; da produtividade para a circulação de conhecimento; da avaliação simples para a interpretação contextual. Essa evolução pode ser observada na emergência de softwares e métodos de mapeamento científico, nos quais redes de citações, palavras-chave, autores ou instituições são visualizadas como paisagens do conhecimento. Tais mapas não são representações diretas da realidade científica; são construções baseadas em escolhas de dados, algoritmos, normalizações e parâmetros (VAN ECK; WALTMAN, 2010; WALTMAN; VAN ECK; NOYONS, 2010; ARIA; CUCCURULLO, 2017).

A cientometria aumentada por *prompts* insere-se nessa trajetória. Ela não substitui os métodos de rede, os indicadores ou as ferramentas de visualização. Ao contrário, atua como uma camada complementar que auxilia o pesquisador a planejar, interpretar e comunicar essas análises. O *prompt* torna-se uma interface entre a intencionalidade do pesquisador e as capacidades de um sistema de IA. Com ele, pode-se solicitar a expansão de termos de busca, a identificação de possíveis sinônimos, a elaboração de hipóteses para *clusters*, a comparação entre indicadores, a redação de limitações ou a construção de um resumo executivo.

É importante observar que a expressão "aumentada" não significa "melhor" de modo automático. Aumentar pode significar ampliar capacidades, mas também ampliar erros. Um estudo com IA pode ser mais rápido, mas não necessariamente mais rigoroso. Pode produzir relatórios mais sofisticados, mas não necessariamente mais verdadeiros. Pode identificar padrões textuais, mas não necessariamente compreender o contexto epistemológico de um campo. A cientometria aumentada por *prompts* deve ser avaliada pelo ganho de transparência, pela qualidade da validação e pela coerência entre pergunta, método e interpretação.

Na bibliometria tradicional, muitos procedimentos são parametrizáveis: escolher uma janela temporal, definir um limiar mínimo de citações, selecionar um algoritmo de *cluster*ização, aplicar uma medida de normalização, remover duplicatas. Na cientometria com IA, os *prompts* adicionam uma nova camada de parametrização semântica. Ao pedir que o modelo classifique artigos, por exemplo, o pesquisador precisa definir o que conta como categoria temática, que nível de granularidade é desejado, se categorias podem se sobrepor, como tratar artigos interdisciplinares e como registrar incerteza. Esses parâmetros nem sempre aparecem como números; aparecem como instruções linguísticas. Justamente por isso, precisam ser escritos com precisão.

A transição para *prompts* também transforma a relação entre exploração e confirmação. Em etapas exploratórias, a IA pode ajudar a descobrir vocabulários, subtemas e possíveis relações. Em etapas confirmatórias, deve-se evitar que o modelo fabrique confirmação para hipóteses frágeis. A diferença entre exploração e confirmação deve ser marcada explicitamente no protocolo. Um *prompt* exploratório pode pedir: "Liste hipóteses possíveis sem assumir

que sejam verdadeiras". Um *prompt* confirmatório deve exigir: "Avalie se os dados fornecidos sustentam ou não a hipótese, indicando evidências contrárias". Essa distinção é fundamental para impedir que a IA funcione como mecanismo de reforço de expectativas.

A cientometria aumentada por *prompts* também altera a relação entre especialistas de domínio e analistas de dados. Tradicionalmente, um estudo cientométrico robusto beneficia-se da interação entre quem domina os métodos quantitativos e quem conhece a área analisada. Com IA, essa interação pode ser mediada por saídas preliminares que ajudam ambos a dialogar. O modelo pode gerar descrições alternativas de um *cluster*, perguntas para validação por especialistas e sínteses de termos recorrentes. No entanto, o especialista de domínio continua indispensável para avaliar se os rótulos fazem sentido, se a interpretação respeita a história do campo e se a terminologia está sendo usada corretamente.

Outra característica da abordagem por *prompts* é a possibilidade de modularizar o trabalho analítico. Em vez de pedir ao modelo uma interpretação global de um estudo, o pesquisador pode decompor a tarefa em etapas: primeiro, verificar a clareza da pergunta; depois, avaliar a estratégia de busca; em seguida, agrupar termos; posteriormente, nomear *clusters*; por fim, revisar a coerência do relatório. Essa decomposição reduz a chance de respostas excessivamente gerais e aumenta a auditabilidade. Cada *prompt* pode ser associado a uma etapa, a um insumo e a uma saída específica.

A modularização também facilita a comparação de saídas. O mesmo conjunto de resumos pode ser submetido a *prompts* diferentes, ou o mesmo *prompt* pode ser aplicado a amostras distintas. A consistência entre respostas pode ser avaliada. Divergências podem indicar ambiguidade nos dados, inadequação das categorias ou instabilidade do modelo. Em vez de tratar a variação como ruído indesejado, o pesquisador pode usá-la como sinal de que determinada classificação exige revisão humana. Em cientometria, a incerteza não deve ser eliminada artificialmente; deve ser documentada (CHEN *et al.*, 2025; BENDER *et al.*, 2021).

Um ponto central é que *prompts* não substituem scripts, bancos de dados ou métodos estatísticos. Eles interagem com esses elementos. Em um fluxo robusto, a extração de metadados, a construção de redes e o cálculo de indicadores podem continuar sendo realizados por softwares especializados, enquanto a IA auxilia tarefas de planejamento, tradução metodológica, revisão crítica e interpretação textual. A melhor cientometria com IA provavelmente será híbrida: quantitativa e qualitativa, algorítmica e interpretativa, automatizada e curada, conversacional e documentada.

A passagem da bibliometria tradicional à cientometria aumentada por *prompts*, portanto, não deve ser lida como ruptura total, mas como reconfiguração. Os fundamentos permanecem: qualidade dos dados, adequação dos indicadores, interpretação contextual, atenção aos vieses e transparência. O que muda é a presença de um interlocutor artificial capaz de produzir linguagem, categorias e hipóteses. Esse interlocutor pode enriquecer a análise quando tratado como ferramenta crítica. Pode empobrecê-la quando tratado como autoridade.

2.4 Prompts como instrumentos de pesquisa, não apenas comandos

Um comando é uma solicitação pontual. Um instrumento de pesquisa é um procedimento inscrito em um desenho metodológico. A diferença é decisiva. Quando um pesquisador escreve "resuma estes artigos", está emitindo um comando. Quando escreve um *prompt* que especifica objetivo, *corpus*, critérios de seleção, nível de abstração, formato de saída, limites de interpretação, necessidade de indicar incerteza e proibição de inventar dados, está construindo um instrumento metodológico. O mesmo ato linguístico muda de estatuto quando passa a fazer parte de um protocolo documentado e avaliável (CHEN *et al.*, 2025).

Tratar *prompts* como instrumentos de pesquisa significa reconhecer que eles influenciam os resultados. Em modelos de linguagem, pequenas mudanças de formulação podem alterar a resposta. Uma instrução que pede "identifique os temas mais importantes" pode produzir uma lista diferente de outra que pede "identifique os temas mais frequentes nos resumos fornecidos". A palavra "importantes" abre espaço para julgamento valorativo; a palavra "frequentes" orienta a resposta para evidência textual. De modo semelhante, pedir "explique por que o tema cresceu" pressupõe crescimento e causalidade; pedir "liste hipóteses possíveis para o crescimento observado, indicando evidências necessárias" é metodologicamente mais prudente.

Um *prompt* cientométrico robusto deve conter, no mínimo, sete componentes: papel atribuído ao modelo, objetivo da tarefa, descrição dos dados de entrada, critérios de decisão, restrições, formato de saída e exigência de incerteza. O papel atribuído orienta o enquadramento: "atue como analista cientométrico". O objetivo define a tarefa: "avaliar a adequação de uma estratégia de busca". A descrição dos dados delimita o *corpus*: "lista de títulos e resumos extraídos da base X". Os critérios de decisão tornam a avaliação mais objetiva: "considere precisão terminológica, cobertura temática e risco de ruído". As restrições previnem extrapolações: "não invente autores, periódicos ou estatísticas". O formato de saída facilita revisão: "apresente uma tabela com problema, evidência, consequência e sugestão". A exigência de incerteza impede respostas excessivamente confiantes: "classifique cada recomendação como alta, média ou baixa confiança" (CHEN *et al.*, 2025).

Além desses elementos, *prompts* científicos devem explicitar a diferença entre dado fornecido e conhecimento externo. Em muitas interações, o modelo mistura informações presentes no *corpus* com conhecimento aprendido durante seu treinamento. Essa mistura pode ser útil, mas precisa ser controlada. Um *prompt* adequado pode solicitar: "Separe sua resposta em duas partes: observações baseadas exclusivamente nos dados fornecidos e contextualizações baseadas em conhecimento geral, indicando que estas últimas precisam de verificação bibliográfica". Essa separação é especialmente importante em cientometria, porque a interpretação de um campo depende tanto dos metadados analisados quanto do contexto disciplinar.

Prompts também podem funcionar como mecanismos de autocontenção. Ao pedir que o modelo "não conclua além dos dados", "aponte limitações", "liste interpretações alternativas" e "indique o que seria necessário para confirmar a hipótese", o pesquisador induz uma resposta menos dogmática. Evidentemente, tais instruções não garantem ausência de erro. Modelos podem obedecer de forma incompleta ou superficial. Ainda assim, *prompts* bem desenhados reduzem o risco de saídas indevidas e facilitam a revisão humana, pois tornam explícito o padrão esperado de resposta.

Um aspecto metodológico relevante é a versionabilidade. *Prompts* devem ser salvos com identificação, data, finalidade e versão. Um estudo pode usar, por exemplo, "*Prompt* 2.1 - Expansão de vocabulário de busca", "*Prompt* 4.3 - Agrupamento de palavras-chave", "*Prompt* 7.2 - Interpretação de *clusters*". Quando o *prompt* é alterado, a modificação deve ser registrada. Essa prática permite que outro pesquisador compreenda como a análise foi conduzida. Embora a reprodução exata de saídas de IA possa ser difícil, especialmente em modelos proprietários e atualizados continuamente, a documentação dos *prompts* melhora a transparência e torna o procedimento criticável.

A criticabilidade é um critério essencial. Um método científico precisa poder ser examinado por terceiros. Se a classificação temática de artigos foi feita por IA, o leitor deve saber quais dados foram oferecidos, que instruções foram usadas, como casos ambíguos foram tratados e que revisão humana foi realizada. Sem isso, a IA torna-se uma caixa-preta

metodológica adicional. Com documentação adequada, o *prompt* pode ser analisado como se analisa uma estratégia de busca, um código de classificação ou uma regra de normalização.

Também é necessário diferenciar *prompts* de produção e *prompts* de validação. *Prompts* de produção geram resultados preliminares: categorias, rótulos, sínteses, hipóteses, redações. *Prompts* de validação interrogam esses resultados: "quais afirmações dependem de evidência não apresentada?", "quais categorias se sobrepõem?", "que vieses podem ter sido introduzidos?", "que conclusões parecem fortes demais?". Um fluxo cientométrico maduro deve incluir ambos. Usar IA apenas para produzir é insuficiente; é preciso usá-la também como instrumento de crítica, sem delegar a ela a decisão final.

Prompts podem ainda ser desenhados para favorecer triangulação. Um mesmo resultado pode ser analisado de três perspectivas: estatística, semântica e sociológica. Por exemplo, um *cluster* de coocorrência de palavras-chave pode ser interpretado com base em sua densidade interna, seus termos mais frequentes e sua relação com instituições ou países predominantes. Um *prompt* pode pedir que o modelo construa essa triangulação, mas também que indique onde faltam dados. Dessa forma, a IA ajuda a organizar a interpretação sem substituir a análise substantiva (DONTHU *et al.*, 2021; ZUPIC; ČATER, 2015).

Outra função dos *prompts* é a tradução entre públicos. A cientometria produz resultados úteis para comunidades diversas: pesquisadores, bibliotecários, pró-reitores, gestores de inovação, agências de fomento, formuladores de política científica e público geral. A IA pode adaptar a linguagem de um mesmo achado para diferentes públicos. Contudo, adaptação não deve significar distorção. Um *prompt* responsável deve exigir que o modelo mantenha a mesma base factual, explicite simplificações e evite transformar correlação em recomendação normativa. A comunicação científica assistida por IA deve preservar a integridade do resultado.

Prompts como instrumentos de pesquisa exigem, por fim, uma mudança de postura. O pesquisador não "pergunta à IA" como quem consulta uma autoridade. Ele constrói uma situação analítica, fornece insumos, delimita condições, avalia saídas e integra os resultados ao método. A qualidade da resposta depende da qualidade da pergunta, dos dados, das restrições e da revisão. A IA é parte do arranjo metodológico; não é sua garantia.

2.5 O papel humano: curadoria, validação, interpretação e responsabilidade

A incorporação de IA à cientometria não reduz a responsabilidade do pesquisador. Ao contrário, a amplia. Em fluxos assistidos por modelos de linguagem, o trabalho humano desloca-se para quatro funções centrais: curadoria, validação, interpretação e responsabilidade. Essas funções são interdependentes. A curadoria define o que entra no sistema; a validação examina o que sai; a interpretação conecta resultados a perguntas e contextos; a responsabilidade assume publicamente as escolhas, limitações e consequências do estudo.

A curadoria começa antes da IA. Ela envolve a escolha das bases de dados, a definição do recorte temporal, a construção da estratégia de busca, o estabelecimento de critérios de inclusão e exclusão, a remoção de duplicatas, a padronização de nomes e a seleção de campos bibliográficos. Se a base é incompleta, se a estratégia de busca é enviesada ou se os registros estão mal tratados, a IA apenas processará problemas com maior fluência. Não há *prompt* capaz de compensar integralmente uma coleta mal desenhada. Em cientometria, a qualidade da interpretação depende da qualidade do *corpus*.

A curadoria também envolve decidir que dados podem ser oferecidos ao modelo. Questões de privacidade, direitos autorais, sigilo institucional e sensibilidade estratégica podem limitar o uso de determinadas plataformas. Estudos que analisam dados internos de programas

de pós-graduação, avaliações institucionais ou estratégias de colaboração devem considerar se é adequado inserir tais informações em sistemas externos. A cientometria com IA exige, portanto, uma governança dos dados: o pesquisador deve saber que material está usando, onde ele será processado, que riscos existem e que alternativas locais ou anonimizadas podem ser adotadas.

A validação é a segunda função. Toda saída de IA deve ser tratada como hipótese de trabalho, não como resultado definitivo. Se o modelo agrupa palavras-chave, os agrupamentos devem ser revisados. Se nomeia *clusters*, os rótulos devem ser confrontados com os termos e documentos mais representativos. Se resume tendências, a síntese deve ser comparada com séries temporais e evidência documental. Se sugere riscos metodológicos, estes devem ser avaliados à luz do protocolo real. A validação pode combinar revisão humana, amostragem, comparação entre modelos, repetição com *prompts* alternativos e checagem contra fontes primárias (LUND; WANG, 2023; LIEBRENZ *et al.*, 2023).

A validação deve ser especialmente rigorosa quando a IA lida com referências bibliográficas. Modelos de linguagem podem inventar títulos, autores, periódicos, volumes, páginas e DOI. Em um campo como a cientometria, no qual a referência é simultaneamente objeto e evidência, esse risco é grave. Nenhuma referência gerada por IA deve ser incluída sem verificação em fonte confiável. Da mesma forma, citações textuais, dados estatísticos e afirmações sobre publicações devem ser conferidos. A fluência bibliográfica da IA pode produzir uma aparência de erudição que precisa ser sistematicamente auditada.

A interpretação é a terceira função humana. Indicadores e mapas não falam por si. Uma rede densa pode expressar colaboração consolidada, mas também concentração institucional; um crescimento de publicações pode indicar emergência temática, mas também ampliação de indexação; baixa citação pode revelar irrelevância, mas também juventude do campo, idioma periférico ou orientação aplicada. A IA pode listar alternativas interpretativas, mas a escolha entre elas exige conhecimento do domínio, leitura da literatura e compreensão das condições sociais de produção científica. O pesquisador deve resistir à tentação de aceitar a narrativa mais elegante como a mais plausível.

A responsabilidade é a quarta função. O uso de IA não transfere autoria, responsabilidade ética ou responsabilidade metodológica para a ferramenta. O pesquisador continua responsável por erros, omissões, vieses e conclusões. Em artigos e relatórios, o uso de IA deve ser descrito de forma transparente quando afetar etapas substantivas do trabalho. Deve-se indicar se a IA foi usada para apoio à redação, classificação temática, sumarização de resumos, geração de código, revisão de estratégias de busca ou interpretação preliminar. Transparência não é formalidade; é condição para que leitores avaliem a robustez do estudo.

A responsabilidade inclui reconhecer assimetrias. Modelos de IA são treinados em corpora que refletem desigualdades linguísticas, geográficas, disciplinares e culturais. Em cientometria, essas desigualdades podem se somar aos vieses das próprias bases bibliográficas, que frequentemente favorecem literatura em inglês, periódicos indexados internacionalmente e áreas com maior tradição de publicação em artigos. O uso acrítico de IA pode reforçar a invisibilidade de literatura local, regional, interdisciplinar ou publicada em formatos menos valorizados. O pesquisador deve perguntar não apenas o que o modelo vê, mas o que ele tende a não ver.

A dimensão humana também aparece na negociação de categorias. Nenhum campo científico se organiza de modo perfeitamente estável. Termos mudam, conceitos disputam significados, comunidades usam vocabulários distintos e temas emergem em zonas de fronteira. A IA pode produzir categorias aparentemente limpas, mas a realidade científica é

frequentemente ambígua. A revisão humana deve preservar essa ambiguidade quando ela é substantiva. Forçar todo documento a uma categoria exclusiva pode ser conveniente para visualização, mas inadequado para compreender campos interdisciplinares (BORNMANN; DANIEL, 2008; MERTON, 1968).

Outro ponto importante é a responsabilidade sobre a automação. À medida que modelos se tornam capazes de gerar código, scripts e relatórios, aumenta a tentação de construir *pipelines* totalmente automatizados. A automação pode ser útil para tarefas repetitivas, mas deve haver pontos de controle humano, especialmente em decisões que afetam inclusão/exclusão de documentos, rótulos interpretativos, métricas avaliativas e recomendações institucionais. A cientometria pode influenciar políticas de financiamento, avaliação de pesquisadores e estratégias organizacionais. Por isso, seus resultados não devem ser produzidos por automação opaca.

A relação entre humano e IA deve ser vista como uma relação de supervisão crítica. O modelo pode ampliar possibilidades, mas o pesquisador deve formular o problema, delimitar a evidência, testar alternativas, reconhecer limites e assumir conclusões. A autoridade científica não deriva da capacidade de gerar texto; deriva da capacidade de sustentar uma afirmação com método, dados e argumentação. A IA pode auxiliar a argumentação, mas não substitui a responsabilidade epistêmica do autor (BENDER *et al.*, 2021; LUND; WANG, 2023).

2.6 Limites da IA generativa em estudos científicos

O primeiro limite da IA generativa em estudos científicos é a possibilidade de alucinação. O termo designa saídas que parecem plausíveis, mas são falsas, imprecisas ou não sustentadas pelos dados. Em cientometria, a alucinação pode assumir várias formas: criação de referências inexistentes, atribuição equivocada de autoria, invenção de estatísticas, descrição incorreta de métodos, identificação falsa de tendências ou interpretação causal sem evidência. Como modelos de linguagem são treinados para produzir continuidade textual provável, e não para verificar a verdade de cada afirmação, sua fluência não deve ser confundida com confiabilidade factual.

O segundo limite é a instabilidade. A mesma pergunta pode gerar respostas diferentes em momentos distintos, especialmente quando parâmetros de geração, versão do modelo ou contexto de entrada variam. Essa instabilidade é problemática para a reprodutibilidade. Em estudos científicos, procedimentos devem ser suficientemente documentados para que outros possam compreender e, quando possível, reproduzir ou auditar os resultados. Na IA generativa, a reprodução exata pode ser difícil, mas a documentação de *prompts*, dados de entrada, versões e critérios de validação reduz a opacidade.

O terceiro limite é a dependência do contexto. Modelos respondem ao que recebem. Se o pesquisador fornece uma amostra enviesada de resumos, a síntese será enviesada. Se omite a estratégia de busca, o modelo não poderá avaliar sua cobertura. Se apresenta *clusters* sem indicar seus termos mais relevantes, a rotulação será especulativa. A IA não possui acesso mágico ao desenho metodológico. Ela precisa de insumos claros e suficientes. Quando faltam informações, um modelo responsável deveria declarar insuficiência; porém, modelos frequentemente preenchem lacunas com inferências plausíveis. Cabe ao *prompt* exigir que lacunas sejam explicitadas.

O quarto limite é a dificuldade de lidar com autoridade disciplinar. Modelos de linguagem podem reconhecer padrões gerais, mas podem falhar em distinções finas de áreas específicas. Em campos científicos especializados, termos semelhantes podem ter significados diferentes, e conceitos podem estar inseridos em tradições teóricas distintas. Um modelo pode agrupar

termos por proximidade lexical, ignorando divergências epistemológicas. Também pode tratar controvérsias como consensos, ou consensos como debates abertos. A revisão por especialistas de domínio continua indispensável.

O quinto limite é o viés de treinamento e de cobertura. A literatura disponível para treinamento de modelos e as bases bibliográficas usadas em estudos cientométricos não representam a ciência de modo equitativo. Há vieses de idioma, país, disciplina, acesso aberto, prestígio editorial e presença digital. Uma análise de ciência global mediada por IA pode reproduzir a centralidade de países, instituições e línguas já dominantes. Sem desenho metodológico atento, a IA pode naturalizar desigualdades estruturais como se fossem padrões neutros de excelência ou relevância.

O sexto limite é a opacidade técnica. Muitos modelos de linguagem são sistemas proprietários, com detalhes incompletos sobre dados de treinamento, atualizações, filtros, parâmetros e mecanismos de alinhamento. Essa opacidade contrasta com princípios de reprodutibilidade e transparência científica. Mesmo quando o pesquisador documenta o *prompt*, pode não saber exatamente como o modelo produziu a resposta. Por isso, a IA deve ser usada de preferência em tarefas nas quais a saída possa ser verificada externamente: classificação revisável, síntese confrontável com textos, geração de rascunhos, identificação de riscos e produção de hipóteses (BENDER *et al.*, 2021; JI *et al.*, 2023; FARQUHAR *et al.*, 2024).

O sétimo limite é a sedução da completude. Modelos de linguagem tendem a responder de forma completa, mesmo quando a pergunta é incompleta. Em estudos cientométricos, isso pode levar a relatórios que parecem finais, mas foram construídos sobre dados insuficientes. Um *prompt* crítico deve incentivar a recusa parcial: "se os dados forem insuficientes, declare insuficiência e indique quais dados faltam". A capacidade de não responder, ou de responder com incerteza, é uma característica desejável em contextos científicos.

O oitavo limite é a confusão entre análise textual e análise bibliométrica. A IA generativa é forte em linguagem, mas nem toda conclusão cientométrica pode ser derivada de linguagem. Indicadores de citação, redes de colaboração, métricas normalizadas e padrões temporais exigem dados estruturados e cálculos adequados. Pedir a um modelo que "avalie o impacto" de um conjunto de artigos sem fornecer citações, datas, campos e normalização é metodologicamente inadequado. A IA pode ajudar a interpretar indicadores, mas não deve inventá-los.

O nono limite é a possibilidade de retroalimentação e homogeneização textual. Se muitos pesquisadores usam IA para gerar títulos, resumos, palavras-chave e revisões, a própria literatura científica pode tornar-se mais padronizada em estilo e vocabulário. Isso pode afetar análises futuras de coocorrência de termos, detecção de temas e mineração de textos. A cientometria com IA deve estar atenta a esse fenômeno: a IA não apenas analisa a literatura; ela pode passar a moldar a literatura que será analisada.

O décimo limite é ético e institucional. Estudos cientométricos podem ser usados para avaliação, ranqueamento, alocação de recursos e formulação de políticas. Se a IA participa da produção de diagnósticos, recomendações ou classificações, há risco de automatizar decisões com impacto sobre pessoas e instituições. A prudência exige que resultados assistidos por IA sejam apresentados como subsídios, não como veredictos. A cientometria deve informar decisões, não substituí-las por autoridade algorítmica (LUND; WANG, 2023; LIEBRENZ *et al.*, 2023).

Esses limites não significam que a IA deva ser excluída da cientometria. Significam que sua incorporação deve ser metodologicamente disciplinada. A IA é útil quando ajuda a explicitar

perguntas, organizar informações, sugerir alternativas, melhorar a comunicação e ampliar a capacidade exploratória. É perigosa quando substitui verificação, validação e julgamento. A fronteira entre uso produtivo e uso problemático não está na ferramenta em si, mas no modo como ela é integrada ao desenho de pesquisa.

2.7 Prompts do capítulo

Os *prompts* a seguir são apresentados como modelos iniciais. Eles devem ser adaptados conforme o tema, a base de dados, o tipo de análise e o público do estudo. Em todos os casos, recomenda-se registrar a versão do *prompt*, a data de uso, o modelo utilizado, os dados fornecidos e as decisões humanas tomadas após a saída.

Prompt 2.7.1 Transformar uma ideia vaga em problema cientométrico

Atue como especialista em Cientometria e metodologia de pesquisa.

Tenho uma ideia inicial de estudo: [DESCREVA A IDEIA EM LINGUAGEM LIVRE].

Transforme essa ideia em um problema cientométrico operacionalizável.

Apresente:

(1) tema geral;

(2) possível pergunta principal;

(3) três perguntas secundárias;

(4) unidade de análise;

(5) fontes de dados possíveis;

(6) período temporal recomendado;

(7) indicadores adequados;

(8) tipos de rede ou mapa que poderiam ser usados;

(9) limitações previsíveis;

(10) dados que ainda preciso definir antes de iniciar o estudo.

Não invente resultados. Quando fizer suposições, marque-as como hipóteses.

Este *prompt* é útil no início do projeto, quando o pesquisador ainda possui uma formulação ampla. Sua função não é substituir o desenho de pesquisa, mas gerar uma primeira estrutura de decisão. A saída deve ser revisada criticamente, pois o modelo pode sugerir perguntas amplas demais, bases de dados inadequadas ou indicadores que não respondem à pergunta. O pesquisador deve verificar se a pergunta proposta é empiricamente observável e se os dados necessários estão disponíveis.

Prompt 2.7.2 Definir escopo de estudo

> Atue como analista cientométrico.
>
> Quero estudar [TEMA] em [ÁREA OU CAMPO].
>
> Ajude-me a definir o escopo do estudo.
>
> Considere: delimitação conceitual do tema, termos relacionados, termos que devem ser excluídos, período temporal, tipos de documento, idiomas, bases de dados, critérios de inclusão e exclusão, possíveis vieses de cobertura e justificativa metodológica para cada decisão.
>
> Organize a resposta em uma tabela com quatro colunas: decisão de escopo, justificativa, risco associado e forma de mitigação.

Este *prompt* transforma o escopo em objeto explícito de análise. Em estudos cientométricos, muitos erros surgem de recortes mal definidos. Um tema pode ter variações terminológicas, usos disciplinares distintos e fronteiras instáveis. A IA pode ajudar a mapear essas fronteiras, mas a decisão final deve considerar conhecimento do domínio e testes exploratórios em bases bibliográficas. O escopo não deve ser decidido apenas por conveniência; deve ser coerente com a pergunta e com a finalidade do estudo.

Prompt 2.7.3 Identificar riscos metodológicos iniciais

> Atue como revisor metodológico crítico de um estudo cientométrico.
>
> Analise o seguinte plano preliminar: [INSIRA OBJETIVO, PERGUNTA, BASES, PERÍODO, ESTRATÉGIA DE BUSCA, INDICADORES E ANÁLISES PREVISTAS].
>
> Identifique riscos metodológicos iniciais em cinco dimensões:
>
> (1) formulação da pergunta;
>
> (2) cobertura e viés da base de dados;
>
> (3) estratégia de busca;
>
> (4) escolha de indicadores e redes;
>
> (5) interpretação e comunicação dos resultados.
>
> Para cada risco, informe gravidade, evidência no plano, possível consequência e ação corretiva.
>
> Não presuma informações ausentes; indique quando faltar dado para avaliação.

Esse *prompt* deve ser usado antes da coleta definitiva, pois muitos problemas são mais baratos de corrigir no início do que depois da análise. Ele orienta a IA a atuar como interlocutor crítico e não como gerador de confirmação. A exigência de gravidade, evidência, consequência e ação corretiva torna a saída mais auditável. Ainda assim, o pesquisador deve revisar cada risco identificado, descartar observações genéricas e incorporar apenas recomendações justificadas.

2.8 Quadro de aprofundamento: princípios operacionais para o uso de IA em cientometria

Quadro 2.1 - Princípios operacionais para o uso de IA em cientometria

Princípio	Explicação	Boas Práticas (Exemplos)	Riscos a Evitar	Referências Relacionadas
1º Princípio da finalidade explícita	Antes de usar IA, o pesquisador deve saber que etapa pretende apoiar. A pergunta "como a IA pode ajudar?" é menos produtiva do que "em que tarefa específica a IA pode reduzir esforço sem reduzir controle?".	• Definir objetivo da tarefa e produto esperado. • Alinhar a tarefa à pergunta de pesquisa. • Selecionar apenas tarefas onde a IA agrega valor real. • Estabelecer critérios de sucesso da tarefa.	• Delegar decisões metodológicas críticas à IA. • Gerar estatísticas, indicadores ou dados inexistentes. • Validar causalidades sem base empírica.	Donthu *et al.*, 2021; Zupic; Čater, 2015.
2º Princípio do dado delimitado	A IA deve trabalhar com dados claramente especificados. A distinção entre *corpus* (dados do estudo) e contexto (conhecimento externo) evita misturas não verificáveis.	• Informar exatamente quais dados serão usados. • Indicar o que pode ser usado como contexto externo. • Separar evidências do *corpus* e informações externas na resposta. • Registrar limites do uso de contexto geral.	• Misturar *corpus* com conhecimento externo sem sinalização. • Inserir informações não verificadas como se fossem do estudo. • Falta de rastreabilidade das fontes utilizadas.	Chen *et al.*, 2025.
3º Princípio da granularidade adequada	*Prompts* muito amplos tendem a gerar respostas genéricas. A decomposição da tarefa em passos menores permite avaliar cada saída e diminuir riscos de erros propagados.	• Decompor a tarefa em etapas específicas. • Definir entrada, ação e saída esperada de cada passo. • Solicitar formatos estruturados (tabelas, listas). • Especificar extensão e nível de detalhamento.	• *Prompts* vagos e tarefas excessivamente amplas. • Saídas genéricas e pouco úteis. • Erros que se acumulam ao longo do processo.	Chen *et al.*, 2025.
4º Princípio da validação proporcional ao risco	Nem toda saída exige o mesmo nível de verificação. Quanto maior o impacto da saída, maior deve ser a intensidade da validação humana.	• Classificar a saída quanto ao impacto potencial. • Definir nível de verificação adequado. • Usar amostras de validação, avaliação de concordância e revisão por pares quando necessário. • Registrar a estratégia de validação aplicada.	• Aceitar saídas críticas sem revisão adequada. • Não ajustar a validação ao impacto da decisão. • Superestimar ou subestimar o risco envolvido.	Donthu *et al.*, 2021; Farquhar *et al.*, 2024.

Princípio	Explicação	Boas Práticas (Exemplos)	Riscos a Evitar	Referências Relacionadas
5º Princípio da explicitação da incerteza	Modelos de linguagem tendem a responder de forma assertiva. O *prompt* deve exigir níveis de confiança, ambiguidades e distinção entre afirmações e hipóteses.	• Solicitar níveis de confiança (alto, médio, baixo). • Pedir que ambiguidades e lacunas sejam listadas. • Separar fatos, inferências e hipóteses. • Discutir limitações de dados, métodos e cobertura.	• Apresentar hipóteses como fatos. • Ocultar limitações e ambiguidades. • Geração de falsa sensação de certeza.	Farquhar *et al.*, 2024.
6º Princípio da triangulação	Uma interpretação cientométrica torna-se mais forte quando combina diferentes tipos de evidências e conhecimentos.	• Combinar indicadores, redes, leitura de documentos e contexto. • Consultar especialistas e literatura do campo. • Confrontar padrões observados com outras fontes e métodos. • Registrar convergências e divergências.	• Basear-se em uma única métrica ou evidência. • Transformar associação em causalidade. • Ignorar contexto histórico e disciplinar.	Donthu *et al.*, 2021; Zupic; Čater, 2015.
7º Princípio da documentação	*Prompts*, saídas, revisões e decisões devem ser preservados para garantir transparência, rastreabilidade e possibilidade de auditoria.	Registrar *prompts* finais e datas de uso. Guardar entradas fornecidas e saídas obtidas. Documentar revisões humanas e justificativas. Incluir anexos metodológicos no relatório.	• Falta de registro das decisões. • Impossibilidade de reproduzir ou auditar o processo. • Perda de rastreabilidade das decisões.	Donthu *et al.*, 2021; Chen *et al.*, 2025.
8º Princípio da parcimônia	A IA deve ser usada onde agrega valor. Cálculos, normalizações e visualizações devem ser feitos com ferramentas apropriadas. A IA complementa, não substitui.	Usar IA para apoiar decisões, não para substituir cálculos e métodos. Empregar ferramentas especializadas para análises quantitativas e gráficos. Evitar complexidade desnecessária. Focar no que responde à pergunta de pesquisa.	• Uso excessivo de IA sem ganho real. • Substituir métodos estatísticos por "opiniões de IA". • Desviar foco da pergunta central.	Aria; Cuccurullo, 2017; Van Eck; Waltman, 2010.
9º Princípio da responsabilidade autoral	O pesquisador assume plena responsabilidade por todo o conteúdo do estudo. A IA é ferramenta de apoio, não autora.	Revisar, validar e aprovar todo o conteúdo. Declarar o uso de IA e as atividades realizadas. Assumir responsabilidade por interpretações e conclusões. Considerar impactos éticos e institucionais.	• Transferir responsabilidade para a IA. • Usar IA como justificativa para erros. • Falta de declaração do uso de IA.	Farquhar *et al.*, 2024; Donthu *et al.*, 2021.

Princípio	Explicação	Boas Práticas (Exemplos)	Riscos a Evitar	Referências Relacionadas
10º Princípio da atualização crítica	Modelos, ferramentas, bases e normas mudam rapidamente. É necessário revisar práticas, recomendações e limitações com regularidade.	Acompanhar avanços metodológicos e técnicos. Revisar políticas editoriais e diretrizes éticas. Avaliar novas ferramentas e comparar desempenho. Atualizar protocolos e capacitação da equipe.	• Usar práticas desatualizadas. • Ignorar mudanças nas bases e modelos. • Confiar em resultados obsoletos.	Farquhar *et al.*, 2024.

Fonte: Elaborado pelo autor com base em Aria e Cuccurullo (2017), Van Eck e Waltman (2010), Zupic e Čater (2015), Donthu *et al.* (2021), Farquhar *et al.* (2024) e Chen *et al.* (2025).

2.9 Exemplo aplicado: da ideia vaga ao problema cientométrico

Considere a ideia inicial: "quero estudar inteligência artificial na educação superior". Em sua forma vaga, essa ideia é ampla demais para um estudo cientométrico robusto. Ela não define se o interesse está em produção científica, impactos, colaboração, métodos de pesquisa, adoção institucional, ética, avaliação ou políticas educacionais. Também não delimita período, bases, idioma ou tipos de documento. Uma interação com IA pode ajudar a decompor o tema, mas não deve decidir sozinha o recorte. O pesquisador pode aplicar o *Prompt 2.7.1* e receber alternativas como: evolução temporal da produção sobre IA generativa na educação superior; redes de colaboração internacional; temas emergentes em avaliação automatizada; ou comparação entre abordagens pedagógicas e éticas.

A partir dessa saída, o pesquisador pode escolher uma pergunta mais específica: "Como evoluiu a literatura sobre IA generativa na educação superior entre 2022 e 2026, e quais temas estruturam esse campo?". Essa pergunta já sugere uma janela temporal, uma unidade de análise e tipos de resultados esperados. Ainda assim, precisa ser refinada. O termo "IA generativa" deve incluir "ChatGPT", "*large language models*", "*generative artificial intelligence*", "LLM" e outros termos? A expressão "educação superior" deve incluir "higher *education*", "university teaching", "college students" e "academic writing"? Serão analisados apenas artigos ou também anais de congressos? Serão incluídas bases como Scopus, Web of Science, ERIC, Dimensions ou OpenAlex? (KASNECI *et al.*, 2023)

O *Prompt 2.7.2* pode ajudar a transformar essas questões em decisões de escopo. A IA pode sugerir termos relacionados, exclusões e riscos de ruído, por exemplo, a confusão entre IA generativa aplicada à educação e estudos técnicos de IA publicados por universidades. O pesquisador deve testar a estratégia em bases reais e verificar amostras de resultados. Se a busca recupera muitos artigos sobre engenharia de modelos sem relação educacional, a *string* deve ser ajustada. Se exclui trabalhos relevantes por uso de terminologia alternativa, deve ser ampliada. A IA auxilia, mas o teste empírico da busca é indispensável.

Antes da coleta final, o *Prompt 2.7.3* pode ser usado para identificar riscos metodológicos. O modelo pode apontar que o período 2022–2026 favorece literatura muito recente e ainda pouco citada, tornando inadequada uma análise forte de impacto por citações. Pode indicar que revisões narrativas e artigos de opinião podem dominar a produção inicial, exigindo classificação por tipo de contribuição. Pode sugerir que o termo ChatGPT cria um viés em favor de uma ferramenta específica, enquanto "*large language models*" amplia o escopo. Essas observações não são conclusões; são alertas para revisão do protocolo.

Esse exemplo ilustra a função adequada da IA: ajudar a tornar explícitas decisões que, em muitos estudos, permanecem implícitas. A IA não substitui a busca na base, a leitura de

amostras, o cálculo de indicadores ou a interpretação substantiva. Ela apoia a passagem de uma intenção ampla para um plano analisável. Esse apoio é valioso porque muitos problemas cientométricos nascem mal formulados e só revelam suas fragilidades após a coleta. A interação com *prompts* pode antecipar parte dessa crítica.

2.10 Discussão conceitual: cientometria, IA e a nova materialidade dos métodos

A introdução da IA na cientometria convida a pensar a materialidade dos métodos. Métodos não são apenas ideias abstratas; são realizados por instrumentos, interfaces, formatos de dados, códigos, parâmetros e práticas de documentação. A bibliometria clássica dependia de índices de citação, catálogos, bases bibliográficas e procedimentos de contagem. O mapeamento científico dependeu de algoritmos de rede, técnicas de normalização e ambientes de visualização. A cientometria com IA depende também de *prompts*, modelos, políticas de uso, mecanismos de recuperação de informação e registros conversacionais. Cada camada material condiciona o que pode ser visto e como pode ser interpretado.

Essa nova materialidade desloca parte da metodologia para a linguagem. A forma como o pesquisador escreve uma instrução passa a influenciar a análise. Isso exige competências que combinam domínio substantivo, rigor metodológico e clareza comunicativa. Um bom *prompt* cientométrico se parece com uma boa instrução de codificação qualitativa, uma boa estratégia de busca e uma boa especificação de análise: deve ser claro, delimitado, justificável e revisável. A escrita do *prompt* torna-se parte da escrita do método (CHEN *et al.*, 2025; GARFIELD, 1955; PRICE, 1965; VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017).

Há também uma dimensão hermenêutica. Modelos de linguagem operam sobre padrões textuais e produzem interpretações probabilísticas. Quando usados para rotular *clusters* ou resumir tendências, eles participam de um processo de atribuição de significado. Essa atribuição não é neutra. Um *cluster* pode ser chamado de "aprendizagem adaptativa", "plataformas inteligentes", "personalização algorítmica" ou "automação pedagógica", dependendo do vocabulário privilegiado. Cada rótulo orienta a leitura do campo. Por isso, rótulos gerados por IA devem ser comparados, justificados e, quando necessário, substituídos por termos validados por especialistas (CHEN, 2006; VAN ECK; WALTMAN, 2010).

Outra questão é a relação entre evidência quantitativa e narrativa. A cientometria produz números e mapas, mas sua utilidade frequentemente depende da narrativa que os organiza. A IA generativa tem força narrativa. Ela pode transformar resultados dispersos em textos coerentes. Essa capacidade é útil, porém ambígua: a coerência textual pode exceder a coerência empírica. Um relatório cientométrico responsável deve evitar narrativas totalizantes, reconhecer lacunas e preservar contradições. A IA deve ser orientada a escrever com prudência, não com excesso de certeza.

A cientometria com IA também reacende debates sobre objetividade. Mapas da ciência já eram construções dependentes de escolhas. A IA acrescenta escolhas linguísticas e algorítmicas menos visíveis. Um estudo que usa IA para classificar temas não é automaticamente menos objetivo, desde que explicite e valide o procedimento. A objetividade, nesse contexto, não significa ausência de mediação, mas controle intersubjetivo das mediações. O leitor deve poder examinar a estratégia de busca, os dados, os algoritmos, os *prompts* e as revisões.

Em termos sociológicos, a IA pode alterar práticas de trabalho em grupos de pesquisa, bibliotecas universitárias, observatórios de ciência e agências de avaliação. Tarefas antes restritas a especialistas podem ser parcialmente distribuídas. Bibliotecários podem usar IA para refinar vocabulários; gestores podem solicitar painéis interpretativos; pesquisadores podem gerar rascunhos de relatórios; estudantes podem explorar campos emergentes. Essa

redistribuição cria oportunidades de formação, mas também exige políticas institucionais. Sem orientação, diferentes atores podem produzir análises inconsistentes, não auditáveis ou metodologicamente frágeis (BENDER *et al.*, 2021; JI *et al.*, 2023).

Por fim, a cientometria com IA deve ser pensada como prática situada. Não existe um único modelo de uso válido para todas as áreas. Ciências biomédicas, engenharias, humanidades, ciências sociais e artes têm regimes distintos de publicação, citação e comunicação. A IA pode ser mais eficaz em áreas com terminologia estabilizada e resumos padronizados, e menos confiável em áreas com conceitos disputados ou produção multilíngue. O desenho do *prompt* deve respeitar essas diferenças disciplinares. A cientometria rigorosa nunca é apenas técnica; é sensível ao campo que analisa.

2.11 Síntese do capítulo

A pergunta "por que Cientometria com IA?" não deve ser respondida com entusiasmo tecnológico genérico. A resposta adequada é metodológica: porque a IA pode ampliar a capacidade de formular, explorar, organizar, interpretar e comunicar análises da ciência, desde que integrada a protocolos transparentes e validados. A IA torna mais rápidas certas tarefas, mas também torna mais necessário o controle crítico. Ela amplia a escala da leitura, mas não substitui a compreensão. Ela produz texto, mas não garante verdade. Ela sugere padrões, mas não determina significado.

O capítulo defendeu cinco ideias principais. Primeiro, a mediação por IA altera a prática cientométrica ao introduzir interfaces conversacionais, aceleração iterativa, apoio semântico e novos riscos de opacidade. Segundo, a bibliometria tradicional não é abandonada; ela é ampliada por uma camada de *prompts* que deve dialogar com indicadores, redes e mapas. Terceiro, *prompts* devem ser tratados como instrumentos de pesquisa, isto é, procedimentos documentados, versionados e avaliáveis. Quarto, o papel humano permanece central nas funções de curadoria, validação, interpretação e responsabilidade. Quinto, a IA generativa possui limites incontornáveis, incluindo alucinação, instabilidade, vieses, opacidade e risco de extrapolação.

A partir dessas ideias, os capítulos seguintes podem aprofundar técnicas específicas. A formulação de perguntas cientométricas será tratada como etapa decisiva. As estratégias de busca serão abordadas como operações que combinam vocabulário humano e expansão assistida por IA. O pré-processamento será discutido como processo de normalização e julgamento. Indicadores, redes, mapas e visualizações serão analisados em sua dimensão técnica e interpretativa. A redação de relatórios será entendida como comunicação responsável de evidências. Em todos os casos, os *prompts* serão apresentados não como atalhos mágicos, mas como peças de um método que exige disciplina.

Em síntese, a Cientometria com IA deve ser compreendida como uma prática de inteligência metodológica aumentada. Seu valor não está em substituir o pesquisador, mas em ampliar sua capacidade de perguntar melhor, revisar mais cedo, explorar alternativas, registrar decisões e comunicar resultados com clareza. Seu risco está em produzir uma aparência de rigor sem os controles correspondentes. A diferença entre essas duas possibilidades será determinada menos pela sofisticação do modelo e mais pela qualidade do desenho de pesquisa.

Aprofundamento metodológico: roteiro de decisão para iniciar um estudo cientométrico com IA

O primeiro eixo de decisão é a pergunta de pesquisa. Em cientometria, perguntas vagas tendem a produzir análises descritivas sem foco. A pergunta deve indicar qual dimensão da ciência será observada: crescimento de produção, impacto, colaboração, estrutura intelectual, circulação internacional, interdisciplinaridade ou emergência temática. A IA pode auxiliar nessa etapa ao propor formulações alternativas, mas o pesquisador deve verificar se cada alternativa corresponde a dados disponíveis e a métodos apropriados. Perguntas boas são aquelas que podem ser respondidas por evidências bibliográficas, sem exigir inferências que o desenho não permite sustentar.

O segundo eixo é a unidade de análise. Um estudo pode tomar como unidade documentos, autores, instituições, países, periódicos, referências citadas, palavras-chave, conceitos ou combinações desses elementos. Cada unidade produz uma visão parcial do campo. Ao analisar documentos, privilegia-se a produção registrada. Ao analisar autores, destacam-se trajetórias e centralidades. Ao analisar instituições, observam-se capacidades organizacionais. Ao analisar termos, mapeia-se vocabulário. A IA pode sugerir unidades adequadas à pergunta, mas deve ser instruída a explicitar as consequências de cada escolha (DONTHU *et al.*, 2021; ZUPIC; ČATER, 2015).

O terceiro eixo é o *corpus* documental. A composição do *corpus* define o universo empírico do estudo. Um *corpus* estreito pode ser coerente para uma pergunta específica, mas insuficiente para representar um campo amplo. Um *corpus* amplo pode aumentar cobertura, mas dificultar a interpretação. A IA pode ajudar a formular critérios de inclusão e exclusão, porém esses critérios devem ser testados em amostras reais de registros. A revisão de documentos recuperados e excluídos é indispensável para avaliar se a estratégia de busca produz o equilíbrio desejado entre sensibilidade e precisão.

O quarto eixo é o vocabulário de busca. Termos científicos são instáveis, variam entre disciplinas e mudam ao longo do tempo. A IA é útil para expandir sinônimos, acrônimos, variantes linguísticas e expressões correlatas, mas pode incluir termos que não pertencem ao campo ou deixar de fora vocabulários locais. Por isso, a expansão terminológica deve ser seguida de verificação. Uma prática recomendável é pedir ao modelo que classifique cada termo sugerido por grau de pertinência e que indique ambiguidades possíveis. O pesquisador deve então testar a *string* em bases bibliográficas.

O quinto eixo é a seleção das fontes de dados. Bases de dados não são equivalentes. Elas diferem em cobertura disciplinar, cobertura geográfica, tratamento de citações, disponibilidade de resumos, qualidade de afiliações e formatos de exportação. Uma análise de ciências biomédicas pode exigir bases especializadas; uma análise interdisciplinar pode se beneficiar de bases amplas; uma análise de produção regional pode requerer bases nacionais. A IA pode listar opções, mas a justificativa final deve considerar a pergunta de pesquisa e as limitações de cobertura.

O sexto eixo é a temporalidade. A escolha da janela temporal influencia fortemente os resultados. Períodos longos favorecem análise de evolução, consolidação e mudança estrutural. Períodos curtos são úteis para temas emergentes, mas reduzem a confiabilidade de métricas baseadas em citação. A IA pode apoiar a identificação de marcos temporais, como surgimento de tecnologias, políticas públicas ou eventos científicos relevantes. Contudo, a janela temporal deve ser justificada por critérios analíticos, não apenas por conveniência de coleta.

O sétimo eixo é o tipo documental. Artigos, revisões, anais de congressos, editoriais, capítulos, livros, *preprints* e patentes representam modalidades diferentes de comunicação. Em algumas áreas, anais de congressos são centrais; em outras, artigos de periódicos dominam. Incluir todos os tipos pode ampliar a visão do campo, mas também misturar gêneros com funções distintas. A IA pode ajudar a prever consequências dessa inclusão, desde que o *prompt* peça explicitamente uma análise por tipo documental e não trate todos os registros como equivalentes.

O oitavo eixo é a normalização. Nomes de autores, instituições e palavras-chave frequentemente aparecem com variações ortográficas, abreviações e traduções. Sem normalização, redes podem se fragmentar e indicadores podem ser subestimados. A IA pode sugerir agrupamentos de variantes, especialmente em palavras-chave e nomes institucionais, mas a revisão humana é obrigatória. Em casos de homônimos, fusões institucionais ou autores com nomes semelhantes, a validação deve recorrer a identificadores, afiliações, ORCID, páginas institucionais ou registros confiáveis.

O nono eixo é a escolha de indicadores. Indicadores devem responder à pergunta. Uma pergunta sobre crescimento exige séries temporais; uma pergunta sobre influência pode usar citações normalizadas; uma pergunta sobre colaboração pede redes de coautoria; uma pergunta sobre interdisciplinaridade exige diversidade de categorias, referências ou periódicos. A IA pode sugerir indicadores, mas o pesquisador deve evitar acumular métricas sem função analítica. Um relatório com muitos indicadores pode parecer completo e, ao mesmo tempo, carecer de argumento científico.

O décimo eixo é a construção de redes. Coautoria, cocitação, acoplamento bibliográfico e coocorrência de termos representam relações distintas. A coautoria mostra copublicação; a cocitação revela associação retrospectiva entre referências; o acoplamento bibliográfico identifica proximidade entre documentos que citam referências comuns; a coocorrência de termos aproxima vocabulários. A IA pode explicar essas diferenças e sugerir usos, mas não deve interpretar todos os mapas como se fossem evidência de colaboração ou proximidade conceitual direta.

O décimo primeiro eixo é a escolha de parâmetros. Limiar mínimo de frequência, método de normalização, algoritmo de *cluster*ização e medida de centralidade alteram a estrutura visível do mapa. Em muitos estudos, esses parâmetros são tratados como decisões técnicas menores, mas podem definir quais atores ou temas aparecem. A IA pode ser usada para criar uma lista de parâmetros a documentar e para explicar seus efeitos prováveis. O pesquisador deve registrar não apenas os valores escolhidos, mas a razão de sua escolha.

O décimo segundo eixo é a rotulação de *clusters*. Um *cluster* é uma formação algorítmica; seu nome é uma interpretação. Modelos de linguagem podem sugerir rótulos úteis a partir de termos, títulos e resumos, mas esses rótulos precisam ser validados. Um bom *prompt* deve solicitar rótulos alternativos, justificativas baseadas em evidências textuais e indicação de documentos centrais. A presença de rótulos concorrentes pode revelar que o *cluster* é heterogêneo ou que contém subtemas que não devem ser comprimidos em uma única expressão.

O décimo terceiro eixo é a análise temporal. A IA pode ajudar a descrever fases de desenvolvimento de um campo, mas deve fazê-lo a partir de evidências explícitas: número de publicações por ano, mudança de termos, surgimento de autores centrais, deslocamento de temas ou alteração de colaboração. Sem dados temporais estruturados, a narrativa de evolução será especulativa. Por isso, *prompts* de análise temporal devem exigir separação entre observação empírica, interpretação provável e hipótese que requer confirmação.

O décimo quarto eixo é a visualização. Gráficos e mapas são instrumentos de leitura, não provas autossuficientes. A IA pode recomendar visualizações adequadas a diferentes resultados, mas precisa receber informações sobre o objetivo comunicacional. Um mapa de rede é útil para relações; uma linha temporal é útil para evolução; um diagrama Sankey pode mostrar fluxos; uma matriz pode comparar categorias. A visualização deve reduzir complexidade sem ocultar limitações. Legendas e notas metodológicas são parte do rigor visual.

O décimo quinto eixo é a interpretação. Interpretar resultados cientométricos exige distinguir descrição, explicação e avaliação. Descrever é dizer o que os dados mostram; explicar é propor razões possíveis; avaliar é atribuir valor ou consequência. A IA tende a misturar esses níveis se não for orientada. Um *prompt* robusto deve pedir respostas separadas em descrição objetiva, interpretação cautelosa, hipóteses alternativas e limitações. Essa estrutura reduz o risco de conclusões fortes demais.

O décimo sexto eixo é a comparação. Estudos cientométricos frequentemente comparam países, instituições, áreas ou períodos. Comparações exigem normalização e contextualização. Comparar números absolutos entre instituições de tamanhos diferentes pode ser enganoso; comparar citações entre áreas com padrões distintos pode ser injusto; comparar períodos recentes e antigos pode favorecer trabalhos mais antigos. A IA pode apontar riscos de comparação, mas o pesquisador deve selecionar métricas compatíveis com a pergunta e com o contexto.

O décimo sétimo eixo é a comunicação. O mesmo resultado pode ser apresentado em artigo científico, relatório executivo, painel institucional ou briefing para gestores. A IA pode adaptar linguagem e extensão, mas a adaptação deve preservar o conteúdo factual. Um risco comum é transformar uma descrição prudente em recomendação assertiva para tornar o texto mais persuasivo. *Prompts* de comunicação devem exigir fidelidade aos dados, manutenção das limitações e indicação clara do nível de confiança.

O décimo oitavo eixo é a validação por especialistas. A IA pode produzir classificações iniciais e sínteses temáticas, mas especialistas de domínio são necessários para avaliar terminologia, relevância e coerência interpretativa. A validação pode ocorrer por amostragem, oficina de revisão, dupla codificação ou comparação com literatura reconhecida. Em relatórios institucionais, a participação de especialistas também aumenta legitimidade e reduz o risco de interpretações descontextualizadas.

O décimo nono eixo é a documentação. Um estudo cientométrico com IA deve registrar *prompts*, versões, dados de entrada, saídas relevantes, revisões humanas e decisões finais. Essa documentação pode ser organizada em diário metodológico, apêndice ou repositório. A documentação não serve apenas à reprodução; serve também à aprendizagem. Ao examinar *prompts* anteriores, a equipe identifica instruções ambíguas, critérios inadequados e melhorias possíveis para estudos futuros.

O vigésimo eixo é a decisão sobre o que não automatizar. Nem todas as tarefas se beneficiam de IA generativa. Cálculos de indicadores devem ser feitos por ferramentas estatísticas ou bibliométricas adequadas; limpeza sensível de dados deve envolver regras verificáveis; decisões avaliativas devem permanecer sob responsabilidade humana. A IA é mais útil quando amplia capacidade crítica e organizacional. Quando sua inclusão aumenta opacidade, risco ou dependência sem ganho claro, a melhor decisão metodológica pode ser não usá-la.

Aprofundamento epistemológico: o que significa conhecer a ciência por meio de rastros digitais

A cientometria parte da premissa de que a ciência produz rastros. Artigos, citações, resumos, palavras-chave, afiliações, financiamentos, dados de uso, menções sociais e colaborações são manifestações documentais de práticas científicas. Esses rastros permitem observar regularidades que seriam invisíveis em escala puramente qualitativa. Entretanto, rastros não são a própria ciência. Eles são inscrições seletivas de processos cognitivos, institucionais e sociais. A entrada da IA nessa análise não elimina a distância entre rastro e fenômeno; ao contrário, adiciona uma nova camada de mediação interpretativa.

Conhecer a ciência por meio de rastros digitais exige reconhecer que cada rastro possui uma lógica de produção. Uma citação pode refletir influência, reconhecimento, crítica, obrigação retórica ou convenção disciplinar. Uma palavra-chave pode ser escolhida pelo autor, padronizada por vocabulário controlado ou influenciada por estratégias de visibilidade. Uma afiliação pode ocultar colaborações informais. Um resumo pode enfatizar resultados e omitir controvérsias. A IA que opera sobre esses elementos herda suas ambiguidades. Quando resume um campo a partir de resumos, trabalha sobre textos já estratégicos e condensados (BORNMANN; DANIEL, 2008).

Essa constatação não desqualifica a cientometria. Pelo contrário, define seu horizonte de validade. Estudos cientométricos são mais robustos quando indicam que tipo de fenômeno seus dados permitem observar. Uma análise de coautoria observa copublicação, não necessariamente colaboração efetiva. Uma análise de cocitação observa associação retrospectiva entre documentos citados, não necessariamente proximidade conceitual em todos os aspectos. Uma análise de termos observa vocabulário, não necessariamente práticas. A IA deve ser orientada a respeitar essas distinções.

Na prática, isso significa inserir nos *prompts* restrições epistemológicas. Em vez de pedir "explique a colaboração entre estes autores", pode-se pedir "interprete a rede de coautoria como evidência de copublicação, evitando inferir colaboração informal não observada". Em vez de pedir "quais são os autores mais importantes?", pode-se pedir "quais autores aparecem como mais centrais segundo as métricas fornecidas, e quais limitações impedem tratar centralidade como importância científica?". Essas formulações protegem a análise contra inferências abusivas.

A IA também deve ser usada para explicitar alternativas interpretativas. Um aumento de publicações pode ter múltiplas causas: crescimento real do interesse, mudança tecnológica, políticas de financiamento, indexação de novos periódicos, modismo terminológico ou alteração na estratégia de busca. Um *prompt* metodologicamente prudente pode solicitar que o modelo liste essas alternativas e indique que dados seriam necessários para discriminá-las. Desse modo, a IA ajuda a ampliar o espaço de possibilidades em vez de fechar prematuramente a interpretação.

Essa abertura é especialmente relevante em temas emergentes. Campos novos costumam apresentar terminologia instável, literatura interdisciplinar, alta proporção de *preprints*, rápida obsolescência e forte presença de discursos prospectivos. A IA pode identificar padrões iniciais, mas a interpretação deve ser provisória. Em tais contextos, recomenda-se evitar conclusões definitivas sobre impacto e maturidade. Indicadores de atenção e produção podem ser mais informativos que citações, mas também mais voláteis. A análise deve declarar essa condição.

Outro aspecto epistemológico diz respeito à relação entre quantificação e valor. Indicadores medem dimensões observáveis, não valores intrínsecos. Um artigo muito citado

pode ser influente, controverso, metodológico ou simplesmente útil como referência. Uma instituição produtiva pode publicar muito por tamanho, por especialização ou por política de incentivo. A IA, se não orientada, pode converter quantidades em juízos de valor. Por isso, *prompts* devem evitar termos como "melhor", "mais relevante" ou "mais importante" sem definir critério.

A cientometria com IA deve, portanto, adotar uma linguagem de qualificação. Em vez de "o autor X é o mais importante", prefere-se "o autor X apresenta maior número de publicações no *corpus* analisado" ou "o autor X ocupa posição central na rede de coautoria segundo a métrica de grau". Em vez de "o tema Y é dominante", pode-se dizer "o tema Y concentra a maior frequência de termos ou documentos neste recorte". A IA pode ser instruída a escrever dessa forma, reduzindo exageros interpretativos.

Por fim, conhecer a ciência por rastros digitais exige compromisso com pluralidade. Nem toda ciência circula nos mesmos canais. Humanidades e ciências sociais podem publicar livros, capítulos e ensaios com maior peso do que artigos indexados. Áreas aplicadas podem produzir relatórios, patentes e normas técnicas. Pesquisas locais podem ser publicadas em idiomas nacionais. Uma cientometria orientada apenas por grandes bases internacionais pode representar parcialmente a produção científica. A IA deve ser usada para identificar essas lacunas, não para apagá-las.

Aprofundamento ético: transparência, autoria e governança do uso de IA

O uso de IA em estudos científicos levanta questões de transparência. Quando um modelo contribui para a formulação de categorias, a redação de seções, a sumarização de literatura ou a classificação de documentos, essa contribuição deve ser descrita. A transparência não exige transformar cada interação em nota de rodapé, mas exige que usos substantivos sejam comunicados de modo claro. O leitor precisa saber se a análise textual foi manual, automatizada, assistida por IA ou validada por especialistas.

A autoria permanece humana. Modelos de linguagem não assumem responsabilidade, não respondem a críticas, não garantem integridade dos dados e não podem ser responsabilizados por erro. Eles podem auxiliar a produção, mas não satisfazem critérios de autoria científica. O autor humano deve revisar integralmente qualquer conteúdo gerado ou modificado por IA. Em cientometria, isso inclui verificar referências, números, nomes de autores, títulos de periódicos, interpretações de mapas e conclusões.

A governança institucional é outro aspecto relevante. Universidades, agências e grupos de pesquisa devem desenvolver diretrizes para o uso de IA em análises bibliométricas e cientométricas. Essas diretrizes podem abordar proteção de dados, escolha de ferramentas, documentação de *prompts*, validação de saídas, declaração de uso e limites para aplicações avaliativas. Sem governança, decisões podem variar de modo arbitrário entre analistas e relatórios, comprometendo comparabilidade e confiança (BENDER *et al.*, 2021; LUND; WANG, 2023; LIEBRENZ *et al.*, 2023).

A questão da privacidade não deve ser subestimada. Embora muitos dados bibliográficos sejam públicos, estudos institucionais podem combinar informações sensíveis, como desempenho de grupos, estratégias de colaboração, avaliações internas e prioridades de pesquisa. Inserir esses dados em plataformas externas pode violar políticas institucionais ou expectativas de confidencialidade. Em casos de dúvida, recomenda-se anonimização, uso de modelos locais ou restrição da IA a dados públicos.

A ética também envolve a prevenção de usos avaliativos simplistas. A cientometria pode ser usada para apoiar decisões, mas não deve reduzir pesquisadores ou programas a métricas isoladas. A IA pode facilitar a geração de rankings, classificações e narrativas avaliativas, aumentando o risco de decisões automatizadas. Um *prompt* responsável deve evitar recomendações normativas quando os dados não as sustentam. Por exemplo, em vez de "indique os grupos que devem receber investimento", o *prompt* deve pedir "descreva padrões observados e limitações para tomada de decisão" (LUND; WANG, 2023; LIEBRENZ *et al.*, 2023).

Outro desafio ético é a desigualdade de acesso. Instituições com melhores ferramentas, assinaturas de bases e infraestrutura de IA podem produzir análises mais sofisticadas, enquanto outras ficam dependentes de soluções limitadas. Isso pode ampliar assimetrias na capacidade de monitorar e comunicar a própria produção científica. O desenvolvimento de práticas abertas, uso de bases acessíveis e compartilhamento de protocolos de *prompts* pode reduzir parcialmente essa desigualdade.

O uso de IA também requer atenção à integridade textual. A geração automática de relatórios pode favorecer padronização excessiva, frases genéricas e ausência de voz analítica. Em textos científicos, clareza não deve significar homogeneidade acrítica. O pesquisador deve editar a saída, inserir contexto, remover generalidades e garantir que cada afirmação tenha função argumentativa. A IA pode fornecer rascunhos, mas o texto final deve refletir julgamento humano.

A governança dos *prompts* é uma dimensão emergente. *Prompts* usados em análises institucionais podem conter pressupostos avaliativos, vieses de linguagem ou critérios implícitos. Por isso, *prompts* relevantes devem ser revisados por pares, especialmente quando sustentam classificações ou relatórios oficiais. Uma biblioteca institucional de *prompts* pode incluir versões aprovadas, exemplos de uso, riscos conhecidos e procedimentos de validação.

Finalmente, a transparência deve alcançar as limitações. Relatórios cientométricos com IA devem explicitar que modelos generativos podem errar, que as saídas foram revisadas e que determinados resultados dependem da qualidade dos dados fornecidos. Essa declaração não enfraquece o estudo; fortalece sua credibilidade. A ciência avança quando seus métodos podem ser examinados, não quando suas incertezas são ocultadas.

Aprofundamento prático: matriz de riscos metodológicos iniciais[1]

Quadro 2.2 - Matriz de riscos metodológicos iniciais

Risco	Evidência típica	Consequência provável	Mitigação recomendada
Pergunta ampla demais	O estudo inicia com tema genérico e sem unidade de análise definida	A coleta recupera literatura heterogênea, dificultando indicadores e interpretação	Usar IA para decompor o tema em perguntas menores e validar a escolha com especialistas
Estratégia de busca enviesada	A *string* privilegia um termo de moda ou uma tecnologia específica	A literatura recuperada reflete vocabulário parcial, não o campo como um todo	Solicitar expansão terminológica e testar amostras de documentos incluídos e excluídos

1 Nota metodológica: os riscos devem ser examinados antes da coleta definitiva, pois correções tardias costumam exigir refazer etapas de busca, limpeza e análise. A IA pode apoiar a matriz inicial, mas o pesquisador deve distinguir riscos reais de observações genéricas e registrar gravidade, probabilidade e ação adotada (DONTHU *et al.*, 2021; ZUPIC; ČATER, 2015; JI *et al.*, 2023).

Risco	Evidência típica	Consequência provável	Mitigação recomendada
Cobertura inadequada da base	A fonte escolhida não representa bem a área, idioma ou tipo documental	Resultados podem subestimar comunidades ou formatos relevantes	Comparar bases, justificar a escolha e declarar limitações de cobertura
Dependência de citações recentes	O estudo avalia impacto de literatura muito nova	Citações insuficientes podem gerar conclusões prematuras	Priorizar produção, atenção, colaboração e análise temática, tratando impacto com cautela
Normalização insuficiente	Autores, instituições ou palavras-chave aparecem com múltiplas variantes	Redes e rankings ficam fragmentados ou distorcidos	Usar procedimentos de limpeza, listas de equivalência e revisão amostral
Rotulação automática acrítica	*Clusters* recebem nomes gerados por IA sem validação	Rótulos podem orientar interpretações erradas	Comparar rótulos alternativos, verificar documentos centrais e consultar especialistas
Inferência causal indevida	O relatório transforma associação em causalidade	Conclusões excedem o desenho observacional	Exigir que *prompts* separem correlação, hipótese e evidência necessária
Omissão de limitações	A redação final enfatiza achados e minimiza fragilidades	Leitores podem superestimar a robustez do estudo	Usar *prompt* específico para seção de limitações e revisão crítica independente
Alucinação bibliográfica	Referências sugeridas pela IA não são verificadas	O relatório incorpora fontes inexistentes ou incorretas	Checar todas as referências em bases confiáveis antes de uso
Falta de documentação	*Prompts* e decisões humanas não são registrados	A análise torna-se pouco auditável	Manter diário metodológico com versões de *prompts*, saídas e decisões

Fonte: Elaborado pelo autor (2026).

Aprofundamento didático: exemplos de uso de prompts ao longo do fluxo inicial

Um primeiro exemplo de uso didático consiste em solicitar à IA que transforme uma demanda institucional em pergunta cientométrica. Demandas institucionais costumam aparecer em linguagem decisória, como "precisamos saber em que área investir" ou "queremos identificar nossos parceiros estratégicos". O papel do pesquisador é converter essa demanda em problema analítico. Um *prompt* adequado pode pedir que o modelo separe objetivo gerencial, pergunta científica, indicadores possíveis e limites de inferência. Essa separação evita que o estudo comece orientado por uma recomendação previamente desejada. A IA ajuda a estruturar o problema, mas a legitimidade do diagnóstico dependerá da qualidade dos dados e da prudência interpretativa.

Um segundo exemplo envolve a revisão de uma estratégia de busca. Depois de criar uma *string* inicial, o pesquisador pode pedir ao modelo que identifique termos ambíguos, termos ausentes e expressões que podem gerar ruído. A resposta deve ser tratada como lista de verificação. O modelo pode sugerir sinônimos úteis, mas também pode incluir termos excessivamente amplos. Por isso, o *prompt* deve exigir que cada sugestão seja acompanhada por justificativa e risco. Uma boa saída não é aquela que simplesmente amplia a busca,

mas aquela que ajuda o pesquisador a compreender o custo metodológico de ampliar ou restringir o vocabulário.

Um terceiro exemplo é a preparação de critérios de inclusão e exclusão. Em estudos cientométricos, tais critérios muitas vezes são menos formalizados do que em revisões sistemáticas, mas continuam sendo importantes. A IA pode auxiliar a redigir critérios baseados em tema, tipo de documento, idioma, período, área e aderência ao objetivo. O *prompt* deve pedir exemplos de registros que seriam incluídos e excluídos, pois exemplos revelam ambiguidades. Se o modelo não consegue explicar por que um documento seria excluído, o critério provavelmente está mal definido. A decisão final deve ser testada em amostra real do *corpus*.

Um quarto exemplo diz respeito à classificação preliminar de resumos. A IA pode classificar documentos em categorias temáticas, metodológicas ou aplicativas. Entretanto, a classificação só é metodologicamente aceitável se houver critérios claros, possibilidade de categoria "ambígua" e revisão humana. Um *prompt* adequado deve solicitar que o modelo apresente justificativa textual, grau de confiança e indicação de casos duvidosos. Essa prática impede que a classificação seja vista como automática e definitiva. Os casos de baixa confiança são particularmente valiosos, pois mostram fronteiras conceituais instáveis do campo analisado.

Um quinto exemplo é a geração de rótulos para *clusters*. Depois que uma rede de coocorrência ou cocitação é produzida por ferramenta especializada, o pesquisador pode fornecer ao modelo os termos ou documentos mais representativos de cada *cluster*. A IA pode sugerir nomes concisos e interpretações possíveis. O *prompt* deve exigir, porém, pelo menos três rótulos alternativos e uma justificativa para cada um. Quando os rótulos variam muito, isso pode indicar que o *cluster* é internamente heterogêneo. A divergência entre rótulos não deve ser ocultada; pode ser incorporada à interpretação como sinal de complexidade temática.

Um sexto exemplo refere-se à redação da metodologia. Modelos de linguagem podem ajudar a transformar procedimentos técnicos em texto claro. O pesquisador pode fornecer base de dados, data de coleta, *string* de busca, filtros, número de registros, critérios de limpeza, software e parâmetros. O *prompt* deve proibir a invenção de etapas não realizadas e exigir linguagem no passado, descrevendo apenas o que foi efetivamente feito. Essa restrição é importante porque a IA tende a completar lacunas com procedimentos metodológicos plausíveis. A seção metodológica deve ser fiel ao processo real, não ao processo ideal imaginado pelo modelo.

Um sétimo exemplo é a revisão de limitações. Muitos relatórios cientométricos dedicam pouca atenção aos limites de seus dados e métodos. A IA pode ser usada como revisor crítico, recebendo o protocolo e os resultados para apontar fragilidades. O *prompt* deve pedir limitações por dimensão: cobertura da base, estratégia de busca, qualidade dos metadados, indicadores, redes, temporalidade, interpretação e generalização. A saída deve ser revisada para evitar frases genéricas, mas costuma ajudar a lembrar problemas que o pesquisador pode naturalizar por familiaridade com o próprio estudo.

Um oitavo exemplo é a construção de sínteses executivas. Gestores precisam de textos curtos, mas cientificamente responsáveis. A IA pode adaptar a linguagem de resultados complexos, desde que o *prompt* exija preservação de incertezas. Em vez de solicitar "faça um resumo persuasivo", convém pedir "produza uma síntese executiva com achados, evidências, implicações e cautelas". Essa estrutura impede que o resumo se transforme em peça de convencimento. A comunicação de resultados cientométricos deve apoiar decisões informadas, não produzir autoridade retórica indevida.

Um nono exemplo é a checagem de proporcionalidade das conclusões. Antes de finalizar um relatório, o pesquisador pode pedir que a IA avalie se cada conclusão é sustentada pelos resultados apresentados. O *prompt* deve solicitar uma tabela com conclusão, evidência correspondente, nível de sustentação e revisão sugerida. Essa etapa funciona como auditoria argumentativa. Conclusões sem evidência suficiente podem ser reescritas como hipóteses ou removidas. A IA não decide a validade final, mas ajuda a tornar visíveis saltos interpretativos que podem passar despercebidos na redação.

Um décimo exemplo é a preparação de materiais suplementares. Estudos com IA devem, sempre que possível, disponibilizar *prompts*, critérios de validação e decisões de curadoria. O modelo pode ajudar a organizar esses materiais em formato legível: lista de *prompts*, finalidade, dados de entrada, saída esperada e validação aplicada. Essa documentação favorece transparência e reuso. Ao mesmo tempo, o pesquisador deve remover informações confidenciais e respeitar licenças de bases de dados. A abertura metodológica deve ser equilibrada com responsabilidade legal e institucional.

Um décimo primeiro exemplo envolve a formação de equipes. Em grupos de pesquisa, diferentes membros podem usar *prompts* de modos inconsistentes. Uma biblioteca comum de *prompts* reduz variação e facilita treinamento. A IA pode ajudar a criar versões didáticas desses *prompts*, explicando quando usar, que dados fornecer e como interpretar a saída. A equipe deve revisar coletivamente os modelos de *prompt*, pois instruções mal formuladas podem induzir vieses sistemáticos. A padronização, nesse caso, não significa rigidez absoluta, mas criação de um vocabulário metodológico compartilhado.

Um décimo segundo exemplo é o uso da IA para reflexão pós-estudo. Após concluir a análise, o pesquisador pode fornecer ao modelo o protocolo, as dificuldades encontradas e as decisões tomadas, solicitando uma avaliação retrospectiva. O objetivo não é alterar resultados, mas aprender com o processo. O modelo pode apontar etapas que deveriam ter sido documentadas melhor, decisões que exigiriam justificativa adicional ou oportunidades de validação futura. Essa prática transforma a IA em apoio à melhoria contínua dos procedimentos cientométricos, fortalecendo a maturidade metodológica da equipe.

Um décimo terceiro exemplo é a revisão de linguagem avaliativa. Termos como excelência, liderança, impacto, relevância e prioridade podem carregar juízos normativos. A IA pode ser instruída a localizar expressões avaliativas no relatório e sugerir versões descritivas, baseadas em métricas explícitas. Essa revisão é útil quando o estudo será usado por gestores, pois evita que indicadores sejam transformados em rankings implícitos. A linguagem científica deve informar com precisão o que foi medido e deixar claro o que permanece fora do alcance do método.

Um décimo quarto exemplo é a análise de coerência entre figuras e texto. Em relatórios cientométricos, é comum que gráficos, mapas e tabelas sejam descritos de modo incompleto ou excessivo. A IA pode receber a descrição dos resultados e verificar se a narrativa corresponde aos elementos visuais. O *prompt* deve pedir identificação de afirmações não representadas na figura, ausência de informações importantes e necessidade de notas metodológicas. Essa checagem melhora a qualidade comunicacional e reduz o risco de interpretações que o leitor não consegue verificar visualmente.

Um décimo quinto exemplo refere-se à adaptação multilíngue. Em temas internacionais, palavras-chave em inglês dominam as bases, mas comunidades locais podem usar termos em português, espanhol ou outras línguas. A IA pode ajudar a construir equivalências terminológicas e a reconhecer falsos cognatos. O *prompt* deve solicitar tradução comentada, não tradução literal. Termos técnicos devem ser acompanhados de observações sobre uso

disciplinar. Essa prática é relevante para evitar que a análise reproduza apenas o vocabulário anglófono e subestime tradições científicas locais.

Um décimo sexto exemplo é a identificação de lacunas de dados. Antes de iniciar a interpretação, o pesquisador pode pedir que a IA examine quais informações seriam necessárias para sustentar cada tipo de conclusão desejada. Se o relatório pretende falar de colaboração internacional, são necessárias afiliações confiáveis. Se pretende falar de impacto, são necessárias citações e normalização por área e tempo. Se pretende falar de temas emergentes, são necessárias séries temporais e vocabulário consistente. Esse exercício impede que perguntas importantes sejam respondidas com dados inadequados.

Um décimo sétimo exemplo é o desenho de uma trilha de auditoria. A IA pode ajudar a elaborar uma lista cronológica de decisões: pergunta inicial, refinamento de escopo, busca, exportação, limpeza, análise, visualização, interpretação, revisão e comunicação. Para cada decisão, o pesquisador registra justificativa, ferramenta usada, participação da IA e validação aplicada. Essa trilha torna o estudo mais transparente e facilita revisões futuras. Em ambientes institucionais, ela também permite comparar estudos realizados por equipes diferentes e construir padrões de qualidade.

Um último exemplo consiste em pedir à IA que diferencie resultados robustos, resultados provisórios e perguntas remanescentes. Essa diferenciação ajuda a encerrar o estudo sem encerrar indevidamente o debate científico. Em áreas dinâmicas, um relatório cientométrico deve ser compreendido como fotografia situada, não como diagnóstico definitivo. A IA pode ajudar a formular essa prudência de modo claro para leitores técnicos e não técnicos.

2.12 Considerações finais

A introdução da inteligência artificial na cientometria é um fenômeno metodologicamente relevante porque modifica o modo como pesquisadores interagem com dados, textos e interpretações. Sua contribuição principal não está em automatizar integralmente a pesquisa, mas em ampliar a capacidade de organizar problemas complexos, explorar alternativas, produzir sínteses preliminares e revisar criticamente decisões. Para que essa contribuição seja científica, e não apenas operacional, é necessário tratar *prompts* como componentes do método.

O capítulo mostrou que a cientometria com IA deve ser construída sobre continuidade e transformação. A continuidade está nos fundamentos já consolidados da área: atenção à qualidade dos dados, uso criterioso de indicadores, compreensão de redes, interpretação contextual e consciência dos vieses. A transformação está na introdução de uma interface semântica capaz de apoiar tarefas antes dependentes de leitura manual intensiva. Essa transformação, entretanto, exige novas competências: engenharia de *prompts*, documentação de interações, auditoria de saídas e governança ética.

A pergunta "por que Cientometria com IA?" pode, portanto, ser respondida da seguinte maneira: porque a análise contemporânea da ciência tornou-se grande demais, textual demais, dinâmica demais e estrategicamente importante demais para dispensar instrumentos de apoio semântico; mas também porque esses instrumentos são arriscados demais para serem usados sem método. A IA deve ampliar a inteligência do pesquisador, não substituí-la. Deve tornar a análise mais transparente, não mais opaca. Deve favorecer perguntas melhores, não apenas respostas mais rápidas.

Nos próximos capítulos, essa perspectiva será operacionalizada em procedimentos específicos. O leitor será convidado a transformar ideias vagas em problemas cientométricos, construir estratégias de busca assistidas por IA, realizar pré-processamento de dados, selecionar indicadores, interpretar redes e mapas, analisar tendências temporais, produzir

visualizações e redigir relatórios. Em cada etapa, os *prompts* serão apresentados como ferramentas de estruturação, crítica e comunicação. A promessa do livro é formar uma competência híbrida: saber medir a ciência, saber interpretar seus rastros e saber dialogar criticamente com sistemas de IA.

Como orientação final para o leitor, recomenda-se que todo estudo cientométrico com IA seja iniciado por uma pergunta simples de auditoria: quais decisões deste estudo seriam diferentes se a IA não tivesse sido usada? A resposta ajuda a identificar onde a ferramenta realmente influenciou o desenho, a interpretação ou a comunicação. Se a IA apenas revisou linguagem, o registro pode ser breve. Se a IA participou da formulação de categorias, da classificação de documentos ou da interpretação de redes, o registro deve ser mais detalhado. Essa gradação evita tanto o ocultamento do uso de IA quanto a burocratização desnecessária. O princípio orientador é que qualquer leitor qualificado deve conseguir compreender como a IA participou do raciocínio científico e quais controles humanos foram aplicados para preservar a confiabilidade da análise.

Capítulo 3

Fundamentos de Cientometria

3.1 Introdução

Este capítulo apresenta os fundamentos conceituais e metodológicos da Cientometria no contexto de uma trilha dedicada ao uso de inteligência artificial, especialmente por meio de *prompts*, para analisar, mapear e interpretar a ciência. O objetivo não é apenas definir termos, mas estabelecer uma gramática de trabalho que permita ao pesquisador compreender o que está sendo contado, por que está sendo contado, que tipo de evidência é produzido e quais decisões interpretativas permanecem sob responsabilidade humana. Em estudos cientométricos, números raramente são neutros: eles dependem de bases de dados, escolhas de indexação, critérios de inclusão, unidades de análise, janelas temporais, normalizações, modelos estatísticos, técnicas de rede e convenções disciplinares. A IA pode ampliar a capacidade de formular perguntas, comparar indicadores e traduzir resultados para diferentes públicos, mas não elimina a necessidade de rigor na construção do objeto empírico (MINGERS; LEYDESDORFF, 2015; WALTMAN, 2016).

A Cientometria deve ser entendida como parte de um conjunto mais amplo de metrias da informação e da comunicação científica. Bibliometria, informetria, webometria e altmetria não são sinônimos perfeitos, embora compartilhem métodos, fontes e problemas. A bibliometria consolidou técnicas de contagem e análise de documentos, referências e citações; a cientometria ampliou o foco para a dinâmica social, cognitiva e institucional da ciência; a informetria generalizou o estudo quantitativo de fenômenos informacionais; a webometria examinou vínculos, presenças e relações na web; e a altmetria incorporou traços digitais de atenção, uso e circulação social da produção acadêmica. Com IA generativa, essas fronteiras tornam-se ainda mais porosas, pois *prompts* podem operar sobre metadados bibliográficos, resumos, textos completos, identificadores persistentes, perfis de autores, redes, indicadores e narrativas de avaliação (PRITCHARD, 1969; MINGERS; LEYDESDORFF, 2015; THELWALL; VAUGHAN; BJÖRNEBORN, 2005; HAUSTEIN *et al.*, 2014).

A relação entre Cientometria e inteligência artificial, entretanto, deve ser construída com prudência metodológica. Modelos de linguagem são úteis para reescrever perguntas, organizar taxonomias, sugerir dimensões analíticas, apoiar a interpretação de padrões e produzir versões comunicáveis de resultados técnicos. Porém, também podem gerar afirmações sem base documental, simplificar excessivamente conceitos, confundir indicadores, inventar referências ou atribuir significados indevidos a correlações. Por isso, a incorporação de *prompts* ao trabalho cientométrico exige uma postura dupla: abertura à automação assistiva e resistência à terceirização do julgamento científico. O pesquisador deve registrar *prompts*, controlar entradas, verificar saídas e explicitar as decisões tomadas depois da interação com a IA.

O capítulo está organizado em cinco eixos. Primeiro, apresenta as relações entre cientometria, bibliometria, informetria, webometria e altmetria. Em seguida, discute as dimensões de produção científica, impacto, colaboração, circulação e visibilidade. A terceira seção examina as unidades de análise mais comuns em estudos cientométricos: artigo, autor, periódico, instituição, país, área e conceito. A quarta seção sistematiza indicadores de produtividade, impacto e colaboração, observando suas aplicações e limitações. A quinta seção distingue medir, mapear, avaliar e interpretar, uma distinção decisiva para evitar usos tecnocráticos de métricas. Ao final, são propostos *prompts* para explicar indicadores a

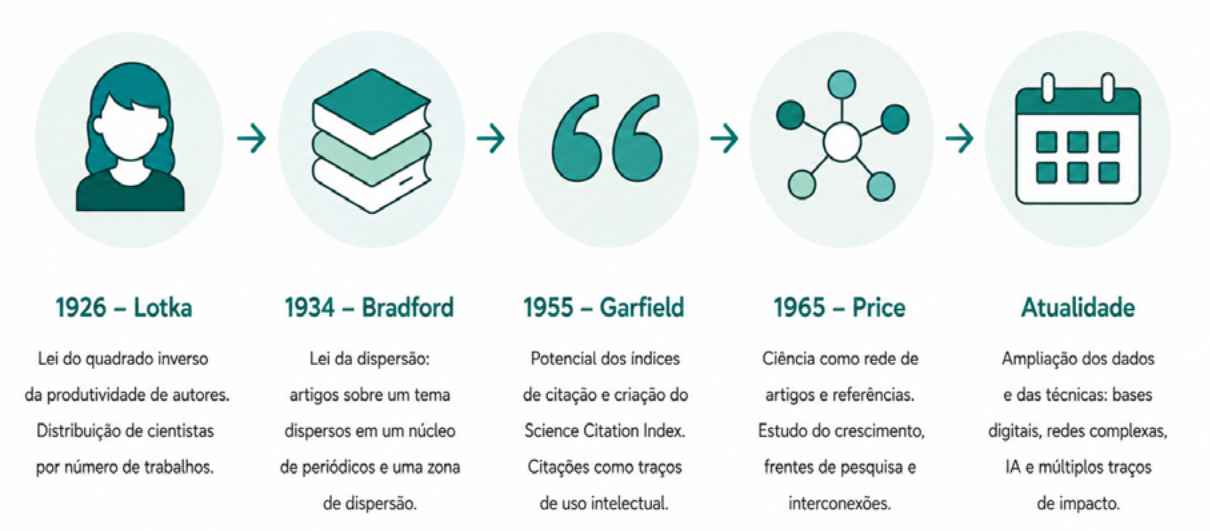
públicos não especialistas, comparar métricas e selecionar indicadores adequados a uma pergunta de pesquisa.

Figura 3.1 - Fundamentos conceituais e metodológicos da Cientometria

Fonte: Imagem gerada por IA, ChatGPT (2026).

3.2 Cientometria, bibliometria, informetria, webometria e altmetria

Figura 3.2 - Evolução histórica dos estudos métricos da informação



Fonte: Imagem gerada por IA, ChatGPT (2026).

Genealogia das metrias

A tradição das metrias da informação nasce de tentativas de reconhecer regularidades quantitativas na produção e circulação de documentos científicos. Lotka investigou a distribuição da produtividade de autores; Bradford tratou da dispersão de artigos em periódicos; Garfield demonstrou o potencial dos índices de citação; Price analisou a ciência como rede de artigos

e referências. Essas formulações não devem ser lidas como leis naturais imutáveis, mas como modelos estatísticos e históricos que precisam ser testados em cada *corpus* (LOTKA, 1926; BRADFORD, 1934; GARFIELD, 1955; PRICE, 1965).

Quadro 3.1 - Famílias de metrias e focos analíticos

Família	Objeto preferencial	Exemplos de evidência	Cuidado interpretativo
Bibliometria	Documentos, referências, autores e periódicos	Publicações, citações, palavras-chave, fontes	Não reduzir qualidade a volume ou prestígio editorial
Cientometria	Ciência como sistema social, cognitivo e institucional	Indicadores, redes, campos, políticas, desempenho	Contextualizar resultados por área, tempo e missão institucional
Informetria	Fenômenos informacionais em sentido amplo	Fluxos, registros, padrões de uso, distribuições	Evitar diluição conceitual do objeto
Webometria	Presença e relações na web	Links, domínios, downloads, repositórios, páginas	Considerar volatilidade e dependência de plataformas
Altmetria	Atenção e circulação online	Menções, salvamentos, notícias, blogs, políticas	Não confundir atenção social com qualidade científica

Fonte: Elaborado pelo autor (2026).

Bibliometria como análise documental

A bibliometria ocupa-se da aplicação de métodos quantitativos a livros, artigos, periódicos, referências, autores e outros registros bibliográficos. Ela permite examinar crescimento de literatura, núcleos de periódicos, padrões de autoria, frequência de palavras, relações de citação e distribuição de documentos por campos. Seu alcance depende da qualidade dos registros, da cobertura das bases e da clareza sobre o que se toma como documento analisável (PRITCHARD, 1969; ARIA; CUCCURULLO, 2017).

Cientometria como estudo da ciência

A cientometria amplia a bibliometria ao estudar a ciência como sistema de comunicação, produção simbólica, organização institucional e dinâmica de inovação. Além de contar publicações e citações, pergunta como campos se formam, como comunidades colaboram, como especialidades emergem e como políticas científicas utilizam indicadores. A ampliação do escopo implica reconhecer que indicadores são instrumentos de conhecimento e, simultaneamente, instrumentos de governo da ciência (MINGERS; LEYDESDORFF, 2015; LEYDESDORFF; MILOJEVIĆ, 2015).

Informetria como conceito guarda-chuva

A informetria é frequentemente compreendida como campo mais geral, dedicado ao estudo quantitativo da informação em quaisquer suportes, sistemas ou fluxos. Ela inclui fenômenos bibliográficos, científicos, tecnológicos, documentais, digitais e comunicacionais,

sem restringir-se à literatura acadêmica indexada. Essa abrangência é útil para estudos híbridos, mas também exige precisão terminológica para que a análise não dilua seu objeto (MINGERS; LEYDESDORFF, 2015).

Webometria e ambiente digital

A webometria surge quando a web se torna espaço de publicação, circulação, ligação e visibilidade de conteúdos científicos e institucionais. Links, domínios, páginas, repositórios, perfis, downloads, menções e interações passam a compor evidências sobre presença e circulação digital. Tais evidências são voláteis, dependem de plataformas e podem refletir estratégias de comunicação tanto quanto impacto científico propriamente dito (THELWALL; VAUGHAN; BJÖRNEBORN, 2005).

Altmetria e traços de atenção

A altmetria propõe captar rastros de atenção e uso de produtos científicos em ambientes como gestores de referência, redes sociais, blogs, notícias, políticas públicas e plataformas de discussão. Ela amplia a noção de impacto ao incluir sinais precoces, sociais ou não acadêmicos de circulação. A interpretação desses sinais requer cautela, porque atenção não equivale automaticamente a qualidade, influência epistemológica ou benefício social (HAUSTEIN *et al.*, 2014; BORNMANN, 2015).

Quadro 3.2 - Quadro comparativo - as cinco metrias

Metria	Foco central	Objeto típico	Fontes comuns	Exemplos de indicadores	Limitações principais
Bibliometria	Aplicação de métodos quantitativos a registros bibliográficos.	Livros, artigos, periódicos, referências, autores, palavras-chave.	Scopus, Web of Science, PubMed, Dimensions, Lens, bases especializadas.	Produção, crescimento, produtividade, leis de Lotka e Bradford, citações, frequência de termos.	Depende da cobertura e qualidade dos registros; reflete o que é indexado, não toda a ciência.
Cientometria	Estudo da ciência como sistema de comunicação, organização e inovação.	Documentos, autores, instituições, países, áreas, redes, temas.	Bases bibliográficas integradas, dados institucionais, políticas científicas, repositórios.	Redes de coautoria, cocitação, acoplamento, interdisciplinaridade, impacto, colaboração, perfis institucionais.	Indicadores são instrumentos de conhecimento e de governo; interpretar exige contexto e teoria.
Informetria	Estudo quantitativo da informação em quaisquer suportes ou fluxos.	Informação científica, tecnológica, documental, digital, comunicacional.	Bases diversas, repositórios, bancos de dados, sistemas de informação.	Produção e uso de informação, fluxos, recuperação, organização e comportamento informacional.	Amplitude conceitual exige precisão terminológica para não diluir o objeto.
Webometria	Estudo da presença e circulação da informação na web.	Páginas, sites, links, domínios, repositórios, perfis institucionais.	Motores de busca, diretórios web, repositórios, ferramentas de coleta de links.	Tamanho web, visibilidade, links externos e internos, páginas, domínios, downloads.	Volátil e dependente de plataformas; mede presença digital, não qualidade científica.

Metria	Foco central	Objeto típico	Fontes comuns	Exemplos de indicadores	Limitações principais
Altmetria	Captação de atenção e uso em ambientes online e não tradicionais.	Menções, tweets, blogs, notícias, políticas públicas, gestores de referência.	*Altmetric*, PlumX, Crossref Event Data, redes sociais, blogs, mídias.	Menções, compartilhamentos, curtidas, comentários, leitura em gestores, notícias, políticas.	Atenção ≠ qualidade ou impacto científico; suscetível a ruído, manipulação e viés.

Fonte: Elaborado pelo autor (2026).

Integração das metrias

Em estudos contemporâneos, os limites entre bibliometria, cientometria, webometria e altmetria raramente são absolutos. Um mesmo projeto pode combinar publicações indexadas, redes de coautoria, citações normalizadas, menções em documentos de política, indicadores de atenção e análises semânticas. A integração metodológica, contudo, só é legítima quando cada fonte é analisada segundo suas propriedades e seus vieses específicos (WALTMAN, 2016; HICKS *et al.*, 2015).

Prompts como mediação terminológica

No contexto desta trilha, *prompts* podem ajudar a traduzir diferenças conceituais entre as metrias, mas não devem substituir a leitura da literatura especializada. Um *prompt* bem desenhado pode solicitar definições comparativas, exemplos de aplicação, limites metodológicos e critérios para escolher uma abordagem. A saída da IA deve ser tratada como material de apoio, porque modelos de linguagem podem confundir tradições conceituais ou apresentar classificações excessivamente esquemáticas (BROWN *et al.*, 2020; JI *et al.*, 2023).

Cientometria como linguagem de evidência

O valor da cientometria reside em construir evidências sobre a ciência a partir de documentos e relações observáveis. Essas evidências incluem contagens, proporções, distribuições, redes, séries temporais, mapas, perfis e comparações entre unidades. A força analítica do campo depende de converter registros dispersos em argumentos transparentes, reprodutíveis e sensíveis ao contexto (MINGERS; LEYDESDORFF, 2015).

Risco de reificação

Um perigo recorrente em todas as metrias é transformar indicadores em substitutos do fenômeno investigado. Número de artigos não é produtividade intelectual plena; citações não são qualidade; downloads não são uso substantivo; menções online não são impacto social comprovado. A formação cientométrica precisa ensinar o leitor a operar indicadores sem confundi-los com a totalidade da vida científica (BORNMANN; DANIEL, 2008; HICKS *et al.*, 2015).

A distinção entre essas metrias é mais do que terminológica. Ela orienta o desenho da coleta, a escolha das bases, a interpretação de indicadores e a forma de comunicar resultados. Um estudo que se declara bibliométrico, mas produz inferências sobre políticas científicas, aproxima-se da cientometria; um estudo que utiliza menções em redes sociais aproxima-se da altmetria; uma análise de links institucionais se situa no campo webométrico; uma investigação quantitativa de fluxos informacionais mais amplos pode ser descrita como informétrica. O

pesquisador precisa reconhecer essas camadas para que o escopo do estudo não seja menor do que suas conclusões, nem maior do que suas evidências.

Em uma obra sobre Cientometria com IA, essa diferenciação adquire papel pedagógico. Modelos generativos são hábeis em produzir definições plausíveis, mas nem sempre preservam a história dos conceitos ou suas fronteiras de uso. Por isso, é recomendável que o analista solicite definições comparativas, peça exemplos de uso correto e incorreto, exija referências bibliográficas e confira os resultados. A IA deve ser treinada pelo contexto do *prompt* a reconhecer que bibliometria, cientometria, informetria, webometria e altmetria formam uma família de métodos, não uma lista de rótulos intercambiáveis.

Por fim, a compreensão dessas tradições ajuda a evitar o reducionismo de tomar a cientometria como mera técnica de ranqueamento. O campo nasceu da necessidade de compreender padrões da literatura e da comunicação científica, mas se tornou também um espaço de reflexão sobre avaliação, política científica, organização do conhecimento e dinâmica social da ciência. Essa dupla vocação - técnica e crítica - é indispensável para usar IA de modo responsável: *prompts* podem acelerar análises, mas os fundamentos definem o que merece ser analisado, como deve ser analisado e até onde se pode concluir.

3.3 Produção científica, impacto, colaboração, circulação e visibilidade

Produção científica como registro

A produção científica refere-se aos produtos formalmente registrados da atividade de pesquisa, como artigos, revisões, comunicações em congresso, livros, capítulos, dados, software e patentes. Na prática cientométrica, artigos de periódicos e anais de conferências costumam ter centralidade por sua maior padronização em bases bibliográficas. Essa centralidade, entretanto, pode subrepresentar áreas em que livros, relatórios, protótipos, bases de dados ou produção técnica possuem maior relevância (MINGERS; LEYDESDORFF, 2015; HICKS *et al.*, 2015).

Produtividade e crescimento

O estudo da produtividade envolve contagens por autor, instituição, país, periódico, área, tema ou período. Taxas de crescimento, produtividade acumulada, concentração de autores e expansão de palavras-chave ajudam a descrever a dinâmica de uma literatura. O crescimento quantitativo não deve ser automaticamente interpretado como maturidade, qualidade ou consolidação teórica, pois pode refletir modas, mudanças de indexação ou incentivos institucionais (LOTKA, 1926; ARIA; CUCCURULLO, 2017).

Impacto como conceito plural

Impacto é uma das categorias mais disputadas da cientometria, porque combina sentidos cognitivos, sociais, tecnológicos, clínicos, econômicos e políticos. Citações medem relações entre documentos, não necessariamente aprovação, uso prático ou qualidade intrínseca. A interpretação do impacto exige considerar a área, a idade do documento, o tipo de publicação, o idioma, a cobertura da base e a função retórica das citações (BORNMANN; DANIEL, 2008; WALTMAN, 2016).

Impacto acadêmico e impacto social

O impacto acadêmico costuma ser operacionalizado por citações, influência em redes de conhecimento, presença em revisões e incorporação em frentes de pesquisa. O impacto

social pode aparecer em documentos de política, diretrizes clínicas, normas técnicas, mídia, cursos, repositórios, patentes ou debates públicos. A passagem de um tipo de impacto a outro não é linear: trabalhos muito citados podem ter baixa repercussão pública, enquanto produtos técnicos podem ser socialmente relevantes sem ampla citação acadêmica (NARIN; HAMILTON; OLIVASTRO, 1997; HAUSTEIN *et al.*, 2014).

Colaboração como relação observável

A colaboração científica pode ser observada por coautoria, coafiliação, cofinanciamento, participação em projetos, compartilhamento de dados ou produção conjunta de infraestrutura. A coautoria é a medida mais comum porque está disponível nos metadados bibliográficos e permite construir redes de autores, instituições e países. Entretanto, coautoria não captura toda colaboração real e pode registrar relações administrativas, hierarquias laboratoriais ou autoria honorária (GLÄNZEL, 2001; WAGNER; LEYDESDORFF, 2005).

Colaboração internacional

A colaboração internacional é frequentemente usada como sinal de inserção global, capacidade de articulação e circulação transnacional do conhecimento. Indicadores de coautoria entre países permitem mapear centros, periferias, pontes, assimetrias e dependências em campos científicos. A análise deve evitar conclusões simplistas, pois áreas, políticas de financiamento, infraestrutura experimental e tradições disciplinares influenciam fortemente a taxa de colaboração internacional (GLÄNZEL, 2001).

Circulação científica

Circulação designa o percurso pelo qual produtos científicos se deslocam entre comunidades, línguas, plataformas, áreas, públicos e usos. Ela pode ser analisada por citações cruzadas, traduções, repositórios, downloads, recomendações, menções em mídias, reuso de dados e presença em documentos técnicos. A circulação amplia o olhar para além da publicação, mas torna a análise mais dependente de fontes heterogêneas e instáveis (THELWALL; VAUGHAN; BJÖRNEBORN, 2005; HAUSTEIN *et al.*, 2014).

Visibilidade e descobribilidade

Visibilidade científica não é sinônimo de impacto, embora possa facilitar o impacto futuro. Um artigo pode ser visível por estar em periódico de alto alcance, em acesso aberto, em repositório bem indexado, em redes sociais ou em resultados de busca. A visibilidade depende de infraestrutura informacional, metadados, língua, prestígio editorial, estratégias de comunicação e padrões algorítmicos de plataformas (HAUSTEIN *et al.*, 2014; WALTMAN, 2016).

Produção invisível

A cientometria precisa reconhecer que parte relevante da ciência permanece invisível aos sistemas de indexação dominantes. Publicações locais, literatura cinzenta, relatórios técnicos, dados de pesquisa, software, documentos em línguas não hegemônicas e contribuições comunitárias podem ser mal representados. Ao formular *prompts* para análise, convém solicitar explicitamente uma seção de limites de cobertura e vieses de base de dados (HICKS *et al.*, 2015; RAFOLS *et al.*, 2012).

Quadro 3.3 - As cinco dimensões cientométricas

1. PRODUÇÃO CIENTÍFICA como registro	2. IMPACTO como influência	3. COLABORAÇÃO como relação	4. CIRCULAÇÃO como trajetória	5. VISIBILIDADE como descobribilidade
Produtos formalmente registrados da atividade de pesquisa. *Ex.: artigos, revisões, anais, livros, capítulos, dados,* software, *patentes.*	Influência que um produto científico exerce em diferentes domínios. *Ex.: citações, uso em políticas, patentes, mídia, práticas profissionais.*	Relações de coprodução, cofinanciamento ou compartilhamento de recursos científicos. *Ex.: coautoria, coafiliação, projetos, dados, infraestrutura.*	Percurso dos produtos científicos entre comunidades, plataformas e diferentes tipos de uso. *Ex.: citações cruzadas, traduções, repositórios,* downloads, *reuso de dados.*	Condições que tornam o produto científico encontrável, acessível e percebido. *Ex.: acesso aberto, indexação, redes sociais, buscas, mídias.*
Indicadores possíveis				
• Número de publicações • Taxa de crescimento • Produtividade por autor, instituição ou país • Concentração (ex.: lei de Lotka) • Expansão de palavras-chave	• Citações (total, normalizadas, por idade) • h-index, m-index, FWCI, percentis • Citações em revisões e diretrizes • Impacto em documentos de política • Uso em patentes e normas técnicas	• Redes de coautoria (autores, instituições, países) • Taxa de colaboração (nº de autores por documento) • Colaboração internacional (países parceiros) • Cofinanciamento e projetos conjuntos • Centralidade, intermediação e densidade da rede	• Citações cruzadas entre áreas e domínios • Downloads e visualizações em repositórios • Menções em documentos técnicos e políticas • Reuso de dados, softwares e datasets • Traduções, edições e republicações	• Presença em buscadores e bases (indexação) • Acesso aberto, DOIs, metadados completos • Menções sociais (tweets, posts, blogs) • Visualizações e alcance em plataformas • Posição em rankings e diretórios
Dados necessários				
• Registros bibliográficos completos • Metadados de autoria e afiliação • Classificação por área, tema e período	• Base de citações adequada à área • Idade do documento (janela temporal) • Tipos de documento e contexto de citação	• Metadados de autoria e afiliação • Identificadores padronizados (ORCID, ROR) • Informações de financiamento e projetos	• Bases bibliográficas e repositórios • Plataformas de políticas, técnicas e patentes • Logs de acesso, reuso e rastros digitais	• Fontes de altmetria e redes sociais • Metadados de acesso e indexação • Dados de tráfego, menções e buscas
Limitações e vieses				
• Sub-representa livros, relatórios e produção técnica em muitas áreas • Variações de indexação e políticas editoriais • Incentivos institucionais podem inflar produção	• Citações ≠ aprovação ou qualidade intrínseca • Desigualdades entre áreas, línguas e bases • Função retórica das citações (negativa, crítica)	• Coautoria não captura toda colaboração real • Autoria honorária ou administrativa • Assimetria de poder nas redes	• Fontes heterogêneas e instáveis • Cobertura desigual por idioma e região • Difícil distinguir uso real de mera disponibilidade	• Visibilidade não é sinônimo de impacto • Algoritmos e estratégias de comunicação • Dependência de plataformas comerciais

Fonte: Elaborado pelo autor (2026).

IA na leitura das dimensões

Modelos de linguagem podem auxiliar a distinguir produção, impacto, colaboração, circulação e visibilidade em relatórios, mas tendem a fundir conceitos quando o *prompt* não estabelece fronteiras claras. Uma boa prática é pedir que a IA organize cada dimensão em definição, indicadores possíveis, dados necessários, limitações e exemplos de interpretação prudente. Essa decomposição reduz ambiguidade e preserva a diferença entre descrever uma evidência e atribuir valor científico a um resultado (SCHULHOFF *et al.*, 2024; JI *et al.*, 2023).

Essas cinco dimensões não são equivalentes. Produção indica existência de produtos; impacto sugere alguma forma de influência; colaboração descreve relações de coprodução; circulação aponta trajetórias de deslocamento; visibilidade indica condições de ser encontrado, lido ou comentado. Um erro comum é transformar uma dimensão em substituta das demais. Um campo pode crescer em produção sem ampliar impacto; pode tornar-se visível sem consolidar contribuição; pode colaborar internacionalmente sem reduzir dependências epistemológicas; pode circular socialmente sem alterar práticas ou políticas.

A IA pode ser útil justamente por permitir que o analista formule matrizes de distinção. Um *prompt* pode solicitar que, para cada dimensão, sejam indicadas fontes possíveis, indicadores adequados, perguntas de pesquisa, limitações e riscos de interpretação. Esse procedimento ajuda a disciplinar a análise e impede que o relatório use expressões vagas como "relevância", "influência" ou "excelência" sem demonstrar a evidência que sustenta cada termo. Em cientometria, vocabulário avaliativo exige lastro metodológico.

Ao mesmo tempo, a análise dessas dimensões deve considerar desigualdades estruturais. Língua de publicação, acesso a financiamento, infraestrutura laboratorial, indexação internacional, redes de colaboração e políticas editoriais afetam a presença de pesquisadores e instituições nos indicadores. Portanto, comparar produção, impacto, colaboração, circulação e visibilidade sem considerar contexto pode reforçar assimetrias já existentes. O uso de *prompts* deve incluir instruções explícitas para identificar vieses de cobertura e para evitar comparações injustas entre campos, países, instituições ou carreiras.

3.4 Unidade de análise: artigo, autor, periódico, instituição, país, área, conceito

Unidade de análise como decisão metodológica

A unidade de análise é a entidade sobre a qual recaem contagens, relações, comparações e interpretações. Em cientometria, a escolha entre artigo, autor, periódico, instituição, país, área ou conceito define o tipo de pergunta possível e o tipo de erro provável. Quando a unidade não é explicitada, indicadores podem ser calculados corretamente e ainda assim responder a uma pergunta diferente daquela formulada pelo pesquisador (WALTMAN, 2016; HICKS *et al.*, 2015).

Artigo como unidade documental

O artigo é a unidade mais comum em bases bibliográficas porque reúne título, autores, afiliações, resumo, palavras-chave, periódico, referências e citações. Ele permite análises de produtividade, impacto, temas, redes de citação, acoplamento bibliográfico e cocitação. Mesmo assim, o artigo não representa sozinho todo o processo científico: ele é um produto final, estabilizado, publicado e dependente de normas editoriais (GARFIELD, 1955; PRICE, 1965).

Autor como unidade biográfica e social

O autor permite estudar trajetórias, produtividade individual, colaboração, influência e especialização temática. Indicadores por autor incluem número de publicações, citações,

índice h, centralidade em redes, diversidade temática e colaboração internacional. A análise de autores é sensível a homônimos, mudanças de nome, variações de assinatura, identificadores incompletos e diferenças de carreira entre campos (HIRSCH, 2005; EGGHE, 2006).

Periódico como unidade editorial

O periódico organiza comunidades, estabelece filtros de publicação, delimita públicos e opera como canal de legitimação. Indicadores de periódico, como fator de impacto, CiteScore ou redes de cocitação, podem revelar posições em sistemas de comunicação científica. O erro clássico é transferir automaticamente o prestígio médio do periódico para a qualidade de cada artigo publicado nele (DECLARATION ON RESEARCH ASSESSMENT, 2013; HICKS *et al.*, 2015).

Instituição como unidade organizacional

A instituição permite analisar capacidade produtiva, perfil temático, colaboração, internacionalização, impacto médio e redes de parceria. Ela é relevante para gestão da pesquisa, planejamento estratégico e avaliação de programas. A análise institucional exige desambiguação cuidadosa de afiliações, pois nomes, siglas, unidades, hospitais, laboratórios e campi podem aparecer de forma inconsistente nas bases (ARIA; CUCCURULLO, 2017; MINGERS; LEYDESDORFF, 2015).

País como unidade geopolítica

O país é usado em estudos comparativos de sistemas científicos, cooperação internacional e inserção global. Indicadores por país permitem observar produção, impacto, especialização relativa, dependência de colaboração externa e redes de circulação. Entretanto, atribuir publicações a países envolve escolhas de contagem inteira ou fracionária e pode ocultar desigualdades internas entre regiões e instituições (GLÄNZEL, 2001; WALTMAN, 2016).

Área do conhecimento como unidade classificatória

A área do conhecimento é uma unidade construída por sistemas de classificação, categorias de periódicos, palavras-chave, algoritmos ou taxonomias institucionais. Ela permite normalizar citações, comparar campos e delimitar especialidades. Porém, áreas são fronteiras móveis; trabalhos interdisciplinares podem ser mal classificados, e rankings disciplinares podem penalizar pesquisas que circulam entre domínios (RAFOLS *et al.*, 2012; HICKS *et al.*, 2015).

Conceito como unidade cognitiva

Conceitos, termos e temas permitem analisar a estrutura intelectual de uma área por meio de palavras-chave, títulos, resumos, tópicos e coocorrências. Redes de termos e mapas temáticos ajudam a identificar frentes de pesquisa, subáreas, mudanças terminológicas e temas emergentes. A análise conceitual exige cuidado semântico: termos iguais podem ter sentidos diferentes, e termos diferentes podem nomear fenômenos equivalentes (SMALL, 1973; CHEN, 2006).

Quadro 3.4 - Unidades de análise: artigo, autor, periódico, instituição, país, área, conceito

1. ARTIGO *Unidade documental*	2. AUTOR *Unidade biográfica e social*	3. PERIÓDICO *Unidade editorial*	4. INSTITUIÇÃO *Unidade organizacional*
O artigo reúne título, autores, afiliações, resumo, palavras-chave, periódico, referências e citações.	O autor permite estudar trajetórias, produtividade, influência, colaboração e especialização temática.	O periódico organiza comunidades, estabelece filtros de publicação e legitima conhecimento.	A instituição reflete capacidade produtiva, perfil temático e redes de colaboração.
Indicadores possíveis			
• Número de publicações • Citações recebidas • Impacto normalizado • Coautorias • Temas (palavras-chave, tópicos) • Redes de citação e acoplamento	• Número de publicações • Citações e índice h • Produtividade por período • Colaboração (nº de coautores, redes) • Diversidade temática • Colaboração internacional	• Fator de impacto / CiteScore • Índice h do periódico • Rede de cocitação de periódicos • Especialização temática • Percentil e quartil (JCR, Scopus) • Tempo de meia-vida das citações	• Número de publicações • Citações totais e normalizadas • Colaboração nacional e internacional • Impacto médio por publicação • Diversificação temática • Parcerias institucionais
Principais usos			
• Crescimento da literatura • Perfis temáticos • Impacto de documentos • Estrutura intelectual do campo	• Trajetórias e carreiras • Redes de colaboração • Lideranças e comunidades • Especialização e interdisciplinaridade	• Estrutura de canais de comunicação • Núcleos de periódicos • Posicionamento de publicações • Avaliação de coleções e acervos	• Gestão e planejamento • Avaliação institucional • Benchmarking e rankings • Identificação de competências
Riscos e limites			
• Reflete apenas o produto publicado • Variações por tipo de documento • Dependência da cobertura da base	• Homônimos e variações de nome • Mudança de afiliação ao longo do tempo • Diferenças entre áreas e culturas	• Prestígio do periódico ≠ qualidade • Diferenças entre áreas e políticas • Mudanças de escopo e editoria	• Inconsistências de afiliações • Fusões, renomeações e unidades • Comparações injustas entre áreas

5. PAÍS *Unidade geopolítica*	6. ÁREA DO CONHECIMENTO *Unidade classificatória*	7. CONCEITO *Unidade cognitiva*
O país permite analisar sistemas científicos, cooperação e inserção internacional.	A área permite normalizar e comparar campos, especialidades e interdisciplinaridade.	Conceitos, termos e temas revelam a estrutura intelectual por meio de vocabulários e significados.
Indicadores possíveis		
• Produção científica • Citações totais e normalizadas • Colaboração internacional • Especialização relativa • Grau de dependência colaborativa • Impacto por milhão de habitantes	• Produção por área • Citações normalizadas por área • Diversidade interdisciplinar • Estrutura de campos (redes) • Fronteiras e sobreposições • Evolução temática por área	• Frequência de termos • Coocorrência de palavras-chave • Redes semânticas • Tópicos (modelos de linguagem) • Emergência de temas • Mudança terminológica
Principais usos		
• Comparações internacionais • Políticas de ciência e tecnologia • Inserção global e diplomacia científica • Fluxos de colaboração	• Comparação entre campos • Mapas de especialidades • Análise de fronteiras • Políticas de fomento por área	• Mapas temáticos • Identificação de frentes • Evolução de conceitos • Descoberta de lacunas

5. PAÍS *Unidade geopolítica*	6. ÁREA DO CONHECIMENTO *Unidade classificatória*	7. CONCEITO *Unidade cognitiva*
	Riscos e limites	
• Contagem inteira vs. fracionária • Desigualdades internas ignoradas • Dependência da cobertura da base	• Classificações imperfeitas • Interdisciplinaridade mal captada • Mudanças nas taxonomias	• Ambiguidade semântica • Sinônimos e polissemia • Dependência dos dados usados

Fonte: Elaborado pelo autor (2026).

Níveis e agregações

Um mesmo estudo pode operar em múltiplos níveis: artigos compõem autores, autores compõem instituições, instituições compõem países e documentos compõem áreas. A agregação permite síntese, mas também produz perda de informação e risco de falácia ecológica. *Prompts* usados para apoiar a análise devem exigir que a IA identifique o nível de cada afirmação e não generalize resultados individuais para instituições ou países sem evidência (WALTMAN, 2016; HICKS *et al.*, 2015).

Desambiguação e qualidade de dados

A escolha da unidade de análise exige operações de limpeza e desambiguação, especialmente para autores, afiliações, periódicos e palavras-chave. Identificadores como DOI, ORCID, ISSN, ROR e identificadores de bases ajudam, mas não eliminam inconsistências. A IA pode sugerir agrupamentos e detectar variantes, mas a validação humana continua necessária para evitar fusões indevidas ou separações artificiais (ARIA; CUCCURULLO, 2017; JI *et al.*, 2023).

A escolha da unidade de análise deve preceder a escolha do software, do indicador e da visualização. Muitos problemas cientométricos surgem quando o pesquisador inicia pela ferramenta disponível e só depois tenta adaptar uma pergunta. Em uma abordagem robusta, a pergunta define a unidade; a unidade define o dado necessário; o dado define a estratégia de coleta; e a coleta condiciona os indicadores possíveis. *Prompts* bem construídos podem ajudar a manter essa cadeia lógica, desde que contenham instruções sobre granularidade, agregação e limites.

Unidades também têm diferentes temporalidades. Artigos acumulam citações ao longo dos anos; autores constroem trajetórias; periódicos mudam políticas e escopos; instituições passam por fusões, renomeações e reestruturações; países alteram sistemas de financiamento; áreas se reconfiguram; conceitos emergem, desaparecem ou mudam de significado. Uma análise cientométrica que ignora temporalidade corre o risco de congelar entidades dinâmicas. Por isso, ao solicitar apoio da IA, convém exigir que a interpretação indique a janela temporal e a estabilidade da unidade analisada.

Outro ponto crucial é a passagem entre níveis. Resultados em nível de artigo não autorizam conclusões automáticas sobre autores; resultados em nível de autor não devem ser generalizados para instituições; métricas de periódico não devem substituir avaliação de documentos; e padrões nacionais não explicam sozinhos trajetórias individuais. A IA pode ser induzida a cometer essas generalizações quando o *prompt* pede conclusões rápidas. Um *prompt* cientométrico responsável deve incluir: "não generalize entre níveis de análise sem justificar a passagem".

3.5 Indicadores de produtividade, impacto e colaboração

Indicadores como representações parciais

Indicadores cientométricos são representações parciais de fenômenos complexos. Eles transformam documentos, relações e eventos em números comparáveis, mas sempre dependem de decisões de coleta, normalização, contagem e interpretação. A pergunta adequada não é se um indicador é verdadeiro em termos absolutos, mas para que ele serve, em que condições é válido e que aspectos deixa de fora (WALTMAN, 2016; HICKS *et al.*, 2015).

Contagem de publicações

A contagem de publicações é o indicador de produtividade mais simples e também um dos mais utilizados. Ela permite observar volume de produção por período, crescimento de uma área, participação de instituições e concentração de autores. Seu limite é evidente: publicar mais não significa necessariamente produzir conhecimento mais relevante, original ou socialmente útil (LOTKA, 1926; MINGERS; LEYDESDORFF, 2015).

Produtividade fracionária

A contagem fracionária distribui o crédito de uma publicação entre autores, instituições ou países participantes. Ela evita que cada unidade receba crédito integral por um mesmo artigo em colaborações extensas, tornando algumas comparações mais equilibradas. Contudo, a fração matemática pode não refletir a contribuição real de cada participante, especialmente em áreas com autoria alfabética ou colaboração laboratorial massiva (WALTMAN, 2016).

Taxas de crescimento

Taxas de crescimento anual, crescimento relativo e crescimento acumulado permitem examinar a expansão temporal de temas e áreas. Esses indicadores são úteis para detectar campos emergentes, estabilização de literaturas e efeitos de políticas científicas. A interpretação deve controlar alterações de base, mudanças terminológicas e efeitos de indexação, pois um aumento aparente pode decorrer de melhor cobertura documental (ARIA; CUCCURULLO, 2017; CHEN, 2006).

Citações brutas

O número bruto de citações é a medida mais difundida de impacto acadêmico, mas também uma das mais problemáticas quando usada isoladamente. Citações acumulam-se com o tempo, variam entre áreas e podem ter funções positivas, negativas, metodológicas ou meramente convencionais. Por isso, citações brutas são mais informativas quando acompanhadas de idade do documento, campo, tipo de publicação e distribuição das citações no conjunto analisado (BORNMANN; DANIEL, 2008; TAHAMTAN; BORNMANN, 2019).

Indicadores normalizados

Indicadores normalizados por campo e ano buscam tornar comparáveis documentos publicados em áreas e períodos com diferentes densidades de citação. A normalização reduz distorções quando se comparam áreas como biomedicina, matemática, ciências humanas e engenharia. Mesmo normalizados, indicadores permanecem dependentes da taxonomia de campos e da qualidade da atribuição disciplinar (WALTMAN, 2016).

Índice h

O índice h combina produtividade e impacto ao indicar que um pesquisador possui h publicações com pelo menos h citações cada. Sua força está na simplicidade e na resistência parcial a valores extremos; sua fraqueza está na dependência da idade da carreira, área, base de dados e padrões de citação. Ele não deve ser usado como medida única de qualidade individual, especialmente em comparações entre disciplinas ou estágios de carreira (HIRSCH, 2005; HICKS *et al.*, 2015).

Índice g e concentração de impacto

O índice g foi proposto para dar maior peso às publicações altamente citadas, respondendo a uma limitação do índice h. Ele permite distinguir perfis em que poucos trabalhos concentram influência expressiva. Como todo indicador de citação, continua sensível a campos, tempo, cobertura de base e práticas de autoria (EGGHE, 2006).

Fator de impacto e métricas de periódicos

Métricas de periódicos podem ser úteis para compreender canais de comunicação, mas são inadequadas como substitutos da avaliação de artigos ou pesquisadores. O fator de impacto representa uma média de citações de artigos recentes de um periódico e sofre com assimetrias de distribuição. A avaliação responsável recomenda não usar métricas de periódico como medida direta da qualidade de trabalhos individuais (DECLARATION ON RESEARCH ASSESSMENT, 2013; HICKS *et al.*, 2015).

Indicadores de colaboração

Indicadores de colaboração incluem número médio de autores por artigo, proporção de artigos em coautoria, colaboração institucional, colaboração internacional e medidas de rede. Redes de coautoria permitem calcular grau, centralidade, densidade, componentes e comunidades. A interpretação precisa reconhecer que coautoria é um *proxy*: ela registra uma relação documental, mas não esgota a colaboração cognitiva ou material (GLÄNZEL, 2001; VAN ECK; WALTMAN, 2010).

Indicadores de rede

Medidas de rede acrescentam uma camada relacional à análise cientométrica, permitindo distinguir volume de posição estrutural. Um autor pode publicar pouco e ainda assim ocupar papel de ponte entre comunidades; uma instituição pode não ser a maior produtora e ainda conectar subcampos. Centralidade, intermediação e modularidade devem ser interpretadas em conjunto com o desenho da rede, o critério de ligação e o método de normalização (KESSLER, 1963; SMALL, 1973; VAN ECK; WALTMAN, 2010).

Indicadores altmétricos

Indicadores altmétricos incluem menções, salvamentos, compartilhamentos, notícias, blogs, documentos de política e outras formas de atenção online. Eles podem captar sinais precoces de circulação e públicos distintos da comunidade acadêmica tradicional. Seu uso em avaliação requer cautela, pois plataformas mudam, métricas podem ser manipuladas e atenção social não equivale necessariamente a impacto científico ou benefício público (HAUSTEIN *et al.*, 2014; BORNMANN, 2015).

Indicadores responsáveis

A literatura de métricas responsáveis defende que indicadores devem apoiar, e não substituir, julgamento qualitativo especializado. Princípios como transparência, contextualização, diversidade de missões, abertura de dados e verificação de efeitos sistêmicos são essenciais. Em ambiente com IA, esses princípios tornam-se ainda mais relevantes, pois a automação pode tornar decisões avaliativas opacas e aparentemente objetivas (HICKS *et al.*, 2015; DECLARATION ON RESEARCH ASSESSMENT, 2013).

Seleção de indicadores por pergunta

A escolha de indicadores deve começar pela pergunta de pesquisa, não pela disponibilidade imediata de números. Perguntas sobre crescimento exigem séries temporais; perguntas sobre influência exigem citações ou uso; perguntas sobre integração exigem redes; perguntas sobre visibilidade exigem fontes de circulação. *Prompts* podem auxiliar a construir uma matriz pergunta-indicador-dado-limite, desde que a resposta seja conferida pelo pesquisador (ARIA; CUCCURULLO, 2017; SCHULHOFF *et al.*, 2024).

Indicadores não devem ser escolhidos por tradição ou facilidade de cálculo, mas por pertinência à pergunta. Contagem de publicações pode ser adequada para mapear crescimento; citações normalizadas podem apoiar comparações de impacto acadêmico; redes de coautoria podem revelar estrutura colaborativa; altmétricas podem indicar circulação inicial ou atenção social. Cada indicador possui uma zona de validade e uma zona de risco. A tarefa do pesquisador é saber onde termina a descrição e onde começa a extrapolação.

Em estudos assistidos por IA, é especialmente importante impedir que o modelo apresente indicadores como se fossem equivalentes. Um *prompt* de seleção de métricas deve exigir que a IA informe: o fenômeno medido, a fórmula ou lógica de cálculo, os dados necessários, a unidade de análise, as vantagens, as limitações, os vieses e a interpretação permitida. Essa estrutura transforma a IA em instrumento de comparação metodológica e reduz o risco de recomendações vagas, como “usar índice h” ou “usar fator de impacto”, sem justificativa.

O uso de indicadores em avaliação demanda critérios ainda mais rigorosos. Indicadores de produtividade podem estimular salami science; indicadores de impacto podem favorecer áreas com alta densidade de citação; métricas de periódicos podem penalizar interdisciplinaridade; indicadores altmétricos podem refletir popularidade efêmera. Assim, o capítulo assume uma posição metodológica clara: indicadores são necessários para compreender padrões, mas insuficientes para decidir mérito científico sem contexto, pluralidade de evidências e julgamento especializado.

3.6 A diferença entre medir, mapear, avaliar e interpretar a ciência

Medir como quantificar

Medir a ciência significa atribuir números a aspectos observáveis da atividade científica, como documentos, citações, autores, instituições, termos ou relações. A medição oferece comparabilidade, síntese e possibilidade de monitoramento temporal. Entretanto, toda medida resulta de uma operação de redução: ela seleciona uma dimensão do fenômeno e deixa outras fora do campo analítico (WALTMAN, 2016).

Mapear como representar relações

Mapear a ciência significa representar relações entre documentos, autores, conceitos, periódicos, instituições ou áreas. Mapas de cocitação, acoplamento bibliográfico, coautoria e coocorrência de termos permitem visualizar estruturas que seriam difíceis de perceber em listas ou tabelas. Mapas, porém, não são fotografias neutras da ciência; são produtos de escolhas de distância, normalização, limiares, algoritmos e rotulação (CHEN, 2006; VAN ECK; WALTMAN, 2010).

Avaliar como formular juízo

Avaliar a ciência implica produzir juízos sobre qualidade, relevância, desempenho, contribuição, mérito, prioridade ou impacto. Indicadores podem informar a avaliação, mas não a esgotam, porque avaliação envolve valores, objetivos institucionais, critérios disciplinares e consequências para pessoas e organizações. A passagem da medida para a avaliação deve ser explícita, justificada e acompanhada de revisão qualitativa (HICKS *et al.*, 2015; DECLARATION ON RESEARCH ASSESSMENT, 2013).

Interpretar como construir sentido

Interpretar é articular medidas, mapas, conhecimento substantivo da área, contexto histórico e hipóteses explicativas. A interpretação transforma resultados em argumentos, reconhecendo incertezas, limites de dados e alternativas plausíveis. Nesse ponto, a IA pode atuar como interlocutor analítico, mas não como autoridade final: ela pode sugerir leituras, não certificar significados (BORNMANN; DANIEL, 2008; JI *et al.*, 2023).

Relação entre as quatro operações

Medir, mapear, avaliar e interpretar formam uma sequência possível, mas não obrigatória. Um estudo pode medir sem avaliar, mapear sem ranquear, interpretar sem produzir decisão avaliativa ou avaliar com base em múltiplas evidências. Confundir essas operações é fonte frequente de injustiça: rankings podem ser apresentados como mapas, contagens como qualidade e visibilidade como impacto (RAFOLS *et al.*, 2012; HICKS *et al.*, 2015).

Exemplo de distinção aplicada

Considere um conjunto de artigos sobre inteligência artificial aplicada à saúde. Medir seria contar publicações por ano; mapear seria representar coocorrências de termos e redes de colaboração; avaliar seria julgar desempenho de instituições; interpretar seria explicar por que certos subtemas cresceram em determinado período. Cada operação requer dados, critérios e linguagem próprios (ARIA; CUCCURULLO, 2017; CHEN, 2006).

IA e risco de colapso semântico

Modelos de linguagem podem colapsar a diferença entre medir, mapear, avaliar e interpretar quando solicitados a produzir respostas executivas muito sintéticas. Por exemplo, podem converter uma maior centralidade em rede em suposta liderança científica sem evidência adicional. *Prompts* devem obrigar a IA a separar descrição empírica, inferência interpretativa, juízo avaliativo e recomendação prática (BROWN *et al.*, 2020; WEI *et al.*, 2022).

Da tabela ao argumento

Um relatório cientométrico de qualidade não termina em tabelas de indicadores. Ele precisa explicar o que cada indicador mostra, que pergunta responde, que pergunta não responde e quais evidências complementares seriam necessárias. Essa passagem da tabela ao argumento é uma competência central do analista e deve ser ensinada desde os fundamentos do campo (WALTMAN, 2016; BORNMANN; DANIEL, 2008).

Responsabilidade interpretativa

A responsabilidade interpretativa significa assumir que resultados cientométricos têm efeitos sobre reputações, recursos, políticas e trajetórias profissionais. Portanto, interpretações devem evitar determinismo métrico, conclusões causais indevidas e comparações descontextualizadas. Com IA, essa responsabilidade aumenta, pois textos fluentes podem dar aparência de segurança a interpretações frágeis (HICKS *et al.*, 2015; JI *et al.*, 2023).

Fundamento para os capítulos seguintes

A distinção entre medir, mapear, avaliar e interpretar é a base para os demais capítulos da trilha. Estratégias de busca, limpeza de dados, redes, mapas temáticos, análise temporal e relatórios executivos só produzem conhecimento confiável quando situados nessas diferenças. Os *prompts* apresentados a seguir foram elaborados para apoiar essa disciplina conceitual e impedir que a IA transforme indicadores em conclusões automáticas (SCHULHOFF *et al.*, 2024).

A distinção entre medir, mapear, avaliar e interpretar é uma salvaguarda contra o uso simplificador da cientometria. Medir é produzir uma quantidade; mapear é representar relações; avaliar é formular juízo orientado por critérios; interpretar é construir sentido a partir de evidências e contexto. Essas operações podem se complementar, mas não devem ser confundidas. Um mapa de rede não é um ranking; uma contagem não é uma avaliação; uma correlação não é uma explicação; uma visualização não é uma interpretação.

A IA generativa aumenta tanto o potencial quanto o risco dessa passagem. Por um lado, *prompts* podem ajudar a transformar resultados técnicos em narrativas compreensíveis, a formular hipóteses e a organizar relatórios. Por outro, modelos de linguagem podem produzir conclusões excessivas com tom confiante, sobretudo quando os dados de entrada são insuficientes. O pesquisador deve pedir ao modelo que separe descrição, inferência e recomendação, e deve revisar cada afirmação à luz dos dados originais.

Essa disciplina é fundamental para a proposta do livro. A trilha "Cientometria com inteligência artificial" não defende a automação da avaliação científica, mas a qualificação do trabalho analítico por meio de *prompts* documentados. O objetivo é ampliar a capacidade de formular perguntas, organizar dados, explicar indicadores, comparar métricas e interpretar mapas, preservando a responsabilidade humana. O capítulo seguinte poderá avançar para desenho de estudos e estratégias de busca, mas os fundamentos aqui apresentados devem permanecer como referência de validade.

Quadro 3.5 - Relação entre perguntas, unidades de análise e indicadores

Pergunta típica	Unidade	Indicadores possíveis	Visualização	Limite principal
Como uma área cresceu?	Artigos por ano	Contagem, taxa de crescimento, crescimento acumulado	Série temporal	Mudanças de indexação e vocabulário
Quem colabora com quem?	Autores ou instituições	Grau, densidade, centralidade, comunidades	Rede de coautoria	Coautoria não captura toda colaboração
Quais trabalhos influenciam o campo?	Artigos ou referências	Citações, cocitação, acoplamento bibliográfico	Rede ou ranking contextualizado	Citação não equivale a aprovação
Que temas emergem?	Termos ou conceitos	Coocorrência, frequência temporal, *burst* de termos	Mapa temático ou linha do tempo	Ambiguidade semântica
Qual a visibilidade social?	Artigos ou produtos	Menções, notícias, políticas, salvamentos	*Dashboard* de atenção	Atenção não é impacto comprovado

Fonte: Elaborado pelo autor (2026).

3.7 Prompts do capítulo

Os *prompts* a seguir foram elaborados para apoiar três operações recorrentes em estudos cientométricos: explicar indicadores a públicos não especialistas, comparar métricas e escolher indicadores adequados a uma pergunta de pesquisa. Eles devem ser utilizados com dados reais, com referências verificadas e com revisão humana da saída. Recomenda-se registrar a versão do modelo, a data de uso, o texto do *prompt*, os dados de entrada e as alterações realizadas pelo pesquisador após a resposta da IA.

Prompt 3.7.1 Prompt para explicar indicadores a públicos não especialistas[2]

Atue como especialista em Cientometria e comunicação pública da ciência.

Explique o indicador [NOME DO INDICADOR] para um público não especialista formado por [PÚBLICO].

A explicação deve conter: definição simples; o que o indicador mede; o que ele não mede; exemplo numérico intuitivo; interpretação correta; erros comuns de interpretação; uma analogia didática; e uma frase de cautela.

Use linguagem clara, sem perder precisão metodológica. Não apresente o indicador como medida direta de qualidade científica.

2 Comentário metodológico: este *prompt* é útil para traduzir indicadores sem reduzir sua complexidade. A saída deve ser conferida com a literatura especializada e adaptada ao público-alvo, sobretudo nos exemplos numéricos e nas frases de cautela.

Prompt 3.7.2 Prompt para comparar métricas[3]

> Atue como metodólogo em Cientometria. Compare as métricas [MÉTRICA 1], [MÉTRICA 2] e [MÉTRICA 3] em uma tabela com as seguintes colunas: fenômeno representado; unidade de análise; dados necessários; modo geral de cálculo; vantagens; limitações; vieses conhecidos; situações em que é adequada; situações em que deve ser evitada; e referências conceituais.
>
> Depois da tabela, escreva uma síntese crítica distinguindo uso descritivo, uso avaliativo e uso interpretativo.

Prompt 3.7.3 Prompt para escolher indicadores adequados a uma pergunta de pesquisa[4]

> Atue como consultor de desenho metodológico em Cientometria.
>
> A pergunta de pesquisa é: [PERGUNTA]. Proponha um conjunto de indicadores adequados, organizados em produtividade, impacto, colaboração, circulação e visibilidade. Para cada indicador, informe: justificativa; unidade de análise; fonte de dados recomendada; janela temporal; necessidade de normalização; visualização possível; principais limites; e verificações de qualidade.
>
> Não recomende métricas incompatíveis com a pergunta. Se a pergunta for vaga, proponha versões mais precisas antes de selecionar indicadores.

Aplicação didática: matriz de escolha de indicadores por prompts

Uma forma prática de usar os fundamentos deste capítulo é construir uma matriz de decisão antes da coleta de dados. O pesquisador pode iniciar com uma pergunta ampla, como "qual é a evolução da pesquisa brasileira sobre inteligência artificial na educação?", e solicitar à IA que decomponha a pergunta em dimensões mensuráveis. A resposta esperada não deve ser uma lista genérica de métricas, mas um quadro que relacione produtividade, impacto, colaboração, circulação e visibilidade com unidades de análise e fontes de dados. Essa matriz serve como anteprojeto metodológico, não como substituto do protocolo de pesquisa.

Em seguida, o pesquisador deve revisar cada indicador sugerido. Para produção, pode ser adequado contar publicações por ano e tipo documental; para impacto, usar citações normalizadas por área e idade; para colaboração, mapear coautoria entre instituições; para circulação, verificar repositórios e documentos de política; para visibilidade, observar menções e salvamentos. Cada escolha precisa ser justificada. Se a base de dados não cobre adequadamente periódicos nacionais ou literatura em português, esse limite deve aparecer no relatório. Se a área analisada possui forte tradição de conferências, anais não podem ser ignorados sem justificativa.

O papel dos *prompts* é tornar esse raciocínio mais explícito. Um bom *prompt* pode perguntar: "que indicadores são incompatíveis com esta pergunta?", "que vieses a base

---

3 Comentário metodológico: a comparação em tabela ajuda a explicitar diferenças entre métricas que frequentemente são usadas como se fossem equivalentes. O pesquisador deve revisar fórmulas, pressupostos e referências antes de incorporar a resposta ao relatório.

4 Comentário metodológico: este *prompt* orienta a escolha de indicadores a partir da pergunta, evitando a seleção por hábito ou disponibilidade. A resposta da IA deve ser tratada como anteprojeto metodológico e conferida com dados, fontes e objetivos do estudo.

escolhida pode introduzir?", "que normalizações são necessárias?", "que nível de análise cada métrica permite?" e "que interpretação seria indevida?". Com isso, a IA deixa de ser apenas ferramenta de redação e passa a funcionar como dispositivo de verificação metodológica. O resultado, contudo, só será cientificamente defensável se o pesquisador conferir dados, literatura e hipóteses.

A matriz também contribui para a comunicação com públicos não especialistas. Gestores, coordenadores de programas, pesquisadores de outras áreas e formuladores de políticas frequentemente demandam respostas claras, mas a clareza não pode significar simplificação indevida. Ao explicar indicadores, convém usar linguagem acessível e, ao mesmo tempo, indicar limites. Dizer que "um artigo foi citado 100 vezes" é menos informativo do que explicar se esse número é alto para a área, para o ano de publicação, para o tipo de documento e para a base utilizada. A IA pode ajudar a produzir essa tradução, desde que instruída a preservar cautelas.

Finalmente, a matriz de escolha de indicadores deve ser tratada como documento vivo. À medida que a coleta avança, podem surgir problemas de duplicidade, baixa cobertura, inconsistência de afiliações, mudança de vocabulário ou impossibilidade de calcular certa métrica. O pesquisador pode então usar *prompts* para revisar o plano, mas deve registrar alterações. A reprodutibilidade não exige que o estudo permaneça idêntico ao plano inicial; exige que mudanças sejam justificadas e documentadas. Essa prática conecta os fundamentos cientométricos à ética da pesquisa assistida por IA.

3.8 Considerações finais

Os fundamentos de Cientometria apresentados neste capítulo estabelecem uma base conceitual para o uso de inteligência artificial em análises de ciência. Bibliometria, cientometria, informetria, webometria e altmetria formam um conjunto de tradições que compartilham métodos quantitativos, mas diferem quanto ao objeto, às fontes, às perguntas e aos tipos de evidência mobilizados. Compreender essas diferenças é essencial para formular *prompts* rigorosos e evitar que a IA gere sínteses conceitualmente imprecisas.

Produção, impacto, colaboração, circulação e visibilidade são dimensões interdependentes, mas não intercambiáveis. A produção registra produtos; o impacto aponta relações de influência ou uso; a colaboração descreve vínculos de coprodução; a circulação evidencia deslocamentos; e a visibilidade indica condições de acesso e atenção. O uso responsável de indicadores exige que cada dimensão seja interpretada segundo sua própria lógica, seus dados e suas limitações.

A unidade de análise é uma decisão metodológica central. Artigos, autores, periódicos, instituições, países, áreas e conceitos não apenas organizam dados; eles definem o que pode ser afirmado. Erros de unidade produzem interpretações injustas ou logicamente inválidas, como avaliar pesquisadores por métricas de periódicos, comparar campos sem normalização ou atribuir liderança científica a atores apenas por centralidade visual em uma rede.

Indicadores de produtividade, impacto e colaboração devem ser escolhidos por pertinência à pergunta, e não por disponibilidade ou tradição. Contagens, índices, redes e métricas alternativas são úteis quando contextualizados, mas podem distorcer a avaliação quando usados isoladamente. A literatura de métricas responsáveis oferece princípios indispensáveis para que indicadores apoiem julgamento qualitativo, em vez de substituí-lo.

Por fim, distinguir medir, mapear, avaliar e interpretar é condição para uma Cientometria com IA metodologicamente madura. *Prompts* podem explicar, comparar, selecionar, organizar

e redigir; contudo, não devem converter descrições em juízos, mapas em rankings ou correlações em causalidades. A responsabilidade científica permanece humana. A IA deve ampliar a capacidade analítica, não obscurecer os pressupostos que sustentam a análise.

Capítulo 4

Inteligência Artificial aplicada à Cientometria

Figura 4 - Uso da IA na Cientometria

Fonte: Imagem gerada por IA, ChatGPT (2026).

4.1 Introdução

Este capítulo integra a trilha "Cientometria com Inteligência Artificial: *Prompts* para Analisar, Mapear e Interpretar a Ciência" e tem por objetivo examinar, em perspectiva conceitual e metodológica, como técnicas de inteligência artificial podem ser aplicadas ao ciclo completo de um estudo cientométrico. A ênfase recai sobre a IA generativa e os modelos de linguagem, mas a discussão inclui também aprendizagem de máquina, mineração de textos, modelos de tópicos, extração de entidades, classificação automática, sumarização e formas de apoio interpretativo. A questão central não é apenas saber se a IA pode executar tarefas úteis, mas sob quais condições essas tarefas permanecem cientificamente válidas, reprodutíveis, auditáveis e epistemicamente responsáveis.

A cientometria sempre dependeu de mediações técnicas. Bases de dados, tesauros, classificações disciplinares, algoritmos de contagem, índices de citação, técnicas de normalização e métodos de visualização já estruturavam a observação da ciência antes da popularização dos modelos generativos. A novidade contemporânea consiste na introdução de sistemas capazes de transformar linguagem natural em instruções operacionais, de resumir *corpora* extensos, de sugerir categorias analíticas, de redigir interpretações e de estabelecer diálogos exploratórios com o pesquisador. Essa mediação torna a análise mais acessível, mas também cria novos riscos: aparente fluência sem fundamento, atribuição indevida de causalidade, invenção de referências, perda da cadeia de evidência e confiança excessiva em respostas plausíveis (GARFIELD, 1955; PRICE, 1965; KESSLER, 1963; SMALL, 1973; VAN ECK; WALTMAN, 2010).

A arquitetura do capítulo acompanha a lógica de um estudo cientométrico aplicado. Primeiro, discute-se o que a IA pode fazer no processo de formulação, busca, coleta, limpeza, análise e comunicação de resultados. Em seguida, delimitam-se as tarefas que não devem ser delegadas sem validação humana. Na sequência, apresentam-se noções fundamentais sobre IA generativa, modelos de linguagem e análise textual. Depois, examinam-se usos específicos em classificação, sumarização, extração de entidades e apoio interpretativo. Por fim, são

discutidos riscos críticos, tais como alucinação, viés, falsa autoridade, erros de citação e perda de rastreabilidade. O capítulo encerra com *prompts* operacionais que permitem explicitar incerteza, separar dado, inferência e hipótese, e auditar afirmações.

O texto assume uma posição deliberadamente equilibrada. A IA é tratada como infraestrutura cognitiva de apoio à pesquisa, não como substituta do método científico. Seu valor é mais alto quando opera sobre dados delimitados, quando recebe instruções explícitas, quando suas respostas são comparadas a fontes primárias e quando seus resultados são incorporados a protocolos de validação. Seu risco é mais alto quando é usada como autoridade epistêmica autônoma, especialmente em contextos de avaliação científica, revisão da literatura, atribuição de mérito, diagnóstico institucional ou formulação de políticas de ciência, tecnologia e inovação.

A aplicação de IA à cientometria deve ser compreendida como uma mudança de regime na relação entre dados, métodos e linguagem. Em abordagens tradicionais, o analista seleciona dados bibliográficos, processa metadados, calcula indicadores, produz visualizações e redige interpretações. Com modelos de linguagem, parte desse trabalho passa a ser mediada por interações em linguagem natural: pede-se ao sistema que proponha palavras-chave, explique padrões, rotule agrupamentos, resuma tendências, identifique riscos ou reescreva resultados para diferentes públicos. Essa mediação modifica a temporalidade do trabalho analítico, reduz barreiras técnicas e amplia possibilidades exploratórias, mas não elimina a necessidade de desenho metodológico explícito.

A literatura de mapeamento científico mostrou que ferramentas computacionais como VOSviewer, CiteSpace e bibliometrix apoiam a construção de redes, mapas e análises evolutivas de domínios científicos. A IA generativa não substitui esses ambientes; ela se articula a eles. Pode ajudar a formular hipóteses antes do cálculo, a interpretar comunidades depois da visualização, a explicar indicadores para públicos não especialistas e a transformar saídas técnicas em narrativas analíticas. Contudo, o mapa, o indicador e a interpretação permanecem dependentes de decisões humanas: quais documentos entrarão na amostra, quais bases serão usadas, quais normalizações serão aplicadas, quais limiares serão adotados e quais conclusões são justificáveis (CHEN, 2006; VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017).

A noção de *prompt*, nesse contexto, deve ser deslocada do campo do comando informal para o campo do instrumento metodológico. Um *prompt* cientométrico não é apenas uma pergunta dirigida a um chatbot; é uma peça do desenho de pesquisa. Ele define papel, tarefa, *corpus*, unidade de análise, restrições, formato de saída, critérios de incerteza e condições de validação. Quando documentado, permite reconstituir a interação analítica. Quando não documentado, torna-se um ponto opaco do procedimento. A cientometria com IA, portanto, exige uma cultura de documentação de *prompts* semelhante à documentação de scripts, parâmetros de software, estratégias de busca e critérios de inclusão (BROWN *et al.*, 2020; WEI *et al.*, 2022).

As seções seguintes partem de uma premissa simples: a IA pode ampliar a capacidade de análise da ciência, mas apenas quando integrada a protocolos de curadoria e controle de qualidade. Sem esse enquadramento, a velocidade de produção textual pode mascarar fragilidade metodológica. Com esse enquadramento, por outro lado, os modelos de linguagem podem contribuir para tornar estudos cientométricos mais transparentes, explicáveis, comparáveis e úteis para pesquisadores, gestores, bibliotecários, editores e formuladores de políticas científicas.

4.2 O que a IA pode fazer em um estudo cientométrico

Em um estudo cientométrico, a IA pode atuar desde a fase anterior à coleta dos dados até a comunicação final dos resultados. Antes da coleta, modelos de linguagem podem auxiliar na decomposição de um tema amplo em problemas de pesquisa, na identificação de termos relacionados, na distinção entre conceitos próximos e na formulação de perguntas empiricamente operacionalizáveis. Esse apoio é particularmente útil quando o pesquisador se encontra diante de uma literatura interdisciplinar, marcada por vocabulários heterogêneos, traduções imperfeitas e variações terminológicas. A IA pode sugerir sinônimos, siglas, expressões equivalentes e contextos disciplinares nos quais determinado conceito aparece com sentidos diferentes. A utilidade dessa etapa, porém, depende da transformação das sugestões em estratégias de busca testáveis em bases bibliográficas.

Na etapa de desenho do estudo, a IA pode apoiar a escolha de unidades de análise. Um mesmo problema pode ser observado por artigos, autores, periódicos, instituições, países, áreas temáticas, palavras-chave, referências citadas ou conceitos extraídos de textos. Ao solicitar que o modelo explicite as consequências de cada unidade, o pesquisador pode tornar visíveis decisões que muitas vezes permanecem implícitas. Por exemplo, uma análise de produção por instituição responde a uma pergunta diferente de uma análise de colaboração por coautoria; uma rede de coocorrência de palavras-chave revela proximidades semânticas distintas daquelas sugeridas por cocitação; uma série temporal de publicações não equivale a uma série de impacto ou de consolidação intelectual.

A IA também pode contribuir para construir e revisar estratégias de busca. Em vez de redigir isoladamente uma *string* booleana, o pesquisador pode pedir que o modelo identifique campos semânticos, variantes linguísticas, grafias alternativas, acrônimos e termos ambíguos. Em seguida, pode solicitar que cada termo seja classificado como central, periférico, amplo, específico ou potencialmente ruidoso. Essa prática não substitui testes em bases como Web of Science, Scopus, PubMed, Dimensions, Lens ou OpenAlex; ela funciona como apoio heurístico. A qualidade da estratégia final deve ser aferida por amostras de resultados recuperados, comparação com artigos sentinela e análise do equilíbrio entre sensibilidade e precisão.

Durante a coleta, a IA pode ajudar a planejar fluxos de trabalho e a traduzir procedimentos em checklists. Pode gerar roteiros para exportação de metadados, orientar a documentação de versões de bases, sugerir campos mínimos de coleta e indicar riscos associados à ausência de resumos, referências citadas ou identificadores persistentes. Quando APIs são utilizadas, modelos de linguagem podem auxiliar na escrita de pseudocódigo, na interpretação de documentação técnica e na construção de rotinas preliminares. Todavia, o dado coletado deve ser verificado por scripts, registros de logs e inspeção manual, pois o modelo pode sugerir parâmetros inexistentes, *endpoints* desatualizados ou procedimentos incompatíveis com a fonte consultada.

No pré-processamento, a IA pode apoiar tarefas que demandam julgamento linguístico e semântico. A normalização de palavras-chave, por exemplo, exige agrupar formas singulares e plurais, traduzir expressões, resolver abreviações, detectar variantes ortográficas e decidir quando dois termos aparentemente próximos devem permanecer separados. Modelos de linguagem são úteis para propor agrupamentos e justificar semanticamente suas escolhas. Ainda assim, a decisão final deve ser conservadora, sobretudo em campos nos quais termos semelhantes pertencem a tradições teóricas distintas. A fusão indevida de categorias pode apagar diferenças epistemológicas relevantes e produzir mapas conceitualmente simplificados.

Outra possibilidade importante é a desambiguação. Autores com nomes semelhantes, instituições que mudam de nome, departamentos reestruturados, afiliações multilíngues

e endereços incompletos criam dificuldades para a análise. A IA pode sugerir regras de padronização, gerar listas de equivalência e identificar casos suspeitos. Entretanto, desambiguação é uma tarefa de alto risco, pois erros pequenos podem alterar redes de colaboração e rankings institucionais. O uso responsável exige cruzamento com identificadores como ORCID, DOI, ROR, GRID, Wikidata ou registros institucionais, além de amostragem manual de qualidade.

Na fase de análise, modelos de IA podem apoiar a leitura interpretativa de resultados estatísticos e visuais. Diante de uma tabela de crescimento anual, podem sugerir hipóteses para picos e quedas. Diante de uma rede de coautoria, podem distinguir entre centralidade, intermediação e densidade. Diante de um mapa temático, podem propor nomes para *clusters* e identificar relações entre temas motores, emergentes ou declinantes. Esse apoio é valioso porque ajuda a transformar resultados técnicos em narrativa analítica. Contudo, as explicações geradas devem ser tratadas como hipóteses interpretativas e não como evidências autossuficientes.

A IA pode ainda favorecer a comunicação dos achados. Resultados cientométricos são frequentemente apresentados a públicos com graus variados de familiaridade metodológica. Um comitê de avaliação, um grupo de pesquisa, uma agência de fomento, uma biblioteca universitária e uma equipe de gestão estratégica têm necessidades comunicativas distintas. Modelos de linguagem podem adaptar o nível de explicação, produzir sumários executivos, elaborar glossários, redigir legendas analíticas e transformar resultados em recomendações cautelosas. Essa adaptação, entretanto, não deve simplificar a ponto de omitir incertezas, limitações e dependências metodológicas.

A principal contribuição da IA em estudos cientométricos talvez não esteja em automatizar uma etapa isolada, mas em criar continuidade entre etapas. Um mesmo conjunto de *prompts* pode acompanhar a pesquisa da pergunta inicial ao relatório final: primeiro transforma o tema em problema; depois sugere fontes; em seguida identifica indicadores; posteriormente interpreta padrões; por fim produz síntese para decisão. Essa continuidade, se documentada, aumenta a transparência do processo. Se não documentada, torna-se uma cadeia opaca de decisões distribuídas entre pesquisador e modelo.

É útil distinguir três modalidades de uso: IA como assistente operacional, IA como assistente analítico e IA como interlocutor crítico. Como assistente operacional, ela ajuda a organizar tarefas, formatar tabelas, gerar checklists e revisar linguagem. Como assistente analítico, sugere classificações, interpreta redes, resume padrões e propõe hipóteses. Como interlocutor crítico, é usada para questionar escolhas metodológicas, apontar vieses, identificar lacunas e exigir evidência. A terceira modalidade é especialmente relevante, pois desloca a IA de produtora de respostas para instrumento de controle da qualidade do raciocínio.

O uso da IA também pode ampliar a escalabilidade da leitura. Em estudos com milhares de resumos, é impraticável que o pesquisador leia cada documento com o mesmo grau de detalhe. Modelos de linguagem e técnicas de processamento de linguagem natural podem auxiliar na extração de objetivos, métodos, populações, tecnologias, domínios de aplicação e resultados declarados. Essa capacidade permite enriquecer matrizes bibliométricas tradicionais com variáveis textuais. Ao mesmo tempo, exige estratégias de validação por amostragem, avaliação intercodificador e comparação com classificações humanas, sobretudo quando os resultados orientam decisões institucionais.

Além disso, a IA pode facilitar a integração entre abordagens quantitativas e qualitativas. A cientometria historicamente mobiliza contagens, redes e indicadores, enquanto a revisão qualitativa examina argumentos, métodos, escolas e controvérsias. Modelos de linguagem

podem atuar na fronteira entre essas tradições, ajudando a transformar *clusters* bibliométricos em conjuntos interpretáveis de debates científicos. A promessa é reduzir a distância entre mapa e sentido. O risco é confundir rótulos gerados automaticamente com compreensão substantiva do campo. Por isso, todo rótulo produzido por IA deve ser confrontado com documentos representativos do *cluster*.

A IA pode auxiliar também na reprodutibilidade comunicacional. Um relatório cientométrico não deve apenas apresentar resultados; deve permitir que outros compreendam como o resultado foi produzido. *Prompts* podem gerar descrições metodológicas padronizadas a partir de um protocolo: base utilizada, data de coleta, *string* de busca, filtros, número de registros, critérios de exclusão, softwares, parâmetros, indicadores e versões. Essa função é valiosa porque a fragilidade de muitos estudos bibliométricos reside justamente na descrição incompleta do método. Entretanto, o modelo não deve inventar detalhes ausentes; ele deve solicitar lacunas ou assinalar campos não informados.

Por fim, a IA pode apoiar decisões prospectivas, mas com cautela. Ao analisar séries temporais, crescimento de temas, explosões de citação ou mudanças terminológicas, modelos podem sugerir cenários e implicações estratégicas. Em política científica, tais cenários podem orientar investimentos, formação de redes e monitoramento de áreas emergentes. Ainda assim, previsão em ciência é sempre condicionada por fatores sociais, institucionais, econômicos e tecnológicos. A IA deve ajudar a explicitar premissas, não a transformar tendências observadas em destino inevitável.

Quadro 4.1 - Funções possíveis da IA no ciclo cientométrico

Etapa	Uso possível da IA	Produto esperado	Validação necessária
Formulação	Transformar tema amplo em perguntas, hipóteses e unidades de análise.	Protocolo inicial e alternativas de desenho.	Revisão por especialista e verificação de viabilidade dos dados.
Busca e coleta	Sugerir termos, sinônimos, *strings* booleanas e checklists de metadados.	Estratégia de busca documentada.	Teste em bases, artigos sentinela e análise de ruído.
Pré-processamento	Normalizar palavras-chave, autores, instituições e categorias.	Tabelas de equivalência e listas de casos ambíguos.	Amostragem manual, identificadores persistentes e registro das decisões.
Análise	Apoiar leitura de indicadores, redes, *clusters* e séries temporais.	Hipóteses interpretativas e narrativas preliminares.	Confronto com resultados calculados, literatura e contexto disciplinar.
Comunicação	Redigir sumários executivos, glossários e explicações para públicos distintos.	Relatórios, legendas e sínteses interpretativas.	Revisão humana, checagem de fontes e preservação das limitações.

Fonte: Elaborado pelo autor (2026).

Planejamento e governança do estudo

A governança do estudo cientométrico assistido por IA começa antes da primeira interação com o modelo. É necessário definir finalidade, público, riscos, tipo de dado e critérios de decisão. Um estudo destinado a apoiar uma revisão exploratória admite graus de incerteza diferentes de um relatório usado para avaliação institucional. A mesma ferramenta pode ser

aceitável em contexto pedagógico e inadequada em decisão de financiamento. Assim, a pergunta não deve ser apenas "a IA consegue fazer?", mas 'a IA pode fazer com controle, evidência e responsabilidade suficientes para esta finalidade?'.

A documentação do estudo deve incluir um plano de uso da IA. Esse plano explicita em quais etapas a ferramenta será empregada, quais dados serão fornecidos, quais saídas serão aceitas, como ocorrerá a validação e quais tarefas permanecerão exclusivamente humanas. Tal plano reduz ambiguidades e oferece ao leitor condições de avaliar a contribuição do modelo. Sem ele, a menção genérica a "uso de IA" pouco informa sobre a qualidade metodológica do processo.

A governança também envolve privacidade e propriedade intelectual. Muitos estudos cientométricos usam dados públicos, mas podem incorporar anotações internas, relatórios estratégicos, bases institucionais ou documentos em desenvolvimento. A submissão desses materiais a serviços externos deve respeitar políticas institucionais e condições de uso. Quando houver dúvida, é preferível anonimizar dados, usar modelos locais ou limitar a IA a materiais já públicos e devidamente citáveis.

Qualidade dos dados e metadados

A qualidade da saída de IA é condicionada pela qualidade do dado de entrada. Em cientometria, registros incompletos, duplicados, mal indexados ou com metadados inconsistentes geram interpretações frágeis. Modelos de linguagem podem suavizar essas inconsistências em narrativas aparentemente coesas. Por isso, antes de pedir interpretação, o pesquisador deve verificar número de registros, campos ausentes, duplicatas, distribuição temporal e cobertura por fonte. A IA pode auxiliar nessa verificação, mas não deve mascarar problemas.

Metadados bibliográficos têm natureza institucional. Eles são produzidos por autores, editores, indexadores e bases. Cada campo carrega decisões e limitações. Palavras-chave de autores refletem estratégias de visibilidade; palavras-chave indexadas refletem vocabulários controlados; afiliações refletem formatos heterogêneos; referências citadas refletem práticas disciplinares. A IA deve ser instruída a considerar essa heterogeneidade quando interpreta padrões. Caso contrário, corre o risco de tratar todos os campos como equivalentes.

A normalização é uma etapa crítica para o uso de IA. Se palavras-chave não foram unificadas, a classificação automática pode multiplicar categorias redundantes. Se instituições não foram desambiguadas, a interpretação de colaboração pode ser distorcida. Se autores homônimos não foram tratados, indicadores individuais podem ser inflados ou fragmentados. Em cada uma dessas situações, *prompts* podem ajudar a sugerir regras, mas a aplicação final deve ser auditável.

Integração com softwares cientométricos

A IA generativa deve ser vista como complemento, e não como substituta, de softwares cientométricos. Ferramentas como VOSviewer, CiteSpace e bibliometrix calculam redes, medidas e mapas com procedimentos explícitos. Modelos de linguagem podem interpretar, explicar e revisar tais resultados, mas não devem gerar métricas sem base calculada. A separação entre cálculo e interpretação é essencial para preservar a confiabilidade do estudo.

Uma integração produtiva ocorre quando o pesquisador exporta resultados de softwares cientométricos e os submete à IA com contexto metodológico. Por exemplo, pode fornecer lista de *clusters*, termos mais frequentes, documentos centrais e parâmetros do mapa. O modelo então sugere rótulos e hipóteses, mas a validade da visualização continua dependente dos

dados e algoritmos originais. Essa arquitetura mantém o cálculo em ferramentas apropriadas e usa a IA para ampliar a inteligibilidade.

Outra integração consiste em usar IA para traduzir resultados técnicos em explicações acessíveis. Indicadores como centralidade de intermediação, modularidade, acoplamento bibliográfico e normalização de citações podem ser difíceis para públicos não especialistas. *Prompts* bem construídos permitem adaptar a explicação a gestores, estudantes ou pesquisadores de outras áreas, sem abandonar ressalvas metodológicas. A comunicação torna-se mais inclusiva, mas deve preservar a precisão conceitual.

Avaliação e métricas responsáveis

A aplicação de IA a estudos de avaliação científica deve observar princípios de métricas responsáveis. Indicadores não podem ser tratados como medidas absolutas de qualidade. Eles são sinais parciais, condicionados por área, idioma, tempo de circulação, tipo documental, cultura de citação e cobertura de base. Modelos generativos podem ajudar a explicar essas limitações, mas também podem reforçar interpretações simplistas se forem solicitados a produzir rankings e juízos de valor sem contexto.

Um uso apropriado da IA em avaliação é a análise crítica de indicadores. O pesquisador pode pedir que o modelo identifique o que determinado indicador mede, o que não mede, em quais áreas é mais ou menos adequado e quais indicadores complementares deveriam ser considerados. Esse tipo de *prompt* transforma a IA em ferramenta de prudência metodológica. Em vez de legitimar um número, ela ajuda a contextualizá-lo.

A IA também pode apoiar a elaboração de pareceres mais equilibrados. Um relatório institucional pode combinar produtividade, impacto, colaboração, formação de redes, diversidade temática, internacionalização e contribuição local. Modelos de linguagem podem auxiliar na composição da narrativa, desde que cada afirmação esteja ancorada em dados. A avaliação responsável combina indicadores e julgamento qualitativo; a IA deve apoiar ambos sem substituir nenhum.

Formação e alfabetização em IA

A adoção de IA em cientometria exige formação específica. Não basta ensinar usuários a escrever comandos; é necessário formar analistas capazes de compreender bases de dados, indicadores, modelos de linguagem, vieses e validação. A alfabetização em *prompts* deve incluir noções de contexto, restrição, formato de saída, controle de incerteza e auditoria. Também deve ensinar a reconhecer respostas plausíveis, mas não fundamentadas.

A formação deve enfatizar que a IA não é uma fonte bibliográfica. Ela pode auxiliar a localizar ideias, formular perguntas e revisar textos, mas referências, dados e afirmações factuais devem ser confirmados. Essa distinção é crucial para estudantes e pesquisadores em início de carreira, que podem ser atraídos pela velocidade da ferramenta. A competência científica inclui saber quando não usar a IA, ou quando interromper a interação para verificar fontes.

Programas de pós-graduação, bibliotecas universitárias e núcleos de apoio à pesquisa podem desempenhar papel decisivo. Oficinas de cientometria com IA devem combinar exercícios práticos de *prompts* com avaliação crítica de erros. Um bom treinamento não exibe apenas respostas corretas; mostra falhas, alucinações, vieses e estratégias de mitigação. O objetivo é formar usuários reflexivos, não operadores automáticos.

Quadro 4.2 - A IA como assistente metodológico

1. Planejamento e Governança do Estudo	2. Qualidade dos Dados e Metadados	3. Integração com Softwares Cientométricos	4. Avaliação e Métricas Responsáveis	5. Formação e Alfabetização em Ia
• Definir finalidade, público, perguntas e decisões esperadas. • Delimitar unidades de análise, fontes e indicadores. • Identificar riscos, limitações e requisitos de qualidade. • Elaborar plano de uso da IA: etapas, dados, saídas aceitas, validação e tarefas exclusivamente humanas. • Considerar privacidade, propriedade intelectual e políticas institucionais.	• Verificar número de registros, campos ausentes, duplicatas e cobertura. • Analisar distribuição temporal, por fontes, áreas e tipos documentais. • Sugerir regras de normalização (autores, instituições, palavras-chave). • Detectar inconsistências, valores atípicos e padrões incomuns.	• Interpretar resultados de redes, mapas e indicadores produzidos por softwares. • Explicar decisões (limiares, modularidade, acoplamento etc.) para diferentes públicos. • Sugerir rótulos para *clusters* e tópicos. • Traduzir achados técnicos em narrativas claras e contextualizadas.	• Analisar o que cada indicador mede e o que ele não mede. • Contextualizar por área, tempo, idioma, tipo documental e base. • Combinar indicadores de produtividade, impacto, colaboração, circulação e visibilidade. • Apoiar pareceres equilibrados e baseados em múltiplas evidências.	• Ensinar a construção de *prompts* eficazes. • Desenvolver leitura crítica das saídas. • Reconhecer vieses, alucinações e limitações dos modelos. • Distinguir texto gerado de informação verificada. • Promover uso ético, responsável e transparente.
O que a IA pode apoiar				
• Gerar planos metodológicos e checklists. • Sugerir perguntas e hipóteses alternativas. • Mapear riscos e estratégias de mitigação. • Documentar decisões e responsabilidades.	• Gerar diagnósticos de qualidade dos dados. • Sugerir dicionários de normalização. • Propor padrões e regras de limpeza. • Indicar casos ambíguos para revisão humana.	• Explicar métricas e algoritmos. • Sugerir rótulos e subtemas com justificativas. • Escrever descrições e legendas analíticas. • Adaptar a linguagem a gestores ou estudantes.	• Explicar limitações e vieses dos indicadores. • Construir matrizes indicador × pergunta. • Sugerir indicadores complementares. • Apoiar a redação de pareceres e relatórios.	• Gerar materiais didáticos e exemplos. • Simular exercícios e casos práticos. • Explicar conceitos de IA e de métricas. • Apoiar feedback e autoavaliação.
Atenção				
—	Aplicar regras e validar continua sendo tarefa humana.	Cálculos e métricas devem ser feitos por ferramentas apropriadas.	Indicadores não substituem julgamento especializado nem contexto.	A IA não é fonte bibliográfica. Sempre confirmar fatos, dados e referências.

Fonte: Elaborado pelo autor (2026).

Reprodutibilidade e auditoria

A reprodutibilidade é um desafio particular quando se usam modelos de linguagem. Diferentemente de scripts determinísticos, *prompts* podem produzir variações. Além disso, plataformas comerciais são atualizadas sem controle do pesquisador. Por essa razão, a reprodutibilidade deve ser entendida como rastreabilidade robusta: registrar o suficiente para que outro pesquisador compreenda o procedimento, mesmo que a saída exata não seja reproduzida palavra por palavra.

Um apêndice de IA pode incluir versões de *prompts*, descrição do modelo, data de uso, dados de entrada, amostras de saída, critérios de validação e decisões humanas. Em estudos submetidos a periódicos, essa transparência facilita revisão por pares. Em relatórios institucionais, permite auditoria posterior. Em ensino, ajuda estudantes a compreenderem a diferença entre resultado gerado e resultado validado.

A auditoria deve ser planejada por amostragem e por criticidade. Tarefas de baixo risco, como reformular frases ou criar glossários, exigem revisão simples. Tarefas de alto risco, como classificar documentos, interpretar impacto ou sugerir política científica, exigem validação mais rigorosa. A intensidade da auditoria deve acompanhar o uso que será feito da saída. Quanto maior a consequência, maior o controle.

4.3 O que a IA não deve fazer sem validação humana

A primeira tarefa que a IA não deve executar sem validação humana é a definição final da pergunta de pesquisa. Modelos de linguagem são eficazes para gerar alternativas, mas a pertinência de uma pergunta cientométrica depende de finalidade, contexto institucional, disponibilidade de dados e consequências interpretativas. Uma pergunta aparentemente elegante pode ser inviável se a base escolhida não indexa adequadamente o campo, se o período é inadequado, se os nomes de instituições são ambíguos ou se os indicadores propostos não capturam a dimensão pretendida. A decisão final exige julgamento de domínio e responsabilidade metodológica (HICKS *et al.*, 2015).

A IA também não deve decidir autonomamente os limites de um *corpus*. Em cientometria, pequenas alterações na estratégia de busca podem modificar drasticamente o conjunto analisado. Incluir ou excluir um termo, adotar uma base em vez de outra, restringir por idioma ou selecionar apenas artigos de periódicos são escolhas que afetam representatividade, viés geográfico, cobertura disciplinar e visibilidade de comunidades científicas. Modelos generativos podem sugerir critérios, mas não devem substituí-los por critérios implícitos. O *corpus* é uma construção metodológica e deve ser justificado em função da pergunta.

Outra fronteira crítica é a atribuição de relevância. A IA pode resumir textos e apontar documentos supostamente centrais, mas centralidade bibliométrica não é sinônimo de relevância epistemológica, social ou estratégica. Um artigo muito citado pode ser controverso, metodologicamente problemático ou relevante por motivos externos ao mérito científico. Um trabalho pouco citado pode ser inovador, recente, regionalmente importante ou localizado em área de baixa densidade citacional. A validação humana deve contextualizar a métrica e evitar que o modelo apenas reproduza hierarquias existentes.

Modelos de linguagem não devem produzir referências bibliográficas sem verificação externa. A literatura recente documenta casos de referências fabricadas, autores inexistentes, títulos plausíveis mas falsos, DOIs inválidos e combinações erradas de periódico, volume e páginas. Em um capítulo ou relatório cientométrico, esse risco é particularmente grave, pois a bibliografia constitui parte da evidência. O uso responsável requer que cada referência seja conferida em fontes confiáveis, como páginas de periódicos, DOI resolver, Crossref, PubMed, Scopus, Web of Science, Google Scholar, ACL Anthology ou anais oficiais de congressos.

A IA tampouco deve interpretar mapas de rede como se visualizações fossem fotografias neutras da realidade científica. Mapas dependem de algoritmos de layout, normalização, limiares de corte, medidas de associação e decisões de agregação. Dois mapas produzidos com o mesmo *corpus*, mas com parâmetros diferentes, podem sugerir estruturas distintas. Modelos generativos podem atribuir sentido excessivo à proximidade espacial, ao tamanho de

nós ou à separação de *clusters*. A validação humana deve recolocar a visualização no âmbito do modelo analítico que a gerou.

O mesmo vale para classificações temáticas. Embora a IA seja útil para classificar resumos em categorias, ela não deve criar taxonomias finais sem avaliação especializada. Categorias podem ser conceitualmente sobrepostas, hierárquicas, historicamente situadas ou politicamente sensíveis. Em áreas interdisciplinares, a mesma expressão pode carregar diferentes significados. Em campos emergentes, rótulos estabelecidos podem ainda não existir. Por isso, classificações automáticas devem ser comparadas a amostras codificadas por humanos e avaliadas por métricas de concordância, além de revisão qualitativa dos casos ambíguos.

A IA não deve inferir causalidade a partir de associação bibliométrica. Crescimento de publicações não prova maturidade científica; aumento de citações não prova qualidade; coautoria não prova colaboração substantiva; proximidade em rede não prova afinidade teórica; coocorrência de palavras não prova integração conceitual. Modelos de linguagem, por sua fluência narrativa, podem converter correlação em explicação. Um protocolo robusto deve instruir explicitamente o sistema a distinguir descrição, inferência e hipótese, marcando a força da evidência disponível.

Outra limitação diz respeito à avaliação de pessoas e instituições. Indicadores cientométricos são frequentemente usados em contextos de carreira, financiamento, reputação e alocação de recursos. Delegar juízos avaliativos a IA sem validação aumenta riscos de injustiça, opacidade e reprodução de vieses. O Manifesto de Leiden estabelece a necessidade de apoiar avaliações quantitativas em julgamento qualitativo, proteger a excelência localmente relevante e reconhecer diferenças entre áreas. A IA deve servir para explicitar limites e alternativas, não para automatizar decisões de mérito.

A IA não deve substituir a leitura dos documentos nucleares de um campo. Em mapas de cocitação, acoplamento bibliográfico ou coocorrência de termos, certos documentos funcionam como referências estruturantes. A interpretação de uma comunidade temática exige leitura de trabalhos representativos, especialmente quando se pretende nomear *clusters* ou reconstruir trajetórias intelectuais. Resumos automáticos ajudam a priorizar leitura, mas não substituem o contato com argumentos, métodos, conceitos e evidências dos textos originais.

Também é inadequado usar IA para limpar dados sem registro das transformações efetuadas. A normalização de nomes, instituições e palavras-chave pode alterar o resultado analítico. Se o pesquisador aceita sugestões de agrupamento sem documentar regras e exceções, a reprodutibilidade é comprometida. A IA pode gerar tabelas de equivalência, mas cada substituição deve ser rastreável: termo original, termo normalizado, justificativa, fonte de apoio, nível de confiança e responsável pela decisão.

Outro ponto de cautela é a tradução automática de conceitos. Em cientometria internacional, termos circulam em diferentes idiomas. A IA pode traduzir palavras-chave, resumos e categorias, mas traduções literais podem apagar especificidades disciplinares. 'Impact', por exemplo, pode se referir a impacto citacional, impacto social, impacto tecnológico ou impacto de política pública. O mesmo ocorre com expressões como *'knowledge transfer', 'research assessment', 'innovation ecosystem'* ou *'responsible metrics'*. A tradução deve ser validada pelo contexto de uso, não apenas pela equivalência lexical.

Não se deve confiar na IA para identificar a atualidade de informações sem acesso a fontes recentes ou sem verificação. Modelos treinados em dados históricos podem apresentar conhecimentos desatualizados sobre bases, políticas editoriais, rankings, cobertura de periódicos ou mudanças institucionais. Em estudos cientométricos, a data da coleta e a versão

da fonte são elementos centrais. Assim, qualquer afirmação sobre estado atual de uma base, disponibilidade de API, política de indexação ou cobertura disciplinar deve ser confirmada em fonte primária ou documentação oficial.

A IA também não deve redigir conclusões normativas fortes quando os dados apenas permitem observações descritivas. É comum que relatórios de ciência e tecnologia sejam usados para recomendar investimentos, encerrar programas, priorizar áreas ou reorganizar grupos. Quando a IA transforma resultados em recomendações, deve ser instruída a indicar premissas, grau de confiança e alternativas. O pesquisador, por sua vez, deve verificar se as recomendações decorrem dos dados ou se são extrapolações retóricas.

Em síntese, a fronteira entre apoio e delegação é o eixo ético do uso da IA em cientometria. Apoiar significa ampliar a capacidade humana de formular, explorar, organizar, revisar e comunicar. Delegar sem validação significa transferir responsabilidade a um sistema que não compreende consequências sociais, institucionais ou epistemológicas de suas respostas. A cientometria com IA deve ser concebida como prática de coanálise supervisionada, na qual o modelo produz alternativas e o pesquisador preserva responsabilidade pela evidência, pelo método e pela interpretação.

Quadro 4.3 - Limites de delegação em estudos cientométricos

Tarefa	Risco da delegação automática	Condição mínima de uso	Responsável final
Definir pergunta de pesquisa	Perguntas elegantes mas inviáveis ou desalinhadas ao contexto.	Revisão do objetivo, dados disponíveis e finalidade do estudo.	Pesquisador ou equipe científica.
Construir corpus final	Viés de cobertura, ruído temático e exclusão indevida.	Teste de estratégia, critérios explícitos e amostra de validação.	Pesquisador e especialista de domínio.
Gerar referências	Referências inexistentes, DOIs falsos ou combinações bibliográficas erradas.	Checagem em fontes bibliográficas externas.	Autor do estudo.
Interpretar causalidade	Conversão de associação em explicação causal.	Separação entre descrição, inferência e hipótese.	Equipe analítica.
Avaliar pessoas ou instituições	Automação de juízos de mérito sem contexto.	Métricas responsáveis e julgamento qualitativo.	Comitê ou instância humana competente.

Fonte: Elaborado pelo autor (2026).

4.4 IA generativa, modelos de linguagem e análise textual

A IA generativa designa um conjunto de técnicas capazes de produzir novos conteúdos, como textos, códigos, imagens, tabelas, resumos e classificações, a partir de padrões aprendidos em grandes volumes de dados. No contexto deste capítulo, o foco recai sobre modelos de linguagem, isto é, sistemas treinados para prever e gerar sequências textuais. A emergência dos *Transformers*, apresentados por Vaswani e colaboradores, reconfigurou o campo ao permitir modelagem eficiente de dependências de longo alcance por mecanismos de atenção. A partir desse avanço, modelos como BERT, GPT, T5, SciBERT e SPECTER ampliaram a capacidade de representar textos científicos, executar tarefas de NLP e adaptar-se a instruções em linguagem natural (VASWANI *et al.*, 2017; BROWN *et al.*, 2020).

Modelos de linguagem podem ser classificados, de forma simplificada, em modelos voltados principalmente à representação e modelos voltados à geração. Modelos como BERT e SciBERT são frequentemente utilizados para produzir representações contextualizadas de textos, úteis em classificação, recuperação de informação e extração de entidades. Modelos generativos, como os baseados na família GPT, são treinados para produzir texto coerente em resposta a instruções. Em aplicações cientométricas, essa distinção importa: nem toda tarefa exige geração livre. Em muitos casos, *embeddings*, classificadores supervisionados ou modelos de tópicos oferecem saídas mais controláveis do que respostas textuais abertas.

A análise textual em cientometria antecede a IA generativa. Modelos estatísticos como LSA, LDA e técnicas de mineração de termos já eram empregados para identificar temas, proximidades semânticas e estruturas conceituais. A IA contemporânea acrescenta novas possibilidades: representações semânticas profundas, sumarização contextual, classificação por instrução, extração de entidades complexas e geração de interpretações. O salto qualitativo está na interface em linguagem natural, que permite ao pesquisador formular tarefas sem programar diretamente cada etapa. Entretanto, a facilidade de interação não elimina a necessidade de compreender o que está sendo estimado, inferido ou gerado (DEERWESTER *et al.*, 1990; BLEI; NG; JORDAN, 2003).

Em estudos cientométricos, títulos, resumos, palavras-chave e referências citadas compõem um material textual heterogêneo. Títulos são condensados e estratégicos; resumos são narrativas de objetivos e resultados; palavras-chave combinam vocabulário dos autores e vocabulário indexado; referências citadas revelam bases intelectuais; textos completos contêm métodos, dados, limitações e discussão. Modelos de linguagem podem processar cada uma dessas camadas, mas a interpretação deve considerar suas diferenças. Um tema ausente de palavras-chave pode estar presente no resumo; uma técnica citada na metodologia pode não ser central ao argumento; uma referência citada pode representar crítica, uso metodológico ou filiação teórica.

A IA generativa é particularmente sedutora porque produz continuidade discursiva. Diante de uma tabela fragmentada ou de um mapa complexo, ela oferece uma narrativa. Essa narrativa pode ser útil para iniciar a interpretação, mas também pode criar falsa coerência. A geração textual não é evidência; é uma hipótese discursiva sobre evidências fornecidas. Em cientometria, portanto, o pesquisador deve distinguir três camadas: a camada dos dados bibliográficos, a camada dos resultados calculados e a camada da interpretação textual. O modelo pode operar nas três, mas cada uma possui critérios distintos de validação.

Um conceito central para o uso de modelos de linguagem é o contexto. A resposta depende dos dados fornecidos, das instruções, da ordem das informações, do formato solicitado, dos exemplos e das restrições. Em tarefas cientométricas, *prompts* vagos tendem a produzir generalidades. *Prompts* robustos especificam *corpus*, período, base, unidade de análise, indicadores, objetivo, público, limites e formato. Por exemplo, pedir "interprete este mapa" é insuficiente; pedir 'interprete uma rede de coocorrência de palavras-chave gerada por associação total, com limiar mínimo de cinco ocorrências, em artigos de 2014 a 2025 sobre energia renovável' aumenta a rastreabilidade da resposta.

Outra distinção importante ocorre entre IA generativa autônoma e IA ancorada em recuperação de informação. Em abordagens de geração aumentada por recuperação, o modelo responde com base em documentos previamente recuperados, trechos citáveis ou bases controladas. Para cientometria, esse paradigma é preferível quando o objetivo é reduzir alucinações e vincular afirmações a evidências. A IA pode, por exemplo, resumir apenas os resumos de um *cluster*, rotular comunidades com base em documentos centrais ou comparar

categorias usando trechos explicitamente fornecidos. Quanto mais ancorada a resposta estiver em dados observáveis, menor o risco de invenção (LEWIS *et al.*, 2020).

Os modelos de linguagem também podem ser integrados a *pipelines* computacionais. Uma análise pode começar com coleta por API, limpeza por scripts, cálculo de indicadores em bibliometrix, visualização em VOSviewer e interpretação assistida por IA. Alternativamente, *embeddings* podem ser gerados para agrupar documentos, e os rótulos dos grupos podem ser refinados por *prompts*. Essa integração exige separar tarefas determinísticas, estatísticas e generativas. Contagens e normalizações devem ser reproduzíveis; classificações probabilísticas devem ser avaliadas; interpretações generativas devem ser revisadas.

No campo de textos científicos, modelos especializados como SciBERT e representações como SPECTER demonstram a importância de treinar ou ajustar modelos em *corpora* acadêmicos. A linguagem científica possui convenções próprias: nomes de métodos, entidades técnicas, siglas, estruturas argumentativas, citações e padrões disciplinares. Modelos gerais podem não distinguir adequadamente usos especializados de termos comuns. Em cientometria aplicada, essa limitação se manifesta na rotulagem de *clusters*, na extração de métodos e na identificação de objetos de estudo. Sempre que possível, é recomendável utilizar modelos, vocabulários ou bases adaptadas ao domínio analisado (BELTAGY; LO; COHAN, 2019; COHAN *et al.*, 2020).

A sumarização é uma das aplicações mais imediatas da IA generativa, mas também uma das mais perigosas. Resumir um conjunto de resumos não significa sintetizar criticamente uma área. O modelo pode enfatizar termos frequentes, suavizar controvérsias, omitir minorias temáticas e apresentar consenso onde há divergência. Em estudos cientométricos, a sumarização deve ser estratificada: resumo de documentos centrais, resumo por *cluster*, resumo por período, resumo por tipo de método e resumo de lacunas. Essa estratificação reduz a tendência a produzir sínteses homogêneas demais (CACHOLA *et al.*, 2020).

A análise textual assistida por IA pode enriquecer indicadores clássicos. Em vez de apenas contar publicações por ano, pode-se classificar cada artigo por método, objeto empírico, escala geográfica, tipo de dado ou tecnologia investigada. Em vez de apenas mapear coocorrência de palavras, pode-se identificar relações entre conceitos, problemas e aplicações. Em vez de apenas medir impacto por citação, pode-se examinar o contexto textual das citações, distinguindo uso confirmatório, metodológico, histórico ou crítico. Tais extensões aproximam cientometria, mineração de textos e estudos sociais da ciência (LUAN *et al.*, 2018; REIMERS; GUREVYCH, 2019).

No entanto, modelos generativos não possuem acesso garantido à verdade factual. Eles produzem respostas estatisticamente prováveis com base em padrões de treinamento e contexto. Mesmo quando parecem seguros, podem errar. Em ambientes científicos, essa propriedade exige procedimentos formais: exigir que o modelo marque incerteza, peça dados ausentes, não invente referências, diferencie evidência direta de suposição e forneça critérios de verificação. O *prompt*, portanto, deve conter cláusulas de segurança epistemológica.

A aplicação da IA à cientometria também depende de infraestrutura. Nem todos os modelos preservam dados, aceitam grandes contextos, permitem auditoria, oferecem controle de parâmetros ou garantem privacidade. Estudos com dados institucionais sensíveis, pareceres internos ou estratégias de pesquisa em desenvolvimento requerem atenção a governança de dados. A escolha entre ferramentas comerciais, modelos locais, plataformas acadêmicas e soluções abertas deve considerar segurança, custo, reprodutibilidade, transparência e requisitos éticos.

Em suma, IA generativa e análise textual ampliam o repertório da cientometria, mas impõem uma alfabetização metodológica adicional. O pesquisador precisa compreender que um modelo pode classificar sem explicar adequadamente, resumir sem representar minorias, interpretar sem evidência suficiente e citar sem verificar. A competência central passa a ser formular tarefas de modo controlado, fornecer dados adequados, limitar o escopo da resposta, auditar resultados e combinar a fluência textual da IA com o rigor das práticas cientométricas.

4.5 Classificação, sumarização, extração de entidades e apoio interpretativo

Classificação, sumarização, extração de entidades e apoio interpretativo são quatro famílias de aplicação especialmente relevantes para a cientometria com IA. Embora frequentemente apareçam combinadas, cada uma responde a uma necessidade distinta. Classificação atribui documentos ou trechos a categorias; sumarização condensa informação; extração de entidades identifica elementos relevantes; apoio interpretativo ajuda a construir sentido sobre padrões observados. Confundir essas tarefas pode comprometer o estudo. Um resumo não é uma classificação; uma entidade extraída não é necessariamente um conceito central; uma interpretação não é evidência bruta.

A classificação pode ser supervisionada, não supervisionada ou orientada por instrução. Em classificação supervisionada, utiliza-se um conjunto de exemplos rotulados para treinar ou calibrar um modelo. Em classificação não supervisionada, técnicas de agrupamento identificam proximidades entre documentos sem categorias prévias. Em classificação por *prompt*, um modelo de linguagem recebe instruções e aplica categorias definidas pelo pesquisador. Cada abordagem tem vantagens. A primeira favorece avaliação quantitativa; a segunda revela estruturas inesperadas; a terceira oferece flexibilidade e rapidez. Em estudos cientométricos, uma combinação é muitas vezes mais robusta do que a adoção exclusiva de uma delas.

Uma aplicação típica consiste em classificar resumos de artigos por tema, método ou área de aplicação. O pesquisador pode fornecer uma taxonomia inicial e pedir que a IA atribua cada documento a uma categoria, justificando a decisão com base em evidências textuais. Para melhorar a confiabilidade, o *prompt* deve solicitar grau de confiança e marcação de casos ambíguos. A saída deve ser avaliada por amostragem, comparação com codificação humana e análise dos erros. Categorias com baixa concordância devem ser revistas, fundidas ou subdivididas.

A sumarização pode operar em diferentes escalas. No nível micro, resume um artigo individual. No nível meso, sintetiza um *cluster*, uma instituição, uma área temática ou um período. No nível macro, descreve a evolução de um campo. A cientometria se beneficia da sumarização quando precisa transformar grandes volumes de registros em narrativas compreensíveis. No entanto, sumarizações agregadas devem preservar diversidade. Um *cluster* pode conter sublinhas, tensões e exceções. O *prompt* deve pedir explicitamente que o modelo identifique temas majoritários, temas minoritários, documentos representativos e divergências internas.

A extração de entidades permite enriquecer metadados bibliográficos. Entidades podem incluir métodos, bases de dados, tecnologias, organismos, populações, países, políticas públicas, instrumentos, teorias, autores citados, instituições ou objetos empíricos. Em áreas biomédicas, por exemplo, podem ser doenças, genes, fármacos ou intervenções. Em ciência da informação, podem ser sistemas, serviços, métricas ou práticas de organização do conhecimento. A extração de entidades permite transformar texto não estruturado em variáveis analisáveis, conectando bibliometria e análise de conteúdo.

A extração de entidades exige cuidado com granularidade. Se as entidades forem amplas demais, perdem valor analítico; se forem específicas demais, tornam-se dispersas e difíceis de interpretar. A IA pode sugerir níveis de agregação, mas o pesquisador deve decidir qual granularidade responde à pergunta. Em uma análise de IA na educação, por exemplo, 'aprendizagem personalizada' pode ser categoria temática, enquanto 'sistema tutor inteligente' pode ser tecnologia, e 'ensino superior' pode ser contexto institucional. Misturar esses níveis em uma única variável produz confusão analítica.

O apoio interpretativo é uma aplicação mais aberta e, por isso, mais sensível. Um modelo pode receber uma tabela de indicadores e produzir uma leitura. Pode receber uma lista de *clusters* e sugerir rótulos. Pode receber uma série temporal e apontar hipóteses para mudanças. Pode comparar dois períodos e indicar transições temáticas. Em todos os casos, a resposta deve ser estruturada em camadas: descrição observável, interpretação plausível, hipóteses a verificar e limitações. Essa estrutura reduz o risco de que o modelo apresente especulações como fatos.

Uma prática recomendável é solicitar que a IA produza interpretações concorrentes. Em vez de aceitar a primeira narrativa, o pesquisador pode pedir três explicações alternativas para um padrão: uma explicação metodológica, uma explicação sociológica e uma explicação relacionada à cobertura da base. Por exemplo, um aumento de publicações pode refletir crescimento real do campo, mudança terminológica, indexação de novos periódicos ou alteração na *string* de busca. A IA é útil quando ajuda a multiplicar hipóteses, não quando fecha prematuramente a interpretação.

Na rotulagem de *clusters*, a IA pode examinar títulos, resumos, palavras-chave e documentos mais centrais para sugerir nomes. A boa rotulagem deve ser curta, específica e teoricamente informativa. Rótulos genéricos, como 'tecnologia', 'gestão' ou 'educação', raramente bastam. O *prompt* deve exigir que cada rótulo venha acompanhado de justificativa, termos característicos, documentos representativos e incertezas. Depois, especialistas devem avaliar se o rótulo representa adequadamente o agrupamento e se não obscurece subtemas relevantes.

A IA também pode apoiar a detecção de lacunas. A partir de uma matriz temática, pode identificar combinações pouco exploradas entre métodos, contextos e objetos. Contudo, lacuna não é simplesmente baixa frequência. Uma baixa frequência pode resultar de irrelevância teórica, impossibilidade metodológica, novidade recente, viés de cobertura ou uso de vocabulário alternativo. Por isso, *prompts* de lacuna devem solicitar justificativa baseada em evidência e indicar se a lacuna é empírica, teórica, metodológica, geográfica, temporal ou de aplicação.

Outra aplicação é a comparação entre áreas ou períodos. A IA pode ajudar a descrever diferenças entre a literatura de 2000-2010 e 2011-2025, ou entre produção de países distintos. Pode identificar mudança de vocabulário, crescimento de métodos computacionais, deslocamento de objetos empíricos ou internacionalização de redes. No entanto, comparações devem controlar diferenças de volume, cobertura, idioma e práticas de citação. O modelo deve ser instruído a não atribuir diferenças apenas a preferências científicas quando fatores de base de dados podem estar envolvidos.

Apoios de IA também são úteis para transformar resultados em recomendações operacionais. Para bibliotecas, pode indicar coleções prioritárias; para programas de pós-graduação, pode sugerir frentes de colaboração; para agências de fomento, pode apontar temas emergentes que requerem acompanhamento; para grupos de pesquisa, pode revelar oportunidades de internacionalização. Entretanto, recomendações devem ser apresentadas

como possibilidades condicionadas. Uma recomendação responsável inclui justificativa, evidência, limitações e possíveis efeitos adversos.

O uso de classificação, sumarização e extração em larga escala exige governança da qualidade. Recomenda-se registrar modelo utilizado, versão, data, parâmetros, *prompt*, amostra processada, critérios de avaliação e taxa de revisão humana. Quando a saída é incorporada a uma base de dados, deve haver campos separados para valor original, valor inferido por IA, grau de confiança e status de validação. Essa separação evita que dados gerados sejam confundidos com metadados originais.

Em termos práticos, a melhor arquitetura é iterativa. Primeiro, a IA propõe categorias, entidades ou resumos. Depois, o pesquisador revisa uma amostra. Em seguida, o *prompt* é ajustado, os critérios são refinados e nova rodada é executada. Por fim, os resultados são avaliados e documentados. Esse ciclo aproxima o uso de *prompts* de procedimentos de codificação qualitativa e análise de dados, nos quais categorias não são apenas aplicadas, mas progressivamente estabilizadas.

Portanto, a IA fortalece a cientometria quando transforma textos dispersos em informação estruturada e quando ajuda a formular narrativas interpretativas sobre evidências calculadas. Seu papel é menos o de substituir o analista e mais o de ampliar sua capacidade de observar padrões, formular hipóteses, testar categorias e comunicar resultados. O ganho metodológico depende da manutenção de uma fronteira clara entre extração, classificação, síntese e interpretação.

4.6 Exemplos aplicados de uso controlado da IA

Caso 1: análise de uma área emergente

Considere um estudo sobre IA generativa na educação superior. A IA pode ajudar a expandir vocabulário de busca, distinguindo termos como "*large language models*", *'generative artificial intelligence'*, "ChatGPT", *"automated feedback"*, "AI tutors" e *"academic integrity"*. Após a coleta, pode classificar os resumos por foco temático: ensino-aprendizagem, avaliação, ética, formação docente, desenho instrucional e política institucional. O pesquisador deve, no entanto, verificar se a estratégia não favorece apenas artigos que mencionam nomes comerciais e se não exclui trabalhos que discutem tecnologias equivalentes com outra terminologia.

Nesse caso, a IA pode resumir *clusters* e sugerir rótulos, mas a análise temporal deve ser cautelosa. A explosão de publicações após 2022 pode refletir tanto crescimento real do tema quanto mudança de vocabulário e visibilidade pública dos modelos generativos. O relatório deve distinguir essas hipóteses. O *prompt* de auditoria pode perguntar: 'quais explicações alternativas existem para o crescimento observado?'. Essa pergunta impede que a narrativa reduza a série temporal a uma única causa.

Caso 2: diagnóstico institucional

Em um diagnóstico de produção científica de uma universidade, a IA pode ajudar a organizar resultados por áreas, identificar redes de colaboração e redigir sínteses para gestores. Entretanto, não deve inferir qualidade de grupos apenas por volume de publicação ou citação. Áreas com menor densidade citacional, produção em livros ou relevância regional podem ser subestimadas. O uso responsável exige indicadores contextualizados, leitura qualitativa e reconhecimento da missão institucional.

A IA pode ser útil para gerar perguntas que o relatório deve responder: quais áreas crescem? Quais colaboram internacionalmente? Quais têm baixa integração interna? Quais

temas emergem? Quais indicadores são inadequados para certas áreas? Ao transformar resultados em recomendações, o modelo deve ser instruído a apresentar alternativas e riscos. Recomendar fusão de grupos, priorização de temas ou mudanças estratégicas sem validação humana seria metodologicamente inadequado.

Caso 3: mapeamento de colaboração internacional

Em uma análise de colaboração internacional, a IA pode apoiar a padronização de países e instituições, explicar medidas de rede e sintetizar padrões de centralidade. Contudo, coautoria internacional não equivale necessariamente a colaboração simétrica. Um artigo com autores de vários países pode refletir consórcio, coleta multicêntrica, colaboração desigual ou participação periférica. A interpretação exige cautela e, em alguns casos, análise qualitativa dos papéis autorais e financiamento.

Prompts podem ajudar a identificar riscos de interpretação: 'liste três leituras possíveis para a centralidade elevada de determinado país, indicando evidências adicionais necessárias'. Essa abordagem estimula interpretação plural e evita conclusão automática de liderança científica. A IA torna-se útil quando amplia a sensibilidade sociológica da análise, não quando transforma redes em rankings simplificados.

4.7 Riscos: alucinação, viés, falsa autoridade, erros de citação e perda de rastreabilidade

Os riscos da IA generativa em estudos científicos não são periféricos; eles afetam diretamente a validade dos resultados. O primeiro e mais discutido é a alucinação. Em termos gerais, alucinação ocorre quando o modelo produz conteúdo plausível, fluente e aparentemente informativo, mas falso, não sustentado pelos dados ou incompatível com a evidência disponível. Em cientometria, isso pode aparecer como referência inexistente, DOI inválido, atribuição errada de autoria, interpretação não sustentada de um *cluster*, descrição falsa de cobertura de base ou invenção de uma tendência não observada.

A alucinação é particularmente perigosa porque se apresenta com forma acadêmica. Um modelo pode produzir uma explicação com vocabulário técnico, tom equilibrado e estrutura convincente, mesmo quando a informação é incorreta. A aparência de rigor não garante fundamento. Por isso, em estudos cientométricos, qualquer saída generativa deve ser verificada contra dados, tabelas, mapas e fontes originais. *Prompts* que solicitam incerteza ajudam, mas não eliminam o problema. A verificação externa permanece indispensável.

O viés constitui outro risco central. Modelos de linguagem refletem padrões de seus dados de treinamento, e bases bibliográficas refletem assimetrias de indexação. Quando esses dois níveis se combinam, o risco de amplificação é significativo. A IA pode favorecer literatura em inglês, áreas com maior visibilidade internacional, periódicos de alto impacto, instituições centrais e autores amplamente citados. Pode também marginalizar produção regional, literatura em outros idiomas, práticas científicas locais e periódicos não indexados em bases dominantes. A cientometria com IA deve, portanto, incorporar análise de cobertura e sensibilidade cultural.

A falsa autoridade ocorre quando a resposta da IA é tratada como parecer especializado. Modelos generativos não possuem responsabilidade epistêmica, não têm compromisso institucional com a verdade e não compreendem consequências de suas recomendações. Ainda assim, sua fluência pode induzir usuários a aceitar conclusões sem verificação. Em ambientes de gestão científica, essa falsa autoridade pode afetar avaliações, rankings, prioridades de financiamento e decisões sobre colaboração. A forma

de mitigação é exigir que toda afirmação venha acompanhada de evidência, limitação e status: dado, inferência ou hipótese.

Erros de citação merecem atenção específica. Pesquisas sobre referências geradas por ChatGPT documentaram fabricação e imprecisão bibliográfica. Em cientometria, a citação é simultaneamente objeto de estudo e elemento de sustentação argumentativa. Um erro de citação pode contaminar a bibliografia, fragilizar o argumento e comprometer a confiança no estudo. Portanto, modelos generativos não devem ser usados como fontes finais de referências. Eles podem sugerir autores ou tópicos a serem verificados, mas a referência final deve ser confirmada em fonte bibliográfica confiável.

Quadro 4.4 - Riscos de uso da IA em estudos cientométricos

1. ALUCINAÇÃO *Informação falsa ou não sustentada*	**2. VIÉS** *Amplificação de assimetrias*	**3. FALSA AUTORIDADE** *Aceitação acrítica da resposta*	**4. ERROS DE CITAÇÃO** *Referências incorretas ou fabricadas*	**5. PERDA DE RASTREABILIDADE** *Opacidade e falta de auditoria*
O modelo produz conteúdo plausível, fluente e aparentemente informativo, mas falso ou incompatível com os dados e as evidências.	Modelos e bases refletem assimetrias de linguagem, área, país, visibilidade e indexação, que podem ser amplificadas pela IA.	A fluência do modelo pode induzir a aceitar conclusões como se fossem parecer especializado, embora não haja responsabilidade epistêmica.	Modelos podem inventar ou imprecisar referências. Como a citação é objeto de estudo e base de argumentação, erros comprometem todo o trabalho.	Sem registro do processo (*prompts*, modelos, entradas e decisões), não é possível reconstruir o caminho analítico.
Exemplos em cientometria				
• Referências inexistentes ou DOIs inválidos • Autoria, título ou periódico incorretos • Atribuição errada de afiliações • Interpretação falsa de *clusters* e mapas • Cobertura da base descrita incorretamente • Tendências ou lacunas inventadas	• Favoritismo à literatura em inglês • Sobrerrepresentação de áreas centrais • Valorização de periódicos de alto impacto • Invisibilidade da produção regional e local • Reforço de hierarquias existentes	• Recomendações apresentadas como fato • Juízos de valor sem evidência explícita • Rankings ou avaliações simplistas • Generalizações sobre "a literatura" • Proposições normativas não justificadas	• Referências inexistentes ou imprecisas • Erros de ano, volume, páginas ou autoria • DOIs ou links inválidos • Atribuição errada de ideias ou dados • Citações fora de contexto	• *Prompts*, versões e datas não registrados • Dados de entrada não documentados • Revisões humanas não descritas • Critérios de decisão não explicitados • Saídas intermediárias perdidas
Principais consequências				
• Contamina a bibliografia e o argumento • Gera interpretações erradas • Compromete a credibilidade do estudo • Pode induzir decisões inadequadas	• Distorce comparações entre áreas e países • Marginaliza comunidades e saberes locais • Reforça desigualdades e exclusões • Limita a diversidade de perspectivas	• Decisões avaliativas mal fundamentadas • Uso da IA como substituto da análise crítica • Interpretações frágeis • Risco para políticas e financiamento	• Bibliografia incorreta e não verificável • Fragilização do argumento científico • Perda de confiança de leitores e revisores • Risco ético e de má conduta	• Impossibilidade de auditoria e de reprodução • Dificuldade de revisão por pares • Intransparência para usuários e gestores • Responsabilidade difusa

Fonte: Elaborado pelo autor (2026).

A perda de rastreabilidade é um risco menos visível, porém estrutural. Quando o pesquisador usa IA em várias etapas sem registrar *prompts*, versões, dados de entrada e decisões de revisão, torna-se impossível reconstruir o caminho analítico. Estudos cientométricos devem permitir auditoria. Isso implica registrar, no mínimo, o *prompt* utilizado, o modelo, a data, o *corpus* fornecido, a saída gerada, as alterações humanas e o destino da informação. A rastreabilidade é ainda mais importante quando a IA contribui para classificação, rotulagem ou interpretação, pois essas etapas influenciam diretamente os resultados.

Há também risco de circularidade. Se a IA é treinada sobre textos que já incorporam determinadas hierarquias, e depois é usada para interpretar indicadores produzidos por bases também hierarquizadas, o estudo pode reforçar estruturas dominantes. Por exemplo, pode confirmar a centralidade de países, periódicos e autores já favorecidos pela indexação internacional. A cientometria crítica deve usar IA para revelar e questionar assimetrias, não para naturalizá-las. *Prompts* de auditoria podem solicitar explicitamente análise de exclusões, invisibilidades e efeitos de cobertura.

Outro problema é a instabilidade das respostas. O mesmo *prompt* pode gerar saídas diferentes em momentos distintos, sobretudo em modelos probabilísticos ou em plataformas que atualizam seus sistemas sem aviso ao usuário. Essa instabilidade afeta reprodutibilidade. Para mitigá-la, recomenda-se salvar respostas, usar parâmetros controlados quando disponíveis, aplicar *prompts* padronizados, reduzir ambiguidade e, quando possível, empregar modelos locais ou versões fixas. Em relatórios, deve-se informar o grau em que a IA foi usada e como as respostas foram validadas.

A IA pode também induzir excesso de generalização. A partir de uma amostra delimitada, pode produzir afirmações sobre 'a literatura' como se o *corpus* representasse todo o campo. Em cientometria, o *corpus* é sempre uma aproximação. Ele depende da base, dos filtros, da língua, do tipo documental e da estratégia de busca. Um estudo sobre artigos indexados na Scopus não equivale a estudo sobre toda a produção científica mundial. O *prompt* deve obrigar o modelo a mencionar o escopo em cada interpretação relevante.

Há risco adicional na interpretação de ausência. Se um tema não aparece na amostra, isso não significa necessariamente que não exista produção sobre ele. Pode haver problemas de vocabulário, indexação, idioma, tipo documental ou cobertura regional. Modelos generativos, ao responderem de modo assertivo, podem transformar ausência de evidência em evidência de ausência. O pesquisador deve exigir que o modelo formule alternativas metodológicas antes de concluir que uma lacuna é real.

A confidencialidade é outro ponto sensível. Estudos cientométricos podem envolver dados públicos, mas também podem incluir relatórios internos, estratégias institucionais, avaliações de grupos, pareceres ou documentos não publicados. Enviar esses dados a ferramentas externas pode violar políticas institucionais, direitos autorais ou acordos de confidencialidade. Antes de usar IA, é necessário avaliar o regime de dados, a política da ferramenta, a possibilidade de anonimização e a adequação jurídica do processamento.

A dependência excessiva da IA pode enfraquecer competências analíticas. Se o pesquisador passa a aceitar classificações, resumos e interpretações sem compreender os indicadores e métodos subjacentes, a cientometria torna-se uma prática de superfície. A formação deve combinar alfabetização em IA com alfabetização em métricas, redes, bases de dados, normalização e epistemologia da avaliação científica. A IA deve ampliar a competência, não substituí-la.

Um risco epistemológico mais profundo é a transformação da interpretação em produção textual automática. A interpretação científica envolve julgamento, contextualização

histórica, comparação teórica, reconhecimento de controvérsias e sensibilidade às limitações do método. Modelos generativos podem produzir textos interpretativos, mas não assumem responsabilidade por eles. A autoria intelectual do relatório permanece humana. O pesquisador deve revisar não apenas a correção factual, mas a adequação argumentativa, ética e política da narrativa.

A mitigação desses riscos exige desenho de *prompts* com salvaguardas. Um *prompt* responsável solicita que o modelo indique incerteza, separe dado e inferência, não invente informações, peça dados ausentes, cite apenas fontes fornecidas quando essa for a regra, identifique possíveis vieses, ofereça explicações alternativas e proponha formas de verificação. Além disso, a saída deve ser submetida a auditoria. O uso de *prompts* de auditoria, apresentados ao final do capítulo, é parte central dessa estratégia.

Em conclusão, riscos como alucinação, viés, falsa autoridade, erro de citação e perda de rastreabilidade não impedem o uso de IA em cientometria, mas definem suas condições de legitimidade. Quanto mais influente for a decisão derivada do estudo, mais rigorosa deve ser a validação. A IA é aceitável como instrumento de apoio quando suas respostas são tratadas como objetos a verificar. Torna-se problemática quando suas respostas são tratadas como evidências finais.

Quadro 4.5 - Matriz de riscos e salvaguardas

Risco	Manifestação em cientometria	Efeito provável	Salvaguarda
Alucinação	Invenção de referências, tendências ou explicações não sustentadas.	Fragilidade científica e perda de confiança.	Verificação externa, uso de fontes fornecidas e *prompt* de auditoria.
Viés	Reprodução de assimetrias de idioma, país, área e indexação.	Invisibilidade de comunidades e produção local.	Análise de cobertura e comparação entre bases.
Falsa autoridade	Resposta fluente tratada como parecer especializado.	Decisões sem base metodológica.	Separar dado, inferência e hipótese; revisão por especialistas.
Erro de citação	DOI inválido, autor incorreto, periódico inexistente.	Bibliografia contaminada e argumento frágil.	Checagem em DOI, periódicos, bases e anais oficiais.
Perda de rastreabilidade	*Prompts*, versões e decisões não registrados.	Impossibilidade de auditoria.	Apêndice metodológico de uso da IA.

Fonte: Elaborado pelo autor (2026).

Validação e calibração de saídas de IA em cientometria

A validação das saídas de IA deve ser planejada de acordo com a natureza da tarefa. Em tarefas de revisão linguística, a verificação pode se concentrar em clareza, estilo e preservação do sentido. Em tarefas de classificação, extração de entidades ou rotulagem temática, a validação deve incluir comparação com julgamento humano, análise de erros, amostragem estratificada e revisão dos casos de baixa confiança. Em tarefas interpretativas, a validação deve confrontar a narrativa produzida com dados originais, indicadores calculados e literatura de referência. A diversidade de tarefas exige diversidade de critérios de controle.

A calibração começa com exemplos. Antes de aplicar um *prompt* a todo o *corpus*, o pesquisador deve selecionar uma amostra pequena, representativa e deliberadamente variada. Essa amostra deve incluir casos típicos, casos fronteiriços, documentos de diferentes períodos, textos de áreas distintas e registros com metadados incompletos. O objetivo é observar como o modelo se comporta diante da heterogeneidade real do material. Se o modelo classifica bem apenas os casos fáceis, mas falha nos casos ambíguos, o *prompt* precisa ser reformulado ou a tarefa deve permanecer sob codificação humana mais intensa.

A validação deve considerar precisão e utilidade. Uma saída pode estar formalmente correta, mas ser pouco útil para a pergunta de pesquisa. Por exemplo, um modelo pode rotular um *cluster* como "inteligência artificial" quando todos os documentos do *corpus* já tratam desse tema. O rótulo é correto, mas analiticamente vazio. A calibração deve, portanto, avaliar especificidade, poder discriminativo e valor interpretativo. Em cientometria, a boa categoria não é apenas verdadeira; ela ajuda a diferenciar padrões relevantes no conjunto analisado.

É recomendável usar esquemas de dupla revisão. Uma primeira rodada de IA pode produzir classificação ou resumo. Uma segunda rodada, com *prompt* de auditoria, pode avaliar a consistência da saída. Finalmente, um pesquisador humano decide. Essa triangulação não elimina erros, mas aumenta a probabilidade de identificá-los. Em estudos de maior consequência, pode-se combinar revisão por especialistas de domínio, bibliotecários, analistas de dados e gestores. A validação se torna, nesse caso, processo coletivo de qualificação da evidência.

A calibração também deve medir estabilidade. O mesmo *prompt* pode ser executado em pequena amostra mais de uma vez, quando a ferramenta permitir, para verificar variações de classificação e interpretação. Se uma categoria muda frequentemente sem alteração dos dados, isso indica instabilidade. O pesquisador pode reduzir a variabilidade especificando critérios mais objetivos, solicitando formato tabular, limitando opções de resposta e exigindo justificativas textuais baseadas em trechos. Saídas abertas tendem a variar mais do que saídas estruturadas.

Outro critério é a auditabilidade da justificativa. Ao classificar um resumo, o modelo deve indicar quais palavras, expressões ou frases sustentam a decisão. Essa evidência textual permite ao pesquisador verificar se a categoria foi atribuída com base em elementos pertinentes ou em associação superficial. A justificativa também facilita a identificação de vieses: o modelo pode estar usando um termo muito geral, ignorando contexto ou atribuindo categoria com base em expectativa externa. Sem justificativa, a classificação permanece opaca.

A validação de sumarizações exige comparar síntese e *corpus*. Uma boa síntese deve preservar temas centrais, reconhecer temas minoritários e não introduzir informação ausente. Deve também diferenciar resultados frequentes de achados excepcionais. Em *corpora* grandes, pode-se revisar amostras de documentos de cada *cluster* e verificar se o resumo os representa. Quando o modelo produz uma síntese aparentemente elegante, mas sem vínculos explícitos com documentos, a confiabilidade é menor. A síntese deve ser tratada como mapa textual, não como substituto da leitura.

A validação de interpretações requer examinar a coerência entre argumento e evidência. Se o modelo afirma que um tema está emergindo, deve haver suporte temporal. Se afirma que um autor é ponte entre comunidades, deve haver medida de centralidade ou evidência de coautoria/cocitação. Se afirma que uma área é interdisciplinar, deve haver indicadores de diversidade de categorias, referências ou periódicos. A revisão humana deve perguntar constantemente: qual dado sustenta esta frase? Quando a resposta não é clara, a frase deve ser reescrita como hipótese ou removida.

Quadro 4.6 - Checklist de verificação humana

CHECKLIST DE VERIFICAÇÃO HUMANA *(antes de incorporar a saída da IA ao estudo)*
☐ As afirmações estão ancoradas em dados e evidências verificáveis?
☐ As referências foram confirmadas em fontes confiáveis?
☐ Há distinção clara entre dado, inferência, hipótese e opinião?
☐ Os vieses de cobertura, idioma, área e base foram considerados?
☐ As generalizações respeitam o escopo e o nível de análise?
☐ A saída foi validada por amostragem ou por especialistas?
☐ As decisões humanas e os ajustes estão documentados?
☐ A conclusão considera limitações e alternativas plausíveis?
☐ O uso desta saída é apropriado ao impacto da decisão a ser tomada?

Registro da verificação

Responsável pela verificação	Data	Saída / etapa verificada

Fonte: Elaborado pelo autor (2026).

Rastreabilidade, documentação e reprodutibilidade do uso de prompts

A rastreabilidade é o princípio que permite reconstruir o caminho entre pergunta, dado, *prompt*, resposta e decisão analítica. Em estudos tradicionais, ela é buscada pela documentação de bases, *strings* de busca, filtros, softwares, versões e parâmetros. Na cientometria com IA, acrescentam-se novos elementos: modelo utilizado, data de interação, *prompt* exato, dados fornecidos, saída recebida, revisão humana e uso final da resposta. Sem esse registro, a IA torna-se uma etapa invisível do método, enfraquecendo a avaliação por pares e a possibilidade de replicação conceitual.

A documentação dos *prompts* deve ser suficientemente completa para que outro pesquisador compreenda a tarefa. Não basta declarar que 'foi usado ChatGPT para apoiar a análise'. É necessário indicar se a IA sugeriu termos de busca, classificou resumos, resumiu *clusters*, revisou interpretações ou redigiu trechos. Cada uma dessas funções possui impacto metodológico diferente. Um uso linguístico tem consequências distintas de um uso classificatório. A transparência deve informar o grau de influência da IA no resultado final.

A reprodutibilidade exata pode ser difícil quando modelos são atualizados e respostas são probabilísticas. Por isso, a meta prática deve ser a reprodutibilidade procedimental: permitir que outro pesquisador compreenda e avalie o procedimento, mesmo que a saída textual varie. Para isso, é útil salvar exemplos de saída, critérios de aceitação e decisões tomadas após revisão. A rastreabilidade, nesse sentido, aproxima-se de um diário metodológico, no qual se registra não apenas o que o modelo respondeu, mas o que a equipe fez com a resposta.

Um apêndice de *prompts* pode ser organizado por etapa do estudo. Na formulação do problema, registrar *prompts* exploratórios. Na busca, registrar *prompts* de expansão terminológica e revisão de *strings*. No pré-processamento, registrar *prompts* de normalização. Na análise, registrar *prompts* de interpretação e rotulagem. Na redação, registrar *prompts* de síntese e adaptação ao público. Na auditoria, registrar *prompts* de verificação de afirmações. Essa organização evidencia que a IA não é evento isolado, mas componente de um fluxo metodológico.

A documentação deve separar dados originais e dados derivados. Palavras-chave exportadas de uma base são dados originais do registro bibliográfico. Grupos de sinônimos propostos por IA são dados derivados. Rótulos de *clusters* sugeridos por modelo são interpretações derivadas. Essa separação evita que categorias geradas sejam confundidas com metadados fornecidos pela fonte. Em bancos de dados, recomenda-se criar campos específicos: valor original, valor normalizado, origem da normalização, grau de confiança e status de validação.

A rastreabilidade também protege o pesquisador contra mudanças de entendimento ao longo do projeto. Estudos cientométricos podem durar semanas ou meses. Durante esse período, a equipe pode ajustar termos, excluir registros, refinar categorias e alterar interpretações. Se as interações com IA não forem registradas, torna-se difícil explicar por que determinado *cluster* recebeu um rótulo ou porque certos termos foram agrupados. O registro sistemático reduz dependência de memória e favorece consistência interna.

Em contextos institucionais, a documentação do uso de IA é parte da governança. Relatórios que orientam decisões sobre financiamento, contratação, avaliação ou estratégia de pesquisa devem permitir auditoria posterior. Se uma recomendação foi parcialmente produzida por IA, isso precisa ser conhecido. Não se trata de desvalorizar o estudo, mas de qualificar sua transparência. A confiança em resultados assistidos por IA depende menos da ocultação da ferramenta e mais da demonstração de controle sobre sua contribuição.

A reprodutibilidade também é fortalecida por formatos estruturados de saída. Quando *prompts* pedem tabelas com campos fixos, é mais fácil comparar respostas, importar dados e auditar decisões. Saídas narrativas longas são úteis para interpretação, mas difíceis de validar sistematicamente. Por isso, uma boa prática é solicitar primeiro uma saída estruturada e depois uma narrativa baseada nessa estrutura. A narrativa deve derivar da tabela, não substituí-la.

Critérios para decidir quando usar, limitar ou evitar IA

A decisão de usar IA deve considerar adequação da tarefa, risco da saída e disponibilidade de validação. Tarefas de baixo risco, como transformar uma explicação técnica em linguagem mais acessível, podem se beneficiar amplamente dos modelos generativos. Tarefas de risco intermediário, como classificar resumos ou agrupar termos, exigem validação por amostragem. Tarefas de alto risco, como avaliar desempenho de pesquisadores, atribuir qualidade a instituições ou recomendar políticas de financiamento, não devem depender de IA sem revisão especializada rigorosa.

Um primeiro critério é a disponibilidade de dados explícitos. A IA é mais útil quando recebe dados delimitados: uma tabela, uma lista de resumos, um conjunto de *clusters*, uma *string* de busca ou um protocolo. Quando o modelo é solicitado a responder com base em conhecimento geral, o risco de erro aumenta. Em cientometria, deve-se privilegiar *prompts* ancorados em evidências fornecidas pelo pesquisador. Quanto mais fechado for o universo de referência, mais controlável tende a ser a resposta.

Um segundo critério é a verificabilidade. Se uma saída pode ser verificada por meio de dados, fontes ou julgamento especializado, o uso da IA é mais aceitável. Se a saída envolve afirmações de difícil verificação, recomendações normativas ou previsões amplas, o cuidado deve aumentar. A pergunta operacional é: como saberemos se esta resposta está correta ou adequada? Se não houver resposta clara, a tarefa não deve ser delegada ou deve ser reformulada em partes verificáveis.

Um terceiro critério é a consequência da decisão. Erros em um exercício didático têm impacto limitado; erros em relatório de avaliação institucional podem afetar reputações e recursos. A intensidade da validação deve acompanhar a consequência social e institucional da saída. Essa lógica é coerente com princípios de avaliação responsável: quanto maior a possibilidade de dano, mais transparente, plural e controlado deve ser o uso de métricas e de IA.

Um quarto critério é a exigência de domínio especializado. A IA pode reconhecer padrões linguísticos, mas nem sempre compreende tradições teóricas, debates disciplinares e contextos históricos. Em áreas com forte carga conceitual, como estudos sociais da ciência, saúde coletiva, educação ou humanidades digitais, a interpretação exige familiaridade com disputas e categorias internas. A IA pode apoiar, mas especialistas devem revisar. Quando não houver especialista disponível, a interpretação deve permanecer mais descritiva e menos normativa.

Também é preciso considerar custo de validação. Se o esforço para verificar a saída da IA é maior do que realizar a tarefa manualmente, o uso pode não se justificar. Isso ocorre, por exemplo, quando o *corpus* é pequeno ou quando a tarefa exige leitura profunda de poucos documentos. A IA é mais vantajosa em tarefas repetitivas, volumosas ou exploratórias, desde que o processo de validação seja proporcional. A busca por eficiência não deve produzir falsa economia metodológica.

Há situações em que evitar IA é a decisão mais responsável. Isso inclui dados confidenciais sem ambiente seguro, tarefas cuja saída não pode ser validada, demandas que exigem juízo ético sensível, geração de referências bibliográficas não verificadas e interpretação causal sem desenho apropriado. A maturidade no uso de IA não se mede pela frequência de uso, mas pela capacidade de escolher quando a ferramenta agrega valor e quando introduz fragilidade.

Por fim, a decisão deve ser comunicada no relatório. Quando a IA foi usada, indicar como. Quando não foi usada em determinada etapa, explicar por que, se isso for metodologicamente relevante. Essa transparência ajuda o leitor a compreender o grau de automação e a robustez do estudo. Em um campo que lida com avaliação e evidência, a própria prática de uso da IA deve ser objeto de prestação de contas.

Boas práticas para integrar IA a relatórios, artigos e diagnósticos cientométricos

A integração da IA ao produto final do estudo deve preservar a distinção entre método, resultado e interpretação. Na seção metodológica, descrevem-se dados, bases, critérios, softwares, indicadores e uso da IA. Na seção de resultados, apresentam-se achados calculados ou classificados. Na discussão, formulam-se interpretações e hipóteses. A IA pode apoiar todas essas seções, mas não deve misturá-las. Um texto gerado que apresenta interpretação como resultado precisa ser revisado e reorganizado.

Relatórios cientométricos devem evitar linguagem excessivamente assertiva quando os dados não autorizam. Expressões como “prova”, “demonstra”, “determina” ou “causa” raramente são adequadas em análises baseadas em indicadores e redes. A IA tende a produzir frases fluentes e conclusivas; por isso, *prompts* devem solicitar linguagem proporcional à evidência.

Formulações como “sugere”, “indica”, “é compatível com”, “pode refletir” e ‘requer verificação adicional’ são mais adequadas quando se lida com inferências.

Uma boa prática é incluir quadro de limitações produzido com apoio de IA, mas revisado pela equipe. Esse quadro pode abordar cobertura de base, recorte temporal, idioma, tipo documental, qualidade dos metadados, limitações dos indicadores, decisões de normalização, parâmetros de visualização e uso de modelos generativos. Ao tornar limitações visíveis, o estudo aumenta sua credibilidade. A IA pode ajudar a lembrar riscos, mas a equipe deve selecionar os que realmente se aplicam.

Em artigos científicos, o uso de IA deve respeitar diretrizes editoriais. Muitos periódicos exigem declaração de uso de ferramentas generativas e não aceitam IA como autora. Mesmo quando não houver exigência formal, é recomendável declarar usos substanciais, especialmente em classificação, sumarização e redação. A autoria implica responsabilidade intelectual, capacidade de responder por dados e compromisso com integridade científica. Modelos de IA não satisfazem esses requisitos.

Em diagnósticos institucionais, a comunicação deve equilibrar clareza e prudência. Gestores precisam de mensagens sintéticas, mas simplificações podem distorcer. A IA pode produzir sumários executivos em linguagem acessível, desde que cada recomendação contenha base empírica, limitações e alternativas. A inclusão de uma seção ‘o que os dados não permitem concluir’ é especialmente útil. Ela previne usos indevidos do relatório e protege a equipe analítica de interpretações excessivas.

Ao apresentar *prompts* no relatório, não é necessário incluir todas as interações completas no corpo do texto. Pode-se descrever a estratégia e disponibilizar *prompts* principais em apêndice. O corpo do relatório deve informar finalidade e validação. O apêndice oferece detalhes para auditoria. Essa solução preserva legibilidade e transparência. Em trabalhos extensos, uma tabela de *prompts* por etapa pode ser mais útil do que longas transcrições.

A integração responsável também requer revisão final de consistência terminológica. Modelos de linguagem podem alternar entre termos próximos, como bibliometria, cientometria, informetria e análise de ciência, sem respeitar distinções conceituais. Podem também variar nomes de indicadores e métricas. Antes da entrega, a equipe deve verificar se os conceitos estão usados de forma estável. Em capítulos de livro e relatórios técnicos, a consistência terminológica é parte da qualidade científica.

Finalmente, a IA deve ser usada para fortalecer a argumentação, não para preenchê-la artificialmente. Um relatório bom não é aquele com maior volume de texto, mas aquele em que cada afirmação contribui para responder à pergunta. A tentação de gerar longas interpretações deve ser controlada por critérios de relevância. A IA pode expandir possibilidades, mas a equipe deve condensar, selecionar e hierarquizar. A inteligência metodológica continua residindo na capacidade humana de julgar o que merece permanecer.

Competências do pesquisador na cientometria assistida por IA

A incorporação da IA não reduz a exigência de competência cientométrica; ao contrário, amplia o conjunto de conhecimentos necessários. O pesquisador precisa compreender indicadores, bases de dados, redes, métodos de normalização e fundamentos de avaliação científica. Precisa também compreender limitações de modelos de linguagem, desenho de *prompts*, princípios de validação e riscos de alucinação. A competência passa a ser híbrida: combina domínio substantivo, literacia informacional, estatística aplicada, leitura crítica de visualizações e governança de sistemas algorítmicos.

Uma competência decisiva é a capacidade de formular boas perguntas. Modelos generativos respondem melhor quando a tarefa está delimitada, mas a delimitação depende do pesquisador. Perguntas vagas geram respostas vagas; perguntas orientadas por unidade de análise, período, fonte e finalidade geram respostas mais úteis. Em cientometria, formular uma pergunta significa antecipar quais dados serão necessários, quais indicadores serão pertinentes e quais interpretações serão possíveis. A IA pode auxiliar esse processo, mas não pode substituir a compreensão do problema de pesquisa.

Outra competência é a leitura crítica de métricas. O pesquisador deve saber que produtividade, impacto, colaboração e visibilidade são dimensões distintas. Deve reconhecer que citações têm significados variados, que indicadores precisam ser normalizados em certos contextos e que redes dependem de escolhas de construção. Sem essa base, a IA pode produzir explicações convincentes que o usuário não tem condições de avaliar. A alfabetização métrica é, portanto, condição para usar IA sem subordinar-se a ela.

A competência em curadoria de dados também se torna central. O pesquisador deve saber selecionar fontes, examinar metadados, controlar duplicatas, documentar filtros e preservar versões. A IA pode sugerir procedimentos, mas não conhece automaticamente as especificidades do *corpus*. A curadoria é o momento em que o estudo define sua relação com a realidade empírica. Se essa etapa for negligenciada, qualquer interpretação posterior, humana ou algorítmica, será comprometida.

A competência em validação exige disposição para desconfiar de respostas bem escritas. O pesquisador deve pedir evidências, revisar amostras, comparar saídas, testar *prompts* e identificar casos em que o modelo parece seguro demais. Essa atitude não é tecnofobia; é método científico aplicado a sistemas generativos. A confiança deve ser conquistada por verificação, não concedida pela fluência. Em ambientes acadêmicos, tal postura precisa ser ensinada como parte da formação em pesquisa.

A competência comunicacional também muda. A IA facilita adaptar textos para públicos distintos, mas cabe ao pesquisador preservar nuances. Um relatório para gestores pode ser claro sem ser simplista; uma explicação para estudantes pode ser acessível sem ser imprecisa; uma síntese executiva pode ser objetiva sem ocultar incertezas. A IA é útil para produzir versões, mas o pesquisador decide qual versão respeita o rigor do estudo.

Finalmente, há uma competência ética. O pesquisador deve reconhecer que análises cientométricas podem afetar reputações, recursos, estratégias institucionais e políticas públicas. Quando a IA participa dessas análises, a responsabilidade não se transfere ao modelo. A equipe humana continua responsável por escolhas, validações, interpretações e consequências. A ética da cientometria com IA consiste em usar a ferramenta para aumentar transparência e qualidade, não para acelerar julgamentos frágeis.

Essas competências sugerem que a formação futura em cientometria deverá incluir laboratórios de *prompts*, exercícios de auditoria, análise de alucinações, comparação entre classificações humanas e automáticas, documentação de *workflows* e discussão de casos de avaliação responsável. A IA não deve ser tratada como atalho, mas como objeto pedagógico. Aprender a usá-la criticamente é parte da própria aprendizagem sobre como a ciência é medida, mapeada e interpretada.

Síntese operacional para projetos de pesquisa

Em projetos de pesquisa, a adoção de IA deve ser convertida em decisões operacionais simples. A equipe pode iniciar respondendo a cinco perguntas: qual tarefa será assistida; quais

dados serão fornecidos; qual formato de saída é esperado; como a resposta será validada; e qual decisão humana será tomada depois da saída. Essas perguntas reduzem a ambiguidade e transformam o uso da IA em parte do desenho de pesquisa. Sem elas, o risco é acumular respostas interessantes, mas metodologicamente desconectadas.

Um fluxo mínimo pode ser organizado em quatro movimentos. Primeiro, preparação do dado: limpeza, seleção e contextualização. Segundo, execução do *prompt*: instrução específica, escopo delimitado e formato estruturado. Terceiro, validação: comparação com dados, revisão por amostra e marcação de incerteza. Quarto, incorporação: decisão sobre aceitar, revisar ou descartar a saída. Esse ciclo deve ser repetido sempre que a IA for usada em uma nova etapa do estudo.

A equipe também deve estabelecer critérios de parada. Nem toda interação com IA melhora o resultado; às vezes, novas rodadas apenas produzem versões mais longas ou mais ornamentadas da mesma interpretação. Critérios de parada podem incluir saturação de categorias, estabilidade da classificação, ausência de novos riscos metodológicos ou validação satisfatória em amostra. A disciplina do encerramento evita que a fluência do modelo prolongue indefinidamente a análise sem ganho científico real.

Outro elemento operacional é a matriz de responsabilidades. Deve ficar claro quem formula *prompts*, quem executa, quem revisa saídas, quem valida referências, quem aprova categorias e quem responde pelo texto final. Essa divisão é especialmente importante em equipes multidisciplinares. Bibliotecários podem contribuir com busca e metadados; cientometristas com indicadores e redes; especialistas de domínio com interpretação; analistas de dados com scripts e reprodutibilidade. A IA se insere nesse ecossistema, mas não substitui papéis.

Por fim, a adoção de IA deve ser avaliada retrospectivamente. Ao final do estudo, a equipe pode registrar quais usos foram produtivos, quais falharam, quais *prompts* precisaram de ajuste e quais riscos foram encontrados. Esse aprendizado alimenta projetos futuros e contribui para construir repertório institucional. A cientometria com IA tende a se tornar mais robusta quando suas práticas são acumuladas, comparadas e criticadas, em vez de tratadas como improvisações isoladas.

Nota de transição para o capítulo seguinte

A discussão desenvolvida neste capítulo prepara o terreno para a engenharia de *prompts* aplicada à cientometria. Se a IA deve ser usada com curadoria, validação, rastreabilidade e responsabilidade, então o *prompt* deixa de ser uma pergunta casual e passa a constituir uma unidade metodológica. O próximo passo consiste em estudar sua anatomia: contexto, papel, tarefa, dados, restrições, formato de saída, critérios de incerteza e procedimentos de auditoria. Em outras palavras, a qualidade da cientometria assistida por IA dependerá não apenas dos modelos disponíveis, mas da capacidade dos pesquisadores de transformar problemas científicos em instruções controladas, documentadas e criticamente avaliáveis.

Essa transição é relevante porque muitas fragilidades atribuídas genericamente à IA decorrem, em parte, de tarefas mal formuladas e de ausência de validação. Bons *prompts* não resolvem todos os riscos, mas tornam o processo mais explícito. Eles obrigam o analista a dizer o que espera, com quais dados, dentro de quais limites e sob quais critérios de aceitação. Por isso, a engenharia de *prompts* deve ser tratada como competência científica aplicada à análise da ciência, e não apenas como habilidade operacional de interação com ferramentas digitais.

4.8 Prompts do capítulo

Os *prompts* a seguir foram concebidos para uso em estudos cientométricos assistidos por IA. Eles devem ser adaptados ao *corpus*, ao indicador, ao mapa ou à pergunta de pesquisa. Recomenda-se registrar a versão utilizada, a data de aplicação, o modelo empregado, os dados de entrada e as decisões de revisão humana.

Prompt 4.8.1 Resposta com nível explícito de incerteza

Finalidade: Usar quando o pesquisador deseja evitar conclusões excessivamente assertivas. O *prompt* é adequado para interpretação de tabelas, mapas e séries temporais, pois obriga a explicitar a força da evidência.

> Atue como especialista em Cientometria.
>
> Analise os dados fornecidos a seguir, mas organize a resposta em quatro colunas:
>
> - afirmação; evidência disponível no dado;
>
> - nível de incerteza (baixo, médio ou alto);
>
> - verificação necessária. Não apresente conclusões que não possam ser sustentadas pelos dados.
>
> Quando houver informação insuficiente, escreva explicitamente 'dados insuficientes para concluir'.
>
> Dados: [INSERIR TABELA, MAPA, LISTA DE *Clusters* OU RESUMOS].

Prompt 4.8.2 Distinção entre dado, inferência e hipótese

Finalidade: Usar na revisão de resultados e discussão. O *prompt* ajuda a separar descrição empírica, interpretação plausível e especulação, reduzindo risco de causalidade indevida.

> Atue como revisor metodológico de um estudo cientométrico.
>
> A partir do texto abaixo, classifique cada afirmação em uma das categorias:
>
> - dado observado;
>
> - inferência interpretativa;
>
> - hipótese a verificar;
>
> - extrapolação indevida.
>
> Para cada item, justifique a classificação e sugira uma formulação mais cautelosa quando necessário.
>
> Texto: [INSERIR INTERPRETAÇÃO OU DISCUSSÃO].

Prompt 4.8.3 Auditoria de afirmações

Finalidade: Usar antes da finalização de relatórios, artigos ou pareceres. O *prompt* transforma a IA em instrumento de controle, não em autoridade substantiva.

Atue como auditor de qualidade científica.

Verifique as afirmações abaixo e produza uma tabela com:

- afirmação; tipo de evidência exigida;
- evidência presente no material fornecido; lacuna identificada;
- risco se a afirmação for mantida;
- recomendação de revisão.

Não use conhecimento externo, a menos que eu forneça fontes.

Afirmações e fontes: [INSERIR TEXTO E FONTES].

Prompt 4.8.4 Classificação assistida com revisão humana

Finalidade: Usar para codificação temática de *corpora*. Requer validação por amostragem e revisão dos casos de baixa confiança.

Classifique os resumos fornecidos segundo a taxonomia abaixo.

Para cada documento, informe categoria principal, categoria secundária, evidência textual, grau de confiança e necessidade de revisão humana.

Se nenhuma categoria for adequada, marque como 'fora da taxonomia' e explique.

Taxonomia: [INSERIR CATEGORIAS].

Resumos: [INSERIR RESUMOS].

Prompt 4.8.5 Rotulagem de clusters

Finalidade: Usar após análise de coocorrência, cocitação ou acoplamento. O objetivo é apoiar a nomeação de comunidades sem transformar o rótulo em verdade definitiva.

A partir dos títulos, palavras-chave e documentos centrais de cada *cluster*, proponha até três rótulos alternativos para cada grupo.

Para cada rótulo, apresente justificativa, termos que o sustentam, documentos representativos e limitações.

Evite rótulos genéricos. Dados dos *clusters*: [INSERIR LISTA].

Prompt 4.8.6 Detecção de riscos metodológicos

Finalidade: Usar no planejamento e na revisão do protocolo. Serve para antecipar fragilidades antes da coleta e análise.

Avalie o desenho do estudo cientométrico abaixo.

Identifique riscos relacionados a base de dados, estratégia de busca, período, idioma, tipo documental, normalização, indicadores, visualização, interpretação e uso de IA.

Para cada risco, indique gravidade, consequência provável e forma de mitigação.

Protocolo: [INSERIR PROTOCOLO].

Os *prompts* apresentados não devem ser entendidos como fórmulas universais. Eles são moldes metodológicos adaptáveis. Sua função principal é forçar a explicitação de elementos que costumam permanecer implícitos: escopo, evidência, incerteza, status epistêmico da afirmação e necessidade de validação. Em aplicações reais, cada *prompt* deve ser acompanhado do *corpus* ou do resultado a ser analisado, pois perguntas genéricas tendem a gerar respostas genéricas.

Um aspecto essencial é a instrução negativa. Dizer ao modelo o que não deve fazer é tão importante quanto dizer o que deve fazer. Em cientometria, instruções como 'não invente referências', 'não afirme causalidade', 'não extrapole além do *corpus*', 'não use conhecimento externo' e 'marque dados insuficientes' funcionam como barreiras de segurança. Elas não eliminam erros, mas reduzem a probabilidade de respostas indevidamente assertivas.

Também se recomenda que *prompts* sejam versionados. Uma pequena mudança na redação pode modificar a saída. Em estudos formais, convém registrar as versões dos *prompts* em apêndice metodológico, juntamente com data de execução, modelo utilizado, parâmetros disponíveis e exemplos de saída. Esse registro transforma o uso da IA em procedimento auditável. Sem ele, a IA permanece como uma caixa-preta discursiva no interior do método.

A validação dos *prompts* deve envolver testes com casos conhecidos. Antes de aplicar um *prompt* a milhares de resumos, o pesquisador pode testá-lo em um conjunto pequeno, previamente codificado por especialistas. A comparação entre saída da IA e julgamento humano permite identificar ambiguidades, categorias mal definidas e instruções insuficientes. Esse procedimento aproxima a engenharia de *prompts* de práticas tradicionais de calibração de instrumentos de pesquisa.

Prompts de auditoria são especialmente relevantes no fechamento do estudo. Depois que a IA ajudou a classificar, resumir ou interpretar, outra instância de uso pode revisar a coerência das afirmações. O pesquisador pode pedir que o modelo procure extrapolações, identifique ausência de evidência e proponha formulações mais cautelosas. Ainda assim, a revisão final deve ser humana, pois envolve julgamento sobre finalidade, público, risco e responsabilidade.

4.9 Considerações finais

A inteligência artificial aplicada à cientometria deve ser compreendida como uma infraestrutura de apoio à formulação, organização, análise e comunicação de estudos sobre a ciência. Seu potencial é amplo: apoiar estratégias de busca, limpar metadados, classificar resumos, extrair entidades, resumir *clusters*, interpretar mapas, redigir relatórios e auditar afirmações. Entretanto, seu uso responsável depende de uma mudança metodológica: *prompts* precisam ser tratados como instrumentos de pesquisa documentados e avaliáveis.

A contribuição mais importante da IA não está em substituir a expertise cientométrica, mas em ampliar a capacidade de examinar *corpora* complexos e formular perguntas mais

ricas. Quando bem usada, ela ajuda a revelar ambiguidades, multiplicar hipóteses, tornar interpretações mais transparentes e comunicar resultados a diferentes públicos. Quando mal usada, pode produzir falsa segurança, referências inexistentes, classificações frágeis e narrativas convincentes sem evidência.

O capítulo defendeu que a cientometria com IA deve se organizar em torno de quatro princípios. O primeiro é a curadoria: dados, termos, fontes e categorias devem ser selecionados e normalizados com cuidado. O segundo é a validação: saídas geradas por IA precisam ser confrontadas com dados, fontes e julgamento especializado. O terceiro é a rastreabilidade: *prompts*, versões, parâmetros e decisões humanas devem ser registrados. O quarto é a responsabilidade: interpretações e recomendações continuam pertencendo ao pesquisador, especialmente quando influenciam avaliação científica e política institucional.

A aplicação de IA generativa à cientometria é, portanto, simultaneamente oportunidade e desafio. A oportunidade reside em tornar a análise da ciência mais escalável, interativa e explicável. O desafio reside em impedir que a fluência textual substitua o rigor metodológico. A resposta não é rejeitar a IA, mas incorporá-la sob protocolos de controle. A próxima etapa da trilha consiste em aprofundar a engenharia de *prompts* como competência metodológica, isto é, como forma de transformar perguntas, dados e critérios de validação em interações analíticas robustas.

Apêndice A.4 - Sugestão para registrar uso de IA

Em estudos cientométricos assistidos por IA, recomenda-se incluir apêndice metodológico com campos mínimos: objetivo do uso da IA; etapa do estudo; ferramenta ou modelo; data de uso; versão, quando disponível; *prompt* utilizado; dados fornecidos ao modelo; parâmetros relevantes; saída produzida; alterações humanas; critério de validação; responsável pela revisão; decisão final sobre incorporação da saída ao estudo.

Esse registro não precisa reproduzir integralmente todos os dados quando houver restrições de confidencialidade, mas deve oferecer rastreabilidade suficiente para avaliação por pares ou auditoria institucional. Quando a IA for usada apenas para revisão linguística, o registro pode ser simples. Quando for usada para classificação, extração ou interpretação, o registro deve ser mais detalhado e incluir amostras de validação.

A transparência sobre uso de IA é parte da ética da pesquisa. Ela permite ao leitor diferenciar resultados calculados por métodos bibliométricos, interpretações humanas e contribuições generativas. Também facilita a aprendizagem coletiva, pois *prompts* bem documentados podem ser aperfeiçoados, comparados e adaptados a novos estudos. A cientometria com IA tende a se consolidar não pela ocultação da mediação algorítmica, mas pela sua explicitação responsável.

Capítulo 5

Engenharia de Prompts para Estudos Cientométricos

5.1 Introdução

A engenharia de *prompts* tornou-se uma prática central na relação entre pesquisadores e modelos de linguagem de grande escala. Em sua formulação mais simples, um *prompt* é uma instrução textual dada a um sistema de inteligência artificial para orientar uma resposta. Todavia, quando inserido no contexto da pesquisa científica, esse conceito precisa ser ampliado. Um *prompt* científico não é apenas uma pergunta; é uma peça de mediação metodológica. Ele define o papel atribuído ao modelo, explicita a tarefa, informa o *corpus* ou os dados disponíveis, impõe restrições interpretativas, solicita formatos de saída e estabelece critérios mínimos de qualidade. Em estudos cientométricos, nos quais a interpretação de indicadores, redes, mapas e tendências depende de escolhas metodológicas sensíveis, essa mediação assume especial importância.

A cientometria opera sobre registros de produção científica, citações, coautorias, palavras-chave, periódicos, instituições, países, áreas de conhecimento e relações conceituais. Tradicionalmente, seus resultados são produzidos por meio de bases bibliográficas, ferramentas estatísticas e softwares de mapeamento, como VOSviewer, CiteSpace, bibliometrix, Gephi e ambientes de programação em R ou Python. A incorporação de modelos generativos não elimina essas ferramentas; ao contrário, exige maior rigor na articulação entre dados estruturados, procedimentos computacionais e interpretação textual. A IA pode apoiar a formulação de perguntas, a normalização terminológica, a classificação temática, a elaboração de narrativas analíticas e a produção de relatórios. Entretanto, ela também pode introduzir erros, reforçar vieses, produzir generalizações indevidas e fabricar referências caso seja usada sem controle.

A engenharia de *prompts*, nesse cenário, deve ser compreendida como disciplina aplicada de desenho de instruções. Ela envolve a construção de comandos capazes de orientar modelos para executar tarefas sob condições especificadas. A literatura recente sobre modelos de linguagem mostrou que a formulação das instruções influencia fortemente o desempenho dos sistemas, seja em tarefas de aprendizagem *few-shot*, raciocínio encadeado, consistência de respostas ou uso de padrões de *prompts*. Trabalhos como os de Brown *et al.* (2020), Wei *et al.* (2022), Wang *et al.* (2023), White *et al.* (2023) e revisões recentes sobre *prompt engineering* demonstram que pequenas variações na forma de solicitar uma tarefa podem alterar a qualidade, a completude, a estrutura e a confiabilidade das respostas. Essa constatação tem implicações diretas para a cientometria: se o *prompt* altera a resposta, então o *prompt* passa a fazer parte do método.

A afirmação de que o *prompt* faz parte do método não é retórica. Em uma revisão cientométrica, por exemplo, pedir a um modelo que "resuma os principais temas" de um conjunto de resumos é metodologicamente diferente de solicitar que ele "classifique cada resumo em categorias previamente definidas, indique evidências textuais, declare grau de confiança e marque casos ambíguos para revisão humana". No primeiro caso, há abertura elevada para inferência livre; no segundo, há estrutura, rastreabilidade e possibilidade de auditoria. Isso vale para a interpretação de redes de coautoria, mapas de cocitação ou tabelas de indicadores. Um *prompt* mal formulado pode converter uma associação em causalidade, transformar uma tendência observada em prognóstico, ou sugerir conclusões não sustentadas pelos dados.

Este capítulo propõe uma abordagem sistemática para a engenharia de *prompts* em estudos cientométricos. O objetivo não é oferecer uma lista de comandos prontos, mas construir uma gramática metodológica. Essa gramática permite que o pesquisador desenhe *prompts* adequados ao tipo de tarefa, ao estágio do estudo, ao nível de incerteza aceitável, ao público destinatário e ao produto final desejado. Por essa razão, os exemplos apresentados ao longo do capítulo não devem ser copiados mecanicamente. Eles devem ser adaptados ao *corpus*, à base de dados, à pergunta de pesquisa, às métricas utilizadas e às decisões de validação estabelecidas no protocolo.

A trilha deste livro, dedicada à Cientometria com Inteligência Artificial, parte do princípio de que a IA deve ampliar a capacidade analítica do pesquisador sem substituir seu juízo crítico. Isso significa reconhecer a utilidade dos modelos de linguagem para tarefas de apoio, mas também suas limitações. Modelos generativos são capazes de produzir textos fluentes, organizar argumentos, sugerir classificações e simular perspectivas analíticas. Contudo, a fluência não equivale a veracidade, a coerência textual não equivale a validade empírica e a segurança retórica não equivale a evidência científica. A engenharia de *prompts* deve, portanto, ser orientada por prudência epistêmica.

Uma engenharia de *prompts* cientometricamente responsável precisa responder a perguntas simples e exigentes: que tarefa o modelo deve executar? Com quais dados? Sob quais pressupostos? Com que nível de liberdade interpretativa? Em que formato? Como a resposta será conferida? Que partes da resposta correspondem a dados observados, inferências derivadas ou hipóteses? Como o *prompt* será registrado para permitir revisão posterior? Como comparar respostas produzidas por diferentes modelos, versões ou parâmetros? Como evitar que o modelo produza uma narrativa convincente, mas metodologicamente frágil?

Ao discutir esses problemas, o capítulo se aproxima de debates clássicos da cientometria sobre objetividade, normalização, interpretação de citações e uso responsável de indicadores. A novidade está em deslocar parte dessas questões para a interface textual entre pesquisador e IA. Assim como escolhas de base de dados, janela temporal, método de normalização e algoritmo de *cluster*ização precisam ser descritas, também os *prompts* relevantes devem ser documentados. Em estudos assistidos por IA, a transparência metodológica inclui registrar instruções, parâmetros, versões de modelos, dados de entrada, saídas relevantes, critérios de revisão e decisões humanas subsequentes (HICKS *et al.*, 2015; WALTMAN, 2016; BORNMANN; LEYDESDORFF, 2014).

O capítulo está organizado em seis seções. A primeira examina a anatomia de um bom *prompt* científico. A segunda detalha os componentes estruturais: contexto, papel, tarefa, dados, restrições e formato de saída. A terceira propõe uma tipologia funcional de *prompts* exploratórios, analíticos, críticos e redacionais. A quarta apresenta cadeias de *prompts* que acompanham o fluxo completo de um estudo cientométrico. A quinta trata da documentação de *prompts* em protocolos reprodutíveis. A sexta discute temperatura, variação, consistência e controle de respostas. Ao final, são oferecidos quatro modelos aplicados: modelo geral de *prompt* cientométrico, *prompt* de revisão crítica, *prompt* para gerar tabela analítica e *prompt* para transformar saída textual em protocolo.

5.2 Anatomia de um bom prompt científico

Figura 5.1 - Anatomia de um *prompt* científico

Fonte: Imagem gerada por IA, ChatGPT (2026).

Um bom *prompt* científico começa antes da escrita da instrução. Ele nasce de uma pergunta de pesquisa bem delimitada, de um entendimento mínimo sobre o *corpus* disponível e de uma definição clara do produto esperado. Em cientometria, isso significa saber se a tarefa se refere a produção, impacto, colaboração, circulação, visibilidade, estrutura intelectual, evolução temática ou política científica. Sem essa delimitação, o modelo tende a responder de modo genérico, reproduzindo fórmulas amplas sobre "principais tendências", "autores relevantes" ou "lacunas de pesquisa" sem vínculo suficiente com os dados analisados.

A primeira dimensão da anatomia do *prompt* é o objetivo. O objetivo informa por que a tarefa existe. Em vez de solicitar "analise estes dados", o pesquisador deve explicitar se pretende identificar temas emergentes, comparar padrões de colaboração, interpretar *clusters*, construir uma justificativa de escopo, revisar limitações metodológicas ou transformar resultados em linguagem acessível. Quanto mais claro o objetivo, menor a probabilidade de a IA preencher lacunas com suposições. Um objetivo bem formulado também ajuda a definir que tipo de resposta deve ser recusado. Se o objetivo é interpretar uma rede de coautoria, o modelo não deve inferir qualidade científica dos autores. Se o objetivo é resumir resumos de artigos, o modelo não deve inventar resultados não presentes nos textos.

A segunda dimensão é o contexto. O contexto informa ao modelo a natureza da investigação, a área científica, o recorte temporal, a base de dados, os critérios de inclusão e exclusão, a unidade de análise e as variáveis relevantes. Em muitos usos cotidianos de IA, o contexto é omitido porque o usuário espera que o modelo complemente a tarefa com conhecimento geral. Na pesquisa científica, essa prática é perigosa. O contexto deve reduzir a dependência de conhecimento implícito e aumentar a fidelidade aos dados fornecidos. Um *prompt* cientométrico deve indicar, por exemplo, se os registros vêm da Web of Science, Scopus, PubMed, Dimensions, OpenAlex, Lens, SciELO ou repositórios institucionais, pois cada fonte possui cobertura, metadados e vieses próprios.

A terceira dimensão é o papel atribuído ao modelo. Pedir que a IA atue como "especialista em cientometria" pode ser útil para direcionar o vocabulário e a estrutura da resposta, mas não deve ser confundido com garantia de competência. O papel é uma moldura operacional, não uma certificação. Em *prompts* científicos, o papel deve ser formulado de modo funcional: "atue como revisor metodológico", "atue como analista de indicadores", "atue como curador terminológico" ou "atue como redator técnico, preservando as limitações dos dados". Essa especificidade torna a resposta mais aderente ao trabalho desejado e reduz a tendência de produzir comentários genéricos.

A quarta dimensão é a tarefa. A tarefa deve ser expressa por verbos observáveis: classificar, comparar, agrupar, explicar, criticar, verificar, resumir, reescrever, transformar, extrair, listar, justificar, ordenar, rotular, identificar ambiguidades. Verbos vagos, como "analisar" ou "comentar", precisam ser desdobrados. Em cientometria, uma mesma palavra pode ocultar tarefas diferentes. "Analisar citações" pode significar calcular impacto, identificar documentos influentes, examinar redes de cocitação, discutir limitações de indicadores ou avaliar concentração de citações por área. O *prompt* deve transformar essa ambiguidade em procedimento.

A quinta dimensão é o conjunto de dados. Modelos de linguagem respondem melhor quando recebem dados organizados e instruções sobre como utilizá-los. Se o pesquisador fornece uma tabela com autores, anos, periódicos e citações, o *prompt* deve explicar a estrutura dessa tabela e indicar que a resposta deve se limitar aos campos fornecidos. Se o dado for textual, como títulos e resumos, o *prompt* deve orientar o modelo a citar evidências textuais. Se o dado for um resultado de software, como uma matriz de coocorrência ou lista de *clusters*, o *prompt* deve informar como os *clusters* foram gerados e quais métricas estão disponíveis. A frase "não use conhecimento externo para completar lacunas" pode ser decisiva quando se deseja re*string*ir a análise ao *corpus*.

A sexta dimensão é a restrição. Restrições são fundamentais porque modelos generativos tendem a responder mesmo quando os dados são insuficientes. Em estudos científicos, é necessário instruir o modelo a declarar insuficiência de evidência, distinguir hipótese de conclusão, evitar causalidade quando houver apenas associação, não inventar referências, não criar nomes de autores ausentes e não extrapolar além do período analisado. Restrições

também podem definir estilo, tamanho, idioma, público e nível de tecnicidade. Em um relatório para gestores, pode-se solicitar linguagem clara, mas sem simplificar indevidamente. Em um artigo acadêmico, pode-se pedir formalidade, citações e cautela interpretativa.

A sétima dimensão é o formato de saída. Um bom *prompt* define como a resposta deve ser apresentada. A saída pode ser uma tabela, uma lista hierarquizada, um resumo executivo, uma seção metodológica, um conjunto de categorias, uma matriz de riscos, um roteiro de validação ou um parágrafo interpretativo. O formato não é mero detalhe estético; ele altera a possibilidade de revisão. Uma resposta em tabela facilita comparar categorias, níveis de confiança e justificativas. Uma resposta em tópicos facilita revisão de escopo. Um texto corrido pode ser mais adequado para a versão final de relatório, mas menos apropriado para validação intermediária. A escolha do formato deve acompanhar a etapa do estudo.

A oitava dimensão é o critério de qualidade. *Prompts* científicos devem explicitar o que conta como boa resposta. Em cientometria, uma boa resposta pode ser aquela que evita generalizações, mantém aderência aos dados, apresenta justificativas, separa níveis de evidência, identifica limitações, usa vocabulário técnico adequado, reconhece ambiguidade e propõe verificações. Esse critério pode ser incorporado diretamente à instrução: "a resposta deve ser baseada apenas nos dados fornecidos; quando houver ambiguidade, marque para revisão; para cada interpretação, indique a evidência que a sustenta; ao final, liste limitações e verificações recomendadas".

A nona dimensão é a rastreabilidade. Um *prompt* cientométrico deve facilitar a reconstrução do caminho analítico. Isso significa pedir que o modelo associe cada conclusão a dados, campos ou trechos. Em uma classificação temática, o modelo deve indicar quais palavras, expressões ou resumos motivaram a atribuição de categoria. Em uma interpretação de indicadores, deve identificar qual tabela ou métrica sustenta cada afirmação. Em uma sugestão de busca bibliográfica, deve explicar por que determinados termos foram incluídos ou excluídos. A rastreabilidade não elimina erros, mas torna os erros identificáveis.

A décima dimensão é a metacognição controlada. Embora seja inadequado atribuir consciência ao modelo, pode ser útil solicitar que ele explicite incertezas, limitações e possíveis alternativas. Em vez de pedir apenas "dê a resposta", o pesquisador pode solicitar: "indique o grau de confiança", "liste interpretações alternativas", "aponte que evidência adicional seria necessária", "separe achados robustos de achados frágeis". Esses comandos não tornam a resposta automaticamente verdadeira, mas criam um espaço textual para revisão crítica.

Quadro 5.1 - As dez dimensões de um bom *prompt* científico

Nº	Dimensão	Pergunta que ela responde	Exemplo de instrução no prompt
01	**Objetivo**	Por que a tarefa existe?	"Identifique temas emergentes nos resumos; não avalie a qualidade dos autores."
02	**Contexto**	Sobre que dados e que recorte?	"*Corpus* de 320 resumos da Scopus, área de Educação, 2015-2024."
03	**Papel**	De que lugar o modelo deve responder?	"Atue como revisor metodológico em cientometria."
04	**Tarefa**	Que ação observável deve executar?	"Classifique cada resumo em uma das seis categorias e justifique."

Nº	Dimensão	Pergunta que ela responde	Exemplo de instrução no prompt
05	**Dados**	O que pode ser usado e como?	"Use apenas os campos título, resumo e palavras-chave fornecidos."
06	**Restrições**	O que não pode ser feito?	"Não invente referências; se a evidência for insuficiente, declare."
07	**Formato de saída**	Como a resposta deve ser apresentada?	"Responda em tabela: categoria, evidência e grau de confiança."
08	**Critério de qualidade**	O que conta como boa resposta?	"Cada interpretação deve indicar a evidência que a sustenta."
09	**Rastreabilidade**	Como reconstruir o caminho analítico?	"Indique as expressões do resumo que motivaram cada categoria."
10	**Metacognição controlada**	O que permanece incerto?	"Liste interpretações alternativas e o que ainda seria necessário verificar."
Regra geral: quanto mais explícito o prompt, menor o espaço para a IA preencher lacunas com suposições — e maior a possibilidade de revisão humana.			

Fonte: Elaborado pelo autor (2026).

A literatura sobre *prompting* mostra que instruções estruturadas podem melhorar o desempenho de modelos em tarefas complexas. A técnica de *chain-of-thought*, por exemplo, evidenciou ganhos em problemas que exigem decomposição, enquanto estratégias de *self-consistency* exploram múltiplos caminhos de raciocínio para aumentar estabilidade. Em cientometria, essas técnicas devem ser adaptadas com cautela. Não se trata de pedir ao modelo que "pense" livremente, mas de pedir que organize etapas verificáveis: identificar o tipo de dado, descrever o padrão observado, propor interpretação limitada, apontar evidência necessária e sugerir validação (WEI *et al.*, 2022; WANG *et al.*, 2023; LIU *et al.*, 2023).

Uma anatomia adequada do *prompt* também deve reconhecer a diferença entre tarefas abertas e fechadas. Tarefas fechadas possuem resposta mais delimitada: transformar uma lista de referências em formato padronizado, classificar registros segundo categorias previamente definidas, identificar duplicatas prováveis ou gerar uma tabela com campos fixos. Tarefas abertas envolvem interpretação, síntese e formulação de hipóteses. Quanto mais aberta for a tarefa, mais restrições e mecanismos de validação devem ser incluídos. Em uma análise de lacunas de pesquisa, por exemplo, a IA pode sugerir lacunas plausíveis, mas somente o pesquisador pode decidir se elas são empiricamente sustentadas pelo *corpus* e relevantes para a área.

O *prompt* científico deve ser escrito para reduzir ambiguidade sem eliminar a complexidade. A ciência não é uma sequência mecânica de respostas certas; envolve julgamento, contexto e interpretação. O objetivo não é aprisionar a IA em comandos rígidos, mas orientar sua colaboração de modo responsável. Bons *prompts* são suficientemente específicos para evitar derivações indevidas e suficientemente flexíveis para permitir que padrões inesperados sejam identificados. Essa tensão entre controle e abertura é central na engenharia de *prompts* para cientometria.

Um exemplo simples ilustra a diferença. O *prompt* "quais são os temas principais deste campo?" provavelmente produzirá uma resposta genérica. Um *prompt* mais adequado seria: "Com base exclusivamente nos títulos, resumos e palavras-chave fornecidos, identifique até sete temas recorrentes. Para cada tema, apresente rótulo, descrição, termos associados,

número aproximado de registros relacionados, exemplos de evidências textuais e grau de confiança. Marque temas ambíguos para revisão humana e não inclua temas sem evidência no *corpus*". A segunda formulação não garante perfeição, mas oferece um desenho mais controlável e revisável.

Outro exemplo envolve interpretação de indicadores. Perguntar "este programa de pós-graduação é bom?" é cientometricamente inadequado, pois mistura avaliação normativa, dados incompletos e critérios de qualidade indefinidos. Uma formulação mais responsável seria: "Com base nos indicadores de produção, colaboração e citação fornecidos, descreva padrões observáveis no período analisado. Não atribua qualidade acadêmica global. Separe produtividade, impacto bibliométrico e internacionalização. Aponte limitações dos indicadores e indique que evidências qualitativas seriam necessárias para avaliação mais abrangente". Esse tipo de *prompt* preserva a diferença entre medir, interpretar e avaliar.

Portanto, a anatomia de um bom *prompt* científico inclui pelo menos dez componentes: objetivo, contexto, papel, tarefa, dados, restrições, formato, critérios de qualidade, rastreabilidade e declaração de incerteza. Esses componentes não precisam aparecer sempre na mesma ordem, mas sua ausência deve ser deliberada, não acidental. Em estudos cientométricos assistidos por IA, o pesquisador deve perguntar, antes de executar qualquer *prompt*: a instrução deixa claro o que deve ser feito? Define quais dados podem ser usados? Impede inferências indevidas? Permite revisar a resposta? Produz uma saída útil ao estágio atual do estudo?

A engenharia de *prompts*, nesse sentido, aproxima-se da escrita de métodos. Assim como um protocolo bibliométrico descreve base de dados, período, estratégia de busca, critérios de limpeza e indicadores, um protocolo com IA deve descrever *prompts*, versões de modelos, parâmetros, dados de entrada, saídas e validações. O *prompt* é a parte visível de uma decisão analítica. Quando bem desenhado, ele transforma a IA em assistente metodológico. Quando mal desenhado, transforma a IA em fonte de ruído interpretativo.

5.3 Contexto, papel, tarefa, dados, restrições e formato de saída

A estrutura mínima de um *prompt* cientométrico pode ser organizada em seis componentes: contexto, papel, tarefa, dados, restrições e formato de saída. Essa estrutura não esgota todas as possibilidades, mas oferece uma base robusta para construir instruções auditáveis. Cada componente cumpre uma função específica na redução de ambiguidade e na produção de respostas verificáveis. Quando esses elementos são omitidos, o modelo tende a operar com base em padrões gerais de linguagem, e não nas exigências particulares do estudo.

O contexto responde à pergunta: em que situação científica a tarefa ocorre? Ele deve indicar o tema, o campo, a etapa do estudo, a finalidade da análise e, quando pertinente, o público destinatário. Em cientometria, o contexto também deve explicitar o tipo de *corpus*. Um *prompt* para analisar artigos de inteligência artificial em saúde pública deve ser diferente de um *prompt* para interpretar redes de coautoria em ciência dos materiais. Mesmo quando a estrutura da tarefa é semelhante, o vocabulário, as ambiguidades e os riscos de interpretação variam entre áreas. Termos como "impacto", "tradução", "rede", "modelo" e "desempenho" assumem significados diferentes conforme o domínio.

O contexto também informa o estágio do fluxo de trabalho. Na fase inicial, a IA pode ser usada para explorar vocabulário, mapear termos alternativos ou formular perguntas. Na etapa de coleta, pode apoiar a revisão de estratégias de busca. No pré-processamento, pode sugerir agrupamentos terminológicos. Na análise, pode auxiliar na leitura de tabelas e redes. Na

interpretação, pode produzir narrativas provisórias. Na comunicação, pode adaptar resultados a diferentes públicos. Cada estágio requer um tipo de *prompt*. Uma instrução adequada para *brainstorming* pode ser inadequada para redação final, porque permite liberdade excessiva.

O papel define a perspectiva operacional do modelo. Em vez de atribuir papéis vagos, recomenda-se especificar a função analítica. "Atue como bibliometrista especializado em normalização de palavras-chave" produz uma orientação diferente de "atue como redator de relatório executivo". O papel deve estar subordinado à tarefa, não ao prestígio simbólico. Dizer "atue como pesquisador sênior" é menos útil do que dizer "atue como revisor metodológico, procurando inconsistências entre pergunta, dados, indicadores e conclusões". A segunda formulação orienta ação observável.

Na prática, papéis podem ser combinados com limites. Um *prompt* pode dizer: "Atue como analista cientométrico, mas não formule conclusões avaliativas sobre qualidade institucional; limite-se a padrões observáveis nos dados". Essa formulação impede que o modelo assuma autoridade avaliativa indevida. Em contextos de gestão de pesquisa, tal restrição é crucial, pois indicadores podem afetar decisões sobre financiamento, reputação e prioridades institucionais. A IA não deve converter dados parciais em julgamento global.

A tarefa é o núcleo executivo do *prompt*. Ela deve definir exatamente o que o modelo deve entregar. A tarefa pode ser simples, como "padronize os nomes das instituições", ou complexa, como "compare três interpretações possíveis para a evolução temporal de *clusters* temáticos". Em ambos os casos, é recomendável decompor a tarefa em passos. Essa decomposição reduz omissões e facilita a revisão. Por exemplo: "primeiro identifique os termos equivalentes; depois proponha um termo preferencial; em seguida, indique casos ambíguos; por fim, apresente uma tabela com justificativa".

A decomposição da tarefa é especialmente útil quando a resposta exige julgamento. Em classificação temática, pode-se pedir que o modelo leia cada resumo, identifique termos centrais, atribua categoria, justifique com evidência textual e declare confiança. Em interpretação de mapas, pode-se pedir descrição estrutural, interpretação possível, limitações e necessidade de validação. Em revisão de metodologia, pode-se pedir identificação de incoerências entre objetivo, base de dados, indicadores e conclusões. Essa estruturação aproxima o *prompt* de um formulário analítico.

O componente dado responde à pergunta: que material deve ser usado? Em estudos cientométricos, os dados podem estar em tabelas, listas de registros, resumos, matrizes, saídas de softwares, gráficos, descrições de redes ou textos metodológicos. O *prompt* deve instruir o modelo a usar apenas os dados fornecidos quando o objetivo for análise empírica. Caso se aceite conhecimento externo, isso deve ser explicitado. Há diferença entre pedir "explique teoricamente o que é cocitação" e pedir "interprete a rede de cocitação extraída deste *corpus*". No segundo caso, o modelo não deve inferir documentos ou relações ausentes.

A organização dos dados influencia a qualidade da resposta. Tabelas com cabeçalhos claros, campos consistentes e unidades explícitas reduzem erros. Quando se envia ao modelo uma lista de *clusters*, é importante informar se os termos foram extraídos de títulos, resumos, palavras-chave dos autores ou palavras-chave indexadas. É necessário indicar se o número associado ao termo representa frequência, centralidade, ocorrência ponderada ou outro indicador. Sem isso, a IA pode interpretar equivocadamente valores numéricos e produzir narrativa errada.

As restrições delimitam o espaço de liberdade. Elas são indispensáveis em *prompts* científicos porque modelos de linguagem são treinados para completar padrões plausíveis, não para reconhecer automaticamente limites metodológicos. Restrições comuns incluem: não

inventar dados, não usar fontes externas, não afirmar causalidade, não extrapolar para além do período analisado, não atribuir intenção a autores, não classificar qualidade sem critérios explícitos, não gerar referências inexistentes e não ocultar incertezas. Essas restrições devem ser escritas de modo direto.

Há restrições de conteúdo, forma e epistemologia. Restrições de conteúdo definem o que pode ser incluído. Restrições de forma definem tamanho, idioma, estrutura, tabela ou texto. Restrições epistemológicas definem o grau de certeza permitido. Em cientometria, esse terceiro tipo é decisivo. O modelo deve ser instruído a distinguir "o dado mostra", "é possível inferir" e "seria necessário investigar". Essa separação impede que hipóteses apareçam como resultados consolidados.

O formato de saída define a materialidade da resposta. Uma tabela com colunas "achado", "evidência", "interpretação", "limitação" e "verificação necessária" é mais auditável do que um parágrafo narrativo. Uma matriz de riscos permite revisão sistemática. Uma lista ordenada pode ser útil para planejamento. Um texto acadêmico pode ser útil para redação final. Em cada caso, o formato deve estar alinhado ao uso. Para etapa exploratória, formatos flexíveis são aceitáveis. Para documentação metodológica, formatos padronizados são preferíveis.

Um *prompt* cientométrico robusto pode ser compreendido como contrato de interação. O pesquisador oferece contexto, papel, tarefa, dados, restrições e formato; o modelo devolve uma resposta dentro desses limites. Como todo contrato, esse arranjo depende de clareza e fiscalização. O modelo pode descumprir instruções, ignorar restrições ou apresentar informações não solicitadas. Por isso, *prompts* críticos e procedimentos de validação precisam acompanhar *prompts* analíticos.

Considere um estudo sobre a evolução da literatura em "inteligência artificial explicável" entre 2015 e 2025. Um *prompt* frágil seria: "analise a evolução do tema". Um *prompt* estruturado seria: "Contexto: estudo cientométrico sobre inteligência artificial explicável, com registros extraídos da Scopus entre 2015 e 2025. Papel: atue como analista cientométrico. Tarefa: interpretar a tabela anual de publicações e os *clusters* de palavras-chave. Dados: use apenas a tabela e a lista de *clusters* fornecidas. Restrições: não afirme causalidade, não invente eventos externos, diferencie crescimento quantitativo de consolidação conceitual. Formato: apresente uma tabela com padrão observado, evidência nos dados, interpretação possível, limitação e verificação recomendada". Essa formulação converte uma solicitação ampla em procedimento controlável.

O mesmo princípio vale para tarefas redacionais. Pedir "escreva a discussão" pode produzir um texto convincente, porém arriscado. Um *prompt* melhor seria: "Escreva uma seção de discussão com base nos resultados abaixo. Cada parágrafo deve relacionar um achado a uma interpretação, mencionar pelo menos uma limitação e evitar extrapolações. Não inclua referências que não estejam na lista fornecida. Quando houver incerteza, use formulações cautelosas, como 'os resultados sugerem' ou 'é plausível interpretar'". A engenharia de *prompts* não elimina a necessidade de revisão, mas melhora a qualidade do primeiro rascunho.

Outro componente importante é a audiência. Em cientometria, o mesmo resultado pode ser comunicado a especialistas, gestores, professores, estudantes ou público amplo. A adaptação de linguagem deve preservar a precisão. Um *prompt* para público não especialista deve solicitar analogias com cuidado e proibir simplificações que distorçam conceitos. Por exemplo: "explique índice h, fator de impacto e citações normalizadas para gestores sem formação bibliométrica, mantendo distinções conceituais e indicando limitações". Essa formulação evita que a IA transforme indicadores em rankings simplistas.

A sequência dos componentes também pode influenciar a resposta. Em geral, recomenda-se iniciar pelo contexto, seguir pelo papel, definir tarefa, apresentar dados, explicitar restrições e finalizar com formato. Essa ordem ajuda o modelo a interpretar a tarefa antes de produzir a saída. Contudo, em *prompts* longos, é útil usar marcadores e títulos: "Contexto", "Papel", "Tarefa", "Dados", "Restrições", "Formato de saída". Essa estrutura facilita tanto a execução pelo modelo quanto a revisão humana do próprio *prompt*.

Na prática, *prompts* cientométricos devem ser versionados. A primeira versão pode produzir resposta insuficiente; o pesquisador então ajusta restrições, formato ou exemplos. Esse processo de iteração é normal. O problema surge quando as versões não são registradas. Se um relatório final foi influenciado por várias interações com IA, mas apenas a última resposta é preservada, perde-se parte da trilha metodológica. Portanto, a engenharia de *prompts* deve incluir um sistema de registro: identificador do *prompt*, data, modelo, versão, parâmetros, dados de entrada, saída e decisão tomada.

A inclusão de exemplos no *prompt* pode melhorar consistência, sobretudo em tarefas de classificação. Quando há categorias previamente definidas, fornecer dois ou três exemplos classificados ajuda o modelo a aplicar critérios semelhantes. Essa prática se aproxima do *few-shot prompting*. Em cientometria, exemplos podem ser usados para normalizar instituições, classificar tipos de colaboração ou rotular temas. Todavia, exemplos devem ser escolhidos com cuidado, pois podem induzir vieses. Se os exemplos forem estreitos, a IA pode forçar casos diferentes nas mesmas categorias (BROWN *et al.*, 2020).

Um aspecto frequentemente negligenciado é o controle de escopo. *Prompts* muito amplos tendem a gerar respostas longas e pouco verificáveis. *Prompts* excessivamente estreitos podem impedir identificação de padrões relevantes. O equilíbrio deve ser definido pelo estágio do estudo. No início, pode-se usar *prompts* exploratórios amplos, desde que a saída seja tratada como hipótese. Nas etapas analíticas e redacionais, os *prompts* devem ser mais restritivos. Essa progressão acompanha a lógica do método científico: exploração, delimitação, teste, interpretação e comunicação.

A engenharia de *prompts* também precisa incorporar a noção de falha aceitável. Em algumas tarefas, uma resposta incompleta é preferível a uma resposta inventada. Por isso, *prompts* devem autorizar o modelo a dizer que não há informação suficiente. A instrução "se os dados não permitirem responder, declare insuficiência de evidência" é simples, mas metodologicamente poderosa. Ela combate a tendência dos modelos a preencher lacunas com plausibilidade verbal.

Em síntese, contexto, papel, tarefa, dados, restrições e formato constituem o esqueleto de um *prompt* científico. O contexto localiza, o papel orienta, a tarefa executa, os dados fundamentam, as restrições protegem e o formato torna revisável. Em cientometria, esses elementos são indispensáveis porque os resultados frequentemente transitam entre descrição quantitativa, interpretação sociológica e decisão institucional. Quanto maior o impacto potencial de uma análise, maior deve ser o rigor na construção do *prompt*.

5.4 Prompts exploratórios, analíticos, críticos e redacionais

Nem todos os *prompts* cumprem a mesma função em um estudo cientométrico. Uma das causas de uso inadequado da IA na pesquisa é tratar todo *prompt* como se fosse um pedido de resposta final. Na verdade, *prompts* podem ser classificados conforme sua função no processo investigativo. Neste capítulo, propõe-se uma tipologia operacional composta por quatro tipos: *prompts* exploratórios, *prompts* analíticos, *prompts* críticos e *prompts* redacionais. Essa tipologia ajuda a definir expectativas, controles e formas de validação para cada interação.

Prompts exploratórios são usados para abrir possibilidades. Eles auxiliam na formulação de perguntas, expansão de vocabulário, identificação de dimensões conceituais, levantamento de indicadores possíveis e reconhecimento inicial de riscos. Seu produto não deve ser tratado como resultado, mas como insumo para reflexão. Um *prompt* exploratório pode perguntar: "quais recortes cientométricos poderiam ser aplicados ao tema sustentabilidade urbana?" A resposta pode sugerir produção, colaboração internacional, redes de palavras-chave, evolução temporal ou relação entre disciplinas. Essas sugestões precisam ser avaliadas pelo pesquisador à luz da literatura, da disponibilidade de dados e da finalidade do estudo.

A principal virtude do *prompt* exploratório é ampliar o campo de visão. Pesquisadores iniciantes podem não conhecer todas as possibilidades analíticas; gestores podem formular demandas vagas; equipes interdisciplinares podem utilizar vocabulários distintos. A IA pode ajudar a traduzir uma preocupação ampla em alternativas metodológicas. Todavia, a exploração também é o momento de maior risco de plausibilidade vazia. O modelo pode sugerir indicadores inadequados, bases de dados sem cobertura suficiente ou categorias conceituais genéricas. Por isso, *prompts* exploratórios devem solicitar justificativas, limitações e condições de aplicabilidade.

Um *prompt* exploratório adequado poderia ser: "A partir do tema [tema], proponha cinco perguntas cientométricas possíveis. Para cada pergunta, indique unidade de análise, dados necessários, indicadores apropriados, visualizações possíveis, riscos metodológicos e grau de viabilidade. Não apresente as sugestões como recomendações finais; trate-as como alternativas para avaliação". Essa formulação reduz o risco de aceitar a primeira resposta como plano definitivo.

Prompts analíticos são usados quando há dados disponíveis e uma tarefa definida. Eles devem ser mais restritivos. Seu objetivo é transformar dados em saídas intermediárias: tabelas interpretativas, classificações, agrupamentos, descrições de padrões, comparações entre indicadores ou sínteses estruturadas. Ao contrário dos *prompts* exploratórios, *prompts* analíticos devem limitar explicitamente o uso de conhecimento externo. Se o objetivo é analisar um *corpus*, a resposta deve estar ancorada nesse *corpus*. Se a IA não encontra evidência suficiente, deve indicar lacuna.

Em uma análise de palavras-chave, um *prompt* analítico pode solicitar: "Agrupe os termos abaixo em categorias temáticas. Use apenas os termos fornecidos. Para cada categoria, apresente rótulo, termos incluídos, justificativa semântica e casos ambíguos. Não crie categorias sem pelo menos três termos associados". Essa instrução transforma a IA em apoio de curadoria terminológica. O pesquisador ainda deve revisar os agrupamentos, mas recebe uma estrutura inicial auditável.

Prompts analíticos são úteis para tarefas que exigem combinação de leitura textual e organização. Modelos de linguagem têm bom desempenho relativo em sumarização, extração de entidades, classificação semântica e reformulação textual, desde que recebam dados claros e critérios definidos. Em cientometria, isso pode apoiar normalização de variações terminológicas, identificação de tópicos em resumos, interpretação de *clusters* de termos, elaboração de descrições de mapas e organização de resultados. Ainda assim, a IA não substitui cálculos, verificações estatísticas ou validação especializada.

Uma regra importante é não pedir ao modelo que calcule indicadores complexos quando ferramentas adequadas estão disponíveis. Embora modelos possam realizar operações simples, eles não são substitutos confiáveis para scripts estatísticos, planilhas ou pacotes especializados. O *prompt* analítico deve preferencialmente solicitar interpretação de saídas já calculadas por ferramentas apropriadas. Por exemplo, em vez de pedir "calcule modularidade

da rede", é melhor usar software de rede para calcular a métrica e pedir à IA que explique, com cautela, o significado do valor no contexto do estudo.

Prompts críticos têm função avaliativa. Eles servem para revisar escolhas metodológicas, identificar inconsistências, detectar extrapolações, avaliar limitações e testar a robustez de argumentos. Em um fluxo cientométrico com IA, *prompts* críticos são tão importantes quanto *prompts* analíticos. A facilidade de produzir texto pode induzir confiança excessiva. Um *prompt* crítico atua como mecanismo de contrapeso, solicitando que o modelo examine o próprio resultado ou uma seção redigida.

Um *prompt* crítico pode perguntar: "Revise a interpretação abaixo e identifique afirmações que excedem os dados. Classifique cada problema como causalidade indevida, generalização excessiva, ausência de evidência, confusão entre indicador e qualidade, ou necessidade de referência. Sugira reformulação cautelosa". Esse tipo de *prompt* é especialmente útil antes de inserir resultados em artigos, relatórios institucionais ou documentos de política científica.

Prompts críticos também podem ser aplicados ao protocolo. Antes de iniciar a coleta de dados, o pesquisador pode pedir: "Avalie a coerência entre pergunta de pesquisa, base de dados, período, estratégia de busca, indicadores e visualizações propostas. Aponte riscos de viés, lacunas de cobertura e decisões que precisam ser justificadas". A resposta ajuda a antecipar problemas. Contudo, a avaliação do modelo deve ser tratada como revisão auxiliar, não como aprovação metodológica.

Prompts redacionais são usados para transformar análises em texto comunicável. Eles podem produzir seções de metodologia, resultados, discussão, resumos executivos, notas técnicas, legendas de gráficos ou apresentações para públicos específicos. A principal exigência é preservar fidelidade aos dados e às limitações. *Prompts* redacionais devem receber material analítico já validado, não dados brutos ambíguos. A IA é útil para organizar linguagem, melhorar coesão e adaptar estilo, mas não deve introduzir resultados novos.

Um *prompt* redacional responsável pode dizer: "Escreva uma seção de resultados com base exclusivamente na tabela interpretativa validada abaixo. Não acrescente resultados. Use tom formal científico. Diferencie descrição de interpretação. Inclua limitações quando necessário. Evite expressões avaliativas como 'melhor', 'superior' ou 'excelente' sem critério explícito". Essa formulação impede que o modelo transforme uma redação em inferência adicional.

Os quatro tipos de *prompts* podem ser encadeados. Em uma etapa inicial, um *prompt* exploratório ajuda a transformar tema em perguntas. Em seguida, *prompts* analíticos organizam dados. Depois, *prompts* críticos revisam resultados e interpretações. Finalmente, *prompts* redacionais produzem textos. Essa sequência reduz a probabilidade de que uma resposta criativa e não validada seja usada diretamente como conclusão. O ciclo ideal alterna geração e crítica, abertura e controle.

A tipologia também ajuda a definir critérios de avaliação. *Prompts* exploratórios devem ser avaliados por diversidade, relevância e viabilidade. *Prompts* analíticos devem ser avaliados por aderência aos dados, consistência de classificação, rastreabilidade e ausência de invenção. *Prompts* críticos devem ser avaliados por capacidade de identificar riscos reais e propor correções úteis. *Prompts* redacionais devem ser avaliados por fidelidade, clareza, coerência e preservação de limites metodológicos.

Outro aspecto é o nível de liberdade. *Prompts* exploratórios podem operar com maior abertura, desde que a saída seja marcada como provisória. *Prompts* analíticos exigem liberdade moderada, pois precisam organizar dados sem inventar. *Prompts* críticos devem ter liberdade para apontar problemas, mas precisam de categorias claras de crítica. *Prompts* redacionais

devem ter liberdade estilística, porém baixa liberdade substantiva. Essa diferenciação pode ser incorporada ao próprio *prompt*: "esta é uma etapa exploratória; apresente alternativas"; ou "esta é uma etapa de redação final; não introduza novos resultados".

A distinção entre os tipos de *prompts* também protege contra a ilusão de completude. Um modelo pode produzir uma resposta bem escrita a um *prompt* exploratório e dar a impressão de que o estudo já está desenhado. Pode gerar uma discussão elegante e dar a impressão de que a interpretação está validada. A tipologia obriga o pesquisador a perguntar: em que etapa estou? Esta resposta é hipótese, análise, crítica ou texto final? Que validação é necessária antes de avançar?

Em estudos cientométricos, *prompts* exploratórios são particularmente úteis para delimitar campos emergentes, nos quais a terminologia ainda é instável. Eles podem sugerir sinônimos, variantes e termos relacionados. Entretanto, esses termos precisam ser testados em bases bibliográficas, pois a IA pode incluir expressões raramente usadas na literatura indexada. *Prompts* analíticos, por sua vez, são úteis após a coleta, quando títulos, resumos e palavras-chave já permitem identificar padrões. *Prompts* críticos são indispensáveis na interpretação de métricas, pois ajudam a evitar o uso normativo acrítico de indicadores. *Prompts* redacionais apoiam a comunicação, especialmente quando os resultados precisam ser adaptados a gestores.

A integração dos quatro tipos deve ser documentada. Um protocolo pode indicar: P1, *prompt* exploratório para formulação de perguntas; P2, *prompt* analítico para classificação de termos; P3, *prompt* crítico para revisão de interpretação; P4, *prompt* redacional para elaboração da seção de resultados. Para cada *prompt*, devem ser registrados dados de entrada, saída, versão do modelo e decisão humana. Essa prática transforma interações dispersas em trilha metodológica.

A literatura sobre padrões de *prompts* sugere que instruções podem ser concebidas como padrões reutilizáveis. Em cientometria, isso significa que equipes podem desenvolver bibliotecas internas de *prompts* para tarefas recorrentes: normalização institucional, explicação de indicadores, revisão de causalidade, elaboração de relatórios, construção de glossários e análise de *clusters*. A reutilização aumenta eficiência, mas não deve levar à automação irrefletida. Cada *prompt*-padrão precisa ser adaptado ao estudo e revisado periodicamente (WHITE *et al.*, 2023; LIU *et al.*, 2023).

Prompts exploratórios, analíticos, críticos e redacionais também podem ser combinados com papéis diferentes. O mesmo conjunto de dados pode ser submetido a um *prompt* analítico com papel de "curador terminológico" e depois a um *prompt* crítico com papel de "revisor de vieses". Essa mudança de perspectiva pode revelar problemas ocultos. Contudo, não se deve interpretar divergências entre respostas como verdade plural automática. Divergências são sinais para investigação, não decisões finais.

Quadro 5.2 - Quatro tipos de *prompt* no trabalho cientométrico

1. EXPLORATÓRIOS *Abrir possibilidades e mapear o campo*	2. ANALÍTICOS *Organizar dados e produzir resultados intermediários*	3. CRÍTICOS *Avaliar, questionar e testar robustez*	4. REDACIONAIS *Transformar análises em texto comunicável*
Função			
Gerar alternativas, expandir vocabulário, levantar dimensões, indicadores e possibilidades metodológicas.	Classificar, agrupar, extrair, comparar, explicar padrões e descrever resultados com base nos dados fornecidos.	Revisar métodos, indicadores, interpretações e resultados; identificar problemas, vieses, excessos de inferência e lacunas.	Redigir seções, resumos, notas técnicas, legendas, apresentações e textos para diferentes públicos.

1. EXPLORATÓRIOS *Abrir possibilidades e mapear o campo*	2. ANALÍTICOS *Organizar dados e produzir resultados intermediários*	3. CRÍTICOS *Avaliar, questionar e testar robustez*	4. REDACIONAIS *Transformar análises em texto comunicável*
Características			
Alta abertura; respostas provisórias. Foco em diversidade e viabilidade. Expõe opções e condições de aplicabilidade. Não produz conclusão; produz insumos.	Aderência estrita ao *corpus* e aos dados. Tarefas definidas e delimitadas. Produz saídas estruturadas e auditáveis. Exige justificativas claras sobre dados e limites.	Perspectiva avaliativa e cética. Identifica riscos e inconsistências. Sugere correções, alternativas e verificações. Essencial antes de conclusões e decisões.	Baixa liberdade substantiva; foco em fidelidade. Adapta linguagem e formato ao público. Preserva limitações e distinção entre níveis de evidência. Texto claro, coerente e tecnicamente preciso.
Exemplo de prompt			
"A partir do tema sustentabilidade urbana, proponha cinco perguntas cientométricas possíveis. Para cada uma, indique unidade de análise, dados necessários, indicadores apropriados, visualizações possíveis, riscos metodológicos e grau de viabilidade. Trate como alternativas para avaliação, não como recomendações finais."	"Classifique os 50 resumos abaixo em categorias temáticas predefinidas: (1) políticas públicas, (2) tecnologia, (3) comportamento, (4) planejamento, (5) outro. Para cada resumo, atribua categoria, justifique com trechos do texto e informe o grau de confiança (alto, médio ou baixo). Não use conhecimento externo. Marque casos ambíguos."	"Revise a interpretação abaixo e identifique afirmações que excedem os dados. Classifique cada problema como: (a) causalidade indevida; (b) generalização excessiva; (c) ausência de evidência; (d) confusão entre indicador e qualidade; (e) necessidade de referência. Sugira reformulação cautelosa e verificações necessárias."	"Redija a seção de resultados com base exclusivamente na tabela validada abaixo. Não interprete além do que os dados mostram. Utilize tom científico formal. Evite termos avaliativos como "melhor", "superior" ou "excelente" sem critério explícito. Inclua limitações relevantes."
Riscos principais			
Sugestões genéricas ou irrelevantes. Indicadores inadequados ou indisponíveis. Cobertura de base superestimada.	Inventar informação não presente nos dados. Sobrerrepresentar categorias dominantes. Erros de classificação e ambiguidade não sinalizada.	Críticas superficiais ou irrelevantes. Não identificar vieses estruturais. Aceitar achados frágeis como robustos.	Adicionar resultados não presentes nos dados. Exagerar conclusões ou minimizar limitações. Estilo inadequado ao público ou à finalidade.

Fonte: Elaborado pelo autor (2026).

A engenharia de *prompts* para cientometria deve, portanto, operar como sistema de funções. O pesquisador não pergunta apenas "qual *prompt* usar?", mas "qual função preciso cumprir agora?". Se a função é descobrir possibilidades, usa-se exploração. Se é organizar dados, usa-se análise. Se é testar robustez, usa-se crítica. Se é comunicar, usa-se redação. Essa clareza funcional é uma das principais condições para usar IA de modo metodologicamente responsável.

5.5 Cadeias de prompts: da pergunta ao relatório

A noção de cadeia de *prompts* refere-se ao uso sequencial e articulado de múltiplas instruções para conduzir uma tarefa complexa. Em vez de solicitar ao modelo que produza, em uma única resposta, um estudo completo, o pesquisador divide o processo em etapas. Cada etapa gera uma saída intermediária que pode ser revisada, corrigida e usada como entrada para a etapa seguinte. Em cientometria, essa abordagem é especialmente

adequada, pois o estudo envolve formulação de pergunta, busca, coleta, limpeza, análise, interpretação e comunicação.

Uma cadeia de *prompts* bem desenhada se aproxima de um fluxo de trabalho. Ela começa com uma demanda inicial, como "mapear a produção científica sobre aprendizagem federada em saúde", e termina com um produto, como relatório, artigo ou apresentação. Entre esses extremos, há decisões: qual base usar, quais termos incluir, qual período analisar, quais indicadores calcular, como tratar duplicatas, como interpretar redes, como comunicar limitações. A IA pode apoiar cada decisão, mas não deve fundi-las em um texto único sem controle.

A primeira etapa de uma cadeia é a transformação da demanda em problema cientométrico. Um *prompt* pode pedir que a IA converta uma ideia vaga em perguntas operacionalizáveis, indicando unidades de análise e dados necessários. A saída deve ser revisada pelo pesquisador. A segunda etapa é a delimitação do escopo: período, bases, idiomas, tipos de documento e critérios de inclusão. A terceira é a construção da estratégia de busca, com termos, sinônimos, operadores booleanos e justificativas. A quarta é a revisão crítica da estratégia, buscando ruídos e lacunas.

Após a coleta, a cadeia passa para o pré-processamento. *Prompts* podem apoiar identificação de duplicatas prováveis, normalização de palavras-chave, padronização de instituições e agrupamento de variantes terminológicas. Nessa etapa, é importante que as saídas sejam tabelas revisáveis. O modelo deve marcar casos ambíguos, não os resolver arbitrariamente. A decisão final deve permanecer humana, pois erros de normalização podem alterar redes, frequências e indicadores.

A etapa analítica inclui interpretação de tabelas e mapas. Um *prompt* pode solicitar descrição de crescimento anual, distribuição por periódicos, colaboração entre países ou estrutura temática. Outro *prompt* pode revisar se a interpretação exagera os dados. A cadeia deve alternar *prompts* de análise e *prompts* de crítica. Essa alternância evita que a IA produza uma narrativa linear demais, ignorando limites, anomalias e incertezas.

A etapa redacional transforma resultados validados em texto. *Prompts* podem gerar seções de metodologia, resultados, discussão e conclusão. Entretanto, a cadeia deve incluir um *prompt* final de auditoria, verificando se o texto menciona apenas resultados efetivamente produzidos, se as limitações estão presentes e se as referências existem. Em estudos com IA, o último *prompt* não deveria ser "melhore o texto", mas "verifique aderência entre texto, dados, método e evidências".

Uma cadeia de *prompts* pode ser representada como sequência: P1 explorar tema; P2 formular perguntas; P3 definir escopo; P4 construir busca; P5 revisar busca; P6 planejar coleta; P7 limpar dados; P8 classificar termos; P9 interpretar indicadores; P10 interpretar redes; P11 revisar inferências; P12 redigir resultados; P13 redigir discussão; P14 auditar relatório. Essa sequência não é fixa, mas ilustra a lógica de decomposição. Cada *prompt* tem função, entrada, saída e validação.

A vantagem principal das cadeias é a redução de opacidade. Quando uma resposta final surge de um único *prompt* longo, torna-se difícil saber quais decisões foram tomadas pelo modelo. Em uma cadeia, decisões ficam distribuídas e documentáveis. O pesquisador pode interromper o processo, corrigir uma classificação, ajustar categorias e retomar. A cadeia transforma o uso de IA em processo iterativo e controlado.

Outra vantagem é a possibilidade de especialização de *prompts*. Um *prompt* para revisar estratégia de busca não precisa saber redigir discussão. Um *prompt* para normalizar palavras-

chave não precisa interpretar política científica. Essa especialização reduz interferências. Modelos de linguagem tendem a misturar tarefas quando recebem pedidos amplos. Cadeias bem desenhadas mantêm fronteiras: primeiro curar dados, depois interpretar; primeiro descrever, depois discutir; primeiro identificar risco, depois redigir.

Cadeias de *prompts* também favorecem colaboração em equipe. Em um grupo de pesquisa, diferentes membros podem revisar etapas específicas: bibliotecários podem avaliar *prompts* de busca; especialistas de domínio podem revisar categorias temáticas; cientometristas podem revisar indicadores; gestores podem revisar clareza do relatório executivo. O registro da cadeia permite distribuir responsabilidade. Isso é preferível a uma interação individual não documentada.

A construção de cadeias deve considerar dependências. Uma saída incorreta em etapa inicial contamina etapas posteriores. Se a estratégia de busca é inadequada, a análise temática será enviesada. Se a normalização institucional é falha, a rede de colaboração será distorcida. Se os *clusters* são rotulados de modo impreciso, a discussão conceitual será frágil. Por isso, pontos de validação devem ser inseridos após etapas críticas. Uma cadeia responsável não apenas produz; ela verifica.

Um método útil é definir "portões de qualidade". Após P4, a estratégia de busca só avança se for revisada por humano ou comparada com artigos sentinela. Após P7, a limpeza de dados só avança se amostras forem conferidas. Após P9, a interpretação de indicadores só avança se as métricas forem compreendidas e contextualizadas. Após P13, o texto só avança se não houver afirmações sem evidência. Esses portões podem ser incorporados ao protocolo.

Cadeias de *prompts* se relacionam com técnicas de decomposição em modelos de linguagem, mas possuem finalidade metodológica própria. Em tarefas de raciocínio, a decomposição melhora desempenho. Em pesquisa científica, a decomposição melhora transparência. A diferença é importante. O objetivo não é apenas obter uma resposta melhor, mas tornar o processo revisável. Uma cadeia cientométrica não deve ocultar etapas em uma resposta final; deve expô-las (WEI *et al.*, 2022; WANG *et al.*, 2023).

A construção da cadeia deve evitar dependência excessiva do texto gerado pelo modelo. Sempre que possível, dados estruturados devem ser preservados fora da conversa, em planilhas, scripts ou arquivos de análise. A IA pode comentar, reorganizar ou interpretar, mas a fonte primária dos resultados deve permanecer acessível. Isso impede que a conversa se torne o único repositório do estudo. Um protocolo reprodutível deve permitir reconstruir as etapas sem depender da memória do chat.

Um exemplo de cadeia aplicada a um estudo de colaboração internacional pode seguir este percurso. Primeiro, o pesquisador pede à IA que transforme a demanda em perguntas: "como evoluiu a colaboração internacional em nanotecnologia no Brasil?" Segundo, solicita revisão de escopo: período, bases, tipos documentais. Terceiro, constrói estratégia de busca. Quarto, coleta dados em ferramenta apropriada. Quinto, usa *prompt* para planejar limpeza de países e instituições. Sexto, calcula indicadores em software. Sétimo, pede à IA interpretação cautelosa das tabelas. Oitavo, usa *prompt* crítico para identificar riscos de confundir coautoria internacional com cooperação efetiva. Nono, redige relatório. Décimo, audita coerência.

Outro exemplo envolve mapas temáticos. A cadeia começa com definição do tema, passa por busca, coleta de palavras-chave, normalização, geração de rede em software, identificação de *clusters* e rótulos. A IA pode apoiar a rotulação dos *clusters*, mas o *prompt* deve solicitar evidências: termos mais frequentes, relações semânticas e alternativas de rótulo. Depois, um *prompt* crítico revisa se os rótulos são excessivamente amplos. Finalmente, um *prompt* redacional transforma a interpretação em texto. Nesse fluxo, a IA não cria o mapa; ela auxilia sua interpretação.

Cadeias de *prompts* também podem incorporar comparação entre respostas. Em tarefas interpretativas, pode-se pedir ao modelo que produza duas interpretações alternativas e depois use um *prompt* crítico para comparar plausibilidade. Outra estratégia é solicitar que o mesmo *prompt* seja executado mais de uma vez com baixa temperatura, comparando consistência. Divergências devem ser registradas. A consistência entre respostas não prova verdade, mas inconsistência sinaliza necessidade de revisão.

A cadeia deve incluir *prompts* de documentação. Após cada etapa, pode-se pedir: "registre a decisão tomada, a justificativa, a entrada utilizada, a saída produzida e a validação necessária". Esse tipo de *prompt* cria metadados metodológicos. Em vez de documentar apenas ao final, o pesquisador documenta durante o processo. Essa prática reduz esquecimento e melhora transparência.

Um risco das cadeias é a acumulação de erros linguísticos ou conceituais. Se a saída de um *prompt* for aceita automaticamente como entrada do próximo, o modelo pode reforçar uma interpretação incorreta. A cadeia precisa de intervenção humana. Em cada etapa, o pesquisador deve decidir se aceita, corrige, descarta ou refaz a saída. A IA não deve controlar a cadeia sem supervisão, especialmente em estudos que podem orientar decisões institucionais.

Outro risco é a proliferação de *prompts* desnecessários. Nem toda tarefa precisa de IA. Cálculos, extrações e visualizações podem ser melhor executados por scripts e softwares especializados. A cadeia deve reservar IA para tarefas em que linguagem, classificação, interpretação ou comunicação agreguem valor. A pergunta metodológica não é "como usar IA em tudo?", mas "em que etapa a IA reduz custo cognitivo sem reduzir rigor?".

Cadeias de *prompts* bem desenhadas produzem dois resultados: o produto científico e o rastro metodológico. O produto pode ser artigo, relatório ou mapa. O rastro inclui perguntas, escopo, *prompts*, dados, saídas, revisões e decisões. Em uma cultura de ciência aberta e métricas responsáveis, esse rastro é tão importante quanto o produto final. Ele permite que outros pesquisadores compreendam como a IA participou do estudo.

Em síntese, a cadeia de *prompts* é a arquitetura operacional da cientometria assistida por IA. Ela substitui a interação improvisada por fluxo controlado. Ao decompor a análise, especializar tarefas, inserir portões de qualidade e registrar decisões, o pesquisador transforma o modelo generativo em instrumento auxiliar. O resultado esperado não é uma automação plena da cientometria, mas uma prática aumentada, na qual a capacidade de organização textual da IA é subordinada ao método científico.

5.6 Como documentar prompts em protocolos reprodutíveis

A documentação de *prompts* é uma das condições fundamentais para que estudos cientométricos com IA possam ser avaliados, reproduzidos ou, ao menos, auditados. A reprodutibilidade plena de interações com modelos generativos pode ser limitada por fatores como atualização de modelos, mudanças em APIs, variação probabilística, parâmetros de geração e indisponibilidade de sistemas proprietários. Ainda assim, a impossibilidade de reprodução perfeita não justifica ausência de documentação. Pelo contrário, quanto maior a instabilidade potencial do sistema, mais rigoroso deve ser o registro.

Em pesquisa científica, documentar é tornar visíveis as condições de produção do resultado. Em um estudo bibliométrico tradicional, espera-se informar base de dados, estratégia de busca, data de coleta, critérios de inclusão, procedimentos de limpeza, indicadores e ferramentas. Em um estudo com IA, deve-se acrescentar *prompts*, modelos, versões,

parâmetros, dados enviados, saídas relevantes, revisões humanas e decisões derivadas. O *prompt* é parte do método porque influencia a resposta; portanto, deve ser preservado.

Um protocolo reprodutível de *prompts* deve responder a pelo menos dez perguntas: qual foi o objetivo do *prompt*? Qual modelo foi usado? Em que data? Com quais parâmetros? Quais dados foram fornecidos? O *prompt* permitiu conhecimento externo? Qual saída foi produzida? Quem revisou? Que decisão foi tomada? A saída foi usada no resultado final? Essas perguntas podem ser organizadas em uma planilha de auditoria ou em seção metodológica do relatório.

A documentação deve incluir a versão literal do *prompt*. Resumos não bastam. Pequenas diferenças de redação podem alterar respostas, especialmente quando envolvem restrições, exemplos, formatos e papéis. Se o *prompt* original dizia "não use conhecimento externo" e o relato metodológico apenas afirma "foi solicitada uma análise", perde-se informação crucial. O *prompt* deve ser preservado como texto integral, preferencialmente com identificador único.

Além do *prompt*, deve-se registrar os dados de entrada. Quando os dados forem extensos, pode-se registrar o arquivo, a versão e o recorte utilizado. Por exemplo: "P7 recebeu a planilha keywords_normalizadas_v2.xlsx, aba termos, linhas 1 a 350". Se o modelo recebeu apenas amostra, isso deve ser indicado. Em tarefas de classificação, é importante registrar se todos os registros foram classificados pelo modelo ou se apenas uma amostra foi usada para criar categorias.

A saída também deve ser preservada. Em alguns casos, não é necessário incluir toda saída no corpo do artigo, mas ela deve estar disponível em repositório, apêndice ou material suplementar, quando possível e eticamente adequado. A saída permite verificar se o pesquisador editou, selecionou ou reinterpretou o conteúdo. Em estudos com dados sensíveis, deve-se considerar anonimização ou restrição de acesso, mas o princípio de rastreabilidade permanece.

A documentação deve diferenciar uso substantivo e uso editorial da IA. Uso substantivo ocorre quando a IA influencia categorias, interpretações, escolhas metodológicas ou conclusões. Uso editorial ocorre quando a IA revisa estilo, clareza ou gramática sem alterar conteúdo. Essa distinção é importante porque o nível de documentação exigido difere. Um *prompt* usado para classificar resumos tem maior impacto metodológico do que um *prompt* usado para melhorar fluidez de parágrafo. Contudo, ambos podem ser registrados.

Um sistema simples de documentação pode usar códigos: P-EXP para *prompts* exploratórios, P-ANA para analíticos, P-CRI para críticos e P-RED para redacionais. Cada código recebe número sequencial. Por exemplo, P-EXP-01: geração de perguntas; P-ANA-03: agrupamento de palavras-chave; P-CRI-02: revisão de causalidade; P-RED-01: redação de resultados. Essa codificação permite citar *prompts* no protocolo sem repetir todo o conteúdo.

O protocolo também deve registrar parâmetros. Em muitos sistemas, parâmetros como temperatura, *top-p*, máximo de tokens, *seed*, modo de resposta e ferramentas ativadas influenciam a saída. A temperatura controla grau de aleatoriedade na amostragem. Temperaturas mais baixas tendem a produzir respostas mais previsíveis; temperaturas mais altas aumentam variação. Entretanto, mesmo configurações determinísticas podem não garantir reprodução absoluta em sistemas fechados. Por isso, registrar parâmetros é necessário, mas não suficiente.

A versão do modelo é outro elemento crítico. Modelos são atualizados, substituídos ou ajustados. Uma resposta produzida por um modelo em 2025 pode não ser reproduzida em 2026. Quando o fornecedor disponibiliza identificador de versão, ele deve ser registrado. Quando não disponibiliza, deve-se informar o nome do sistema e a data. Em estudos de longo prazo, mudanças de modelo durante a análise devem ser evitadas ou documentadas.

A documentação deve incluir critérios de validação humana. Não basta registrar que a IA produziu uma classificação; é preciso indicar como a classificação foi revisada. Houve amostragem? Quantos registros foram conferidos? Quem revisou? Houve concordância entre avaliadores? Casos ambíguos foram resolvidos por consenso? O uso de IA não elimina princípios de validação conhecidos em análise de conteúdo, revisão sistemática e curadoria bibliográfica. Pelo contrário, torna-os mais necessários.

Para *prompts* que produzem categorias, recomenda-se registrar versão do esquema classificatório. Categorias podem mudar durante o estudo. Se o pesquisador começa com cinco categorias e depois as agrupa em três, essa mudança precisa ser documentada. A IA pode sugerir categorias, mas o esquema final deve ser definido e justificado pela equipe. Versões de categorias devem ser salvas com data e motivo de alteração.

Em *prompts* que geram redação, o protocolo deve indicar se o texto foi usado integralmente, parcialmente ou apenas como inspiração. Isso evita atribuição indevida de autoria analítica ao modelo. Em publicações científicas, muitas revistas exigem declaração de uso de IA. Mesmo quando não há exigência formal, a transparência recomenda explicar como a ferramenta foi empregada. A declaração deve ser específica: "modelo de linguagem foi usado para apoiar a organização de categorias e revisão de clareza, com validação humana", e não apenas "foi usada IA".

A documentação de *prompts* também se relaciona com ética e privacidade. Dados enviados a modelos podem conter informações sensíveis, mesmo em cientometria. Afiliações, nomes de autores e dados bibliográficos são públicos em muitos casos, mas bases institucionais podem incluir informações internas, avaliações ou dados não publicados. O protocolo deve indicar medidas de proteção: remoção de dados sensíveis, uso de ambientes aprovados, restrição de envio de arquivos proprietários e conformidade com políticas institucionais.

Outro ponto é a documentação de falhas. Muitas vezes, apenas *prompts* bem-sucedidos são preservados, enquanto tentativas fracassadas desaparecem. Contudo, falhas revelam decisões importantes. Se um *prompt* produziu categorias inadequadas e foi substituído, isso mostra refinamento metodológico. Não é necessário publicar todas as tentativas, mas o protocolo interno deve registrar mudanças relevantes. A transparência científica inclui reconhecer que métodos foram ajustados.

Um protocolo reprodutível pode adotar o seguinte modelo de registro: identificador do *prompt*; etapa do estudo; objetivo; texto integral do *prompt*; dados de entrada; modelo e versão; parâmetros; data; saída; avaliação humana; decisão; impacto no produto final; observações. Esse modelo pode parecer extenso, mas torna-se simples quando incorporado à rotina. Em estudos complexos, a ausência desse registro gera retrabalho e dificulta defesa metodológica.

A documentação deve estar conectada aos arquivos do estudo. O ideal é manter uma pasta estruturada: 01_protocolo, 02_busca, 03_dados_brutos, 04_dados_limpos, 05_analises, 06_*prompts*, 07_saidas_ia, 08_relatorio. Dentro da pasta de *prompts*, cada arquivo pode conter *prompt*, data e relação com dados. Essa organização permite que outro membro da equipe acompanhe o fluxo. Em projetos institucionais, essa prática reduz dependência de memória individual.

É recomendável também usar controle de versões para scripts, *prompts* e relatórios. Ferramentas como *Git* podem ser úteis, mas uma planilha versionada já representa avanço. O importante é não sobrescrever decisões sem registro. *Prompts* são documentos metodológicos; devem ser tratados como tal. Em uma cultura científica acostumada a preservar scripts e bases, preservar *prompts* é extensão natural.

A reprodutibilidade em IA generativa deve ser entendida em camadas. A primeira camada é reprodutibilidade documental: outro pesquisador pode ver o que foi feito. A segunda é reprodutibilidade operacional: outro pesquisador pode executar instruções semelhantes com dados semelhantes. A terceira é reprodutibilidade textual: outro pesquisador obtém resposta idêntica. A terceira pode ser difícil; as duas primeiras são obrigação metodológica. O objetivo mínimo é permitir auditoria.

Prompts documentados também melhoram formação de pesquisadores. Eles funcionam como exemplos de raciocínio metodológico. Ao revisar um *prompt* bem escrito, estudantes aprendem a formular tarefas, definir restrições e separar dados de inferências. Bibliotecas de *prompts* podem se tornar materiais pedagógicos, desde que acompanhadas de explicações e alertas. O *prompt* isolado é ferramenta; o *prompt* comentado é instrumento de ensino.

A documentação deve evitar transformar *prompts* em receitas universais. Um *prompt* que funcionou para determinada base, área e objetivo pode falhar em outro contexto. Por isso, protocolos devem registrar condições de uso. Bibliotecas de *prompts* devem incluir notas: "adequado para classificação inicial de palavras-chave", "não recomendado para avaliação de qualidade", "exige revisão por especialista", "usar apenas com dados anonimizados". Essa curadoria previne aplicação indevida.

Em síntese, documentar *prompts* em protocolos reprodutíveis significa reconhecer que a interação com IA tem efeitos metodológicos. O pesquisador não precisa garantir que todas as respostas serão idênticas no futuro, mas precisa tornar seu processo inteligível, rastreável e criticável. A engenharia de *prompts*, quando documentada, fortalece a cientometria assistida por IA. Quando não documentada, converte-se em caixa-preta textual.

5.7 Temperatura, variação, consistência e controle de respostas

Modelos generativos de linguagem produzem respostas por mecanismos probabilísticos. Mesmo quando parecem dialogar de modo estável, suas saídas dependem de arquitetura, treinamento, instruções, contexto, parâmetros de geração e, em muitos casos, variações internas do sistema. Para estudos científicos, compreender essa característica é essencial. A resposta de um modelo não deve ser tratada como medição determinística. Ela é uma geração condicionada. A engenharia de *prompts* precisa lidar com temperatura, variação, consistência e controle.

A temperatura é um parâmetro associado ao grau de aleatoriedade na seleção de tokens. Em termos simples, temperaturas mais baixas favorecem respostas mais previsíveis, próximas das opções de maior probabilidade; temperaturas mais altas permitem maior diversidade e criatividade. Em tarefas científicas, a escolha da temperatura deve acompanhar a finalidade. Para redação criativa ou exploração de alternativas, alguma variação pode ser útil. Para classificação, extração, revisão metodológica e geração de tabelas, recomenda-se reduzir variação (RENZE; GUVEN, 2024).

Entretanto, baixa temperatura não elimina todos os problemas. Primeiro, porque nem todos os sistemas oferecem controle transparente. Segundo, porque modelos podem mudar internamente. Terceiro, porque pequenas alterações no *prompt* ou nos dados podem alterar a saída. Quarto, porque respostas determinísticas podem ser consistentemente erradas. A estabilidade textual não é sinônimo de validade. Por isso, controle de temperatura deve ser combinado com validação de conteúdo.

A variação pode ser produtiva ou perigosa. Em fase exploratória, respostas variadas ajudam a descobrir recortes, categorias ou hipóteses. O pesquisador pode solicitar "gere

três estratégias alternativas de análise" ou "proponha cinco rótulos possíveis para o *cluster*". Nesse caso, diversidade é desejável. Em fase analítica, a variação pode indicar instabilidade. Se o mesmo conjunto de resumos é classificado de maneiras muito diferentes em execuções sucessivas, o esquema classificatório precisa ser revisto ou a tarefa deve ser transferida para método mais controlado.

A consistência refere-se à estabilidade das respostas diante de *prompts* iguais ou semelhantes. Em cientometria, consistência é importante em tarefas como classificação temática, normalização terminológica e identificação de riscos metodológicos. Uma forma simples de avaliar consistência é repetir a tarefa com o mesmo *prompt* e comparar saídas. Outra é usar *prompts* equivalentes com redação diferente. Divergências devem ser analisadas. Se a divergência afeta resultados substantivos, o uso da IA precisa ser limitado ou reforçado por revisão humana.

A técnica de *self-consistency*, proposta em pesquisas sobre raciocínio em modelos de linguagem, explora múltiplas cadeias de resposta e seleciona respostas convergentes. Em estudos cientométricos, a ideia pode ser adaptada de modo prudente: solicitar múltiplas classificações ou interpretações e identificar convergências. Contudo, a convergência entre respostas de um mesmo modelo não deve ser confundida com verdade independente. Modelos podem convergir sobre erro comum. A auto verificação deve ser complementada por evidência externa e expertise humana (WANG *et al.*, 2023).

O controle de respostas começa pelo *prompt*. Instruções claras, dados organizados, restrições explícitas e formatos fixos reduzem variação indesejada. Um formato tabular com colunas definidas tende a produzir respostas mais comparáveis do que texto livre. Categorias previamente estabelecidas reduzem deriva semântica. Exigir que cada classificação seja justificada por evidência textual aumenta auditabilidade. Pedir grau de confiança permite priorizar revisão humana de casos incertos.

Outra estratégia de controle é a padronização de entradas. Se os dados são enviados em formatos diferentes a cada execução, a variação aumenta. Em uma tarefa de classificação de resumos, por exemplo, cada registro deve seguir o mesmo padrão: ID, título, resumo, palavras-chave. A presença de identificadores evita confusão. A ausência de campos irrelevantes reduz distração. Dados limpos e estruturados melhoram o comportamento do modelo.

O controle também pode envolver amostragem e validação cruzada. Em vez de submeter milhares de registros diretamente à IA, o pesquisador pode trabalhar com amostras, avaliar desempenho, ajustar *prompt* e só então aplicar ao conjunto maior. Em tarefas classificatórias, uma amostra pode ser codificada por humanos e usada como referência para avaliar a IA. Métricas de concordância podem ser empregadas quando apropriado. A IA não deve ser adotada em escala antes de ser testada.

Em pesquisas de alto impacto, recomenda-se separar *prompts* de desenvolvimento e *prompts* de aplicação. *Prompts* de desenvolvimento são ajustados em amostras. *Prompts* de aplicação são congelados e aplicados ao *corpus*. Essa separação evita adaptar o *prompt* excessivamente aos casos já conhecidos, fenômeno análogo ao sobreajuste. O protocolo deve registrar quando o *prompt* foi congelado e quais mudanças posteriores foram feitas, se houver.

O controle de respostas também exige distinção entre precisão e utilidade. Uma resposta pode ser útil para organizar ideias, mas imprecisa como evidência. Uma sugestão de categorias pode ser bom ponto de partida, mas não resultado final. Uma redação pode melhorar clareza, mas introduzir nuance inadequada. O pesquisador deve definir previamente que uso será dado à saída. Em alguns casos, basta utilidade heurística. Em outros, exige-se acurácia verificável.

Um problema comum é a falsa confiança gerada por linguagem formal. Modelos de linguagem podem produzir textos assertivos mesmo diante de dados insuficientes. Para controlar esse risco, *prompts* devem exigir linguagem proporcional à evidência. Expressões como "os dados mostram" devem ser reservadas a padrões observáveis. Expressões como "os resultados sugerem" indicam interpretação. Expressões como "uma hipótese possível" indicam especulação. Essa gradação linguística é parte da qualidade científica.

Em *prompts* redacionais, o controle deve impedir que o modelo acrescente referências ou resultados. Uma instrução eficaz é: "não inclua nenhuma informação que não esteja nos dados ou nas referências fornecidas". Ainda assim, a revisão deve verificar. Erros de citação e referências fabricadas são risco documentado em usos de IA generativa. Em capítulos, artigos e relatórios, todas as referências devem ser conferidas manualmente em bases confiáveis.

A variação também pode ser usada como ferramenta crítica. Pedir ao modelo que produza interpretações alternativas permite identificar múltiplas leituras de um padrão. Por exemplo, crescimento de publicações pode indicar expansão real de um campo, mudança terminológica, aumento de indexação ou efeito de política editorial. Um *prompt* pode solicitar que a IA apresente hipóteses concorrentes e evidências necessárias para distingui-las. Essa variação interpretativa é positiva quando claramente marcada como hipótese.

A consistência deve ser avaliada em diferentes níveis. Há consistência interna, quando a resposta não se contradiz. Há consistência com os dados, quando afirmações correspondem às tabelas. Há consistência metodológica, quando conclusões respeitam indicadores. Há consistência externa, quando a interpretação é compatível com literatura relevante. Um *prompt* crítico pode verificar essas camadas, mas a decisão final exige revisão humana.

O controle de respostas em cientometria deve ser proporcional ao risco. Em atividades de ensino, *brainstorming* ou explicação preliminar, controles podem ser mais leves. Em relatórios institucionais, avaliação de programas, políticas científicas ou decisões de financiamento, controles devem ser rigorosos. Quanto maior a consequência, maior a exigência de documentação, validação e revisão independente.

Também é necessário reconhecer limites práticos. Muitos usuários não têm acesso a parâmetros avançados, logs completos ou modelos versionados. Nesses casos, ainda é possível melhorar controle por meio de *prompts* estruturados, repetição de respostas, comparação entre modelos, revisão humana e registro de data e sistema. A ausência de condições ideais não autoriza uso descuidado. O princípio é fazer o melhor controle possível e declarar limitações.

O controle de respostas deve incluir uma política de não uso. Há tarefas que a IA não deve executar como autoridade final: decidir qualidade de pesquisadores, ranquear instituições sem critérios, atribuir mérito causal, identificar má conduta com base apenas em padrões bibliométricos ou produzir diagnósticos estratégicos sem validação. *Prompts* podem auxiliar análise, mas não devem substituir deliberação científica e institucional.

A literatura sobre alucinações e limitações de LLMs reforça que a geração de linguagem pode produzir conteúdo plausível e falso. Em cientometria, isso se manifesta como referências inexistentes, nomes de autores inventados, interpretações exageradas de citações, explicações causais sem evidência e normalizações erradas. O controle de respostas deve, portanto, combinar engenharia de *prompts*, verificação externa e cultura de ceticismo (BENDER *et al.*, 2021; JI *et al.*, 2023; HUANG *et al.*, 2025).

Em síntese, temperatura, variação, consistência e controle compõem a dimensão operacional da engenharia de *prompts*. A temperatura regula diversidade; a variação pode apoiar exploração ou indicar instabilidade; a consistência precisa ser testada; o controle depende de

instruções, formatos, validação e documentação. Em estudos cientométricos, o objetivo não é eliminar toda incerteza, mas torná-la visível, gerenciável e proporcional ao uso da resposta.

5.8 Prompts do capítulo

Os *prompts* apresentados nesta seção são modelos. Eles devem ser adaptados ao tema, *corpus*, base de dados, etapa do estudo, unidade de análise e finalidade do produto. Em todos os casos, recomenda-se registrar a versão utilizada, os dados fornecidos, a saída produzida e a validação humana realizada. Os exemplos foram construídos para induzir respostas estruturadas, distinguir evidência de inferência e preservar rastreabilidade.

Prompt 5.8.1 Modelo geral de prompt cientométrico

Contexto:

Estou conduzindo um estudo cientométrico sobre [TEMA], no período [PERÍODO], com dados provenientes de [BASE/FONTE]. A unidade de análise principal é [ARTIGO/AUTOR/INSTITUIÇÃO/PAÍS/PALAVRA-CHAVE/PERIÓDICO/REFERÊNCIA]. O objetivo do estudo é [OBJETIVO].

Papel:

Atue como analista cientométrico e revisor metodológico. Sua função é apoiar a organização e a interpretação dos dados, sem substituir a validação humana.

Tarefa: Com base nos dados fornecidos, produza uma análise estruturada que identifique padrões observáveis, possíveis interpretações, limitações e verificações necessárias.

Dados:

Use exclusivamente os dados apresentados a seguir: [INSERIR TABELA, LISTA, RESUMOS, *Clusters* OU RESULTADOS].

Restrições:

Não invente autores, artigos, indicadores, valores, referências ou eventos externos. Não afirme causalidade quando houver apenas associação. Diferencie dado observado, inferência e hipótese. Quando a evidência for insuficiente, declare explicitamente a limitação.

Formato de saída:

Apresente a resposta em tabela com as colunas: achado observado; evidência nos dados; interpretação possível; grau de confiança; limitação; verificação recomendada. Ao final, escreva um parágrafo síntese em tom formal científico.

Esse modelo geral pode ser usado em várias etapas do estudo. Sua principal qualidade é transformar uma solicitação ampla em matriz de interpretação. Ele não pede apenas um texto, mas uma resposta auditável. O pesquisador pode comparar cada interpretação com a evidência indicada e revisar limitações antes de redigir resultados. A coluna de grau de confiança não deve ser entendida como probabilidade estatística; trata-se de indicação heurística que ajuda a priorizar revisão humana.

Prompt 5.8.2 Prompt de revisão crítica

Contexto:

Abaixo está uma interpretação preliminar de resultados de um estudo cientométrico. Preciso verificar se o texto respeita os dados, os indicadores e os limites metodológicos.

Papel:

Atue como revisor crítico de metodologia cientométrica.

Tarefa:

Identifique problemas de interpretação, extrapolação ou redação. Classifique cada problema e proponha correção.

Texto a revisar:

[INSERIR TEXTO].

Dados ou resultados que sustentam o texto:

[INSERIR TABELA, GRÁFICO DESCRITO OU RESULTADOS].

Critérios de revisão:

Procure especialmente: causalidade indevida; generalização excessiva; confusão entre quantidade e qualidade; uso inadequado de citações como medida absoluta de impacto; ausência de normalização por área ou tempo; conclusão sem evidência; linguagem avaliativa não sustentada; referência a dados não apresentados.

Formato de saída:

Apresente uma tabela com as colunas: trecho problemático; tipo de problema; explicação; risco para o estudo; reformulação recomendada. Ao final, liste três recomendações prioritárias para fortalecer a interpretação.

Esse *prompt* é adequado antes da submissão de relatórios, artigos ou capítulos. Sua função é reduzir o excesso interpretativo. Ele também pode ser usado para revisar respostas produzidas por outro *prompt*. Nesse caso, a IA atua como instrumento de crítica auxiliar. A decisão final deve ser humana, mas a tabela de problemas facilita a revisão.

Prompt 5.8.3 Prompt para gerar tabela analítica

Contexto:

Tenho resultados de um estudo cientométrico e preciso transformá-los em uma tabela analítica para apoiar a redação da seção de resultados.

Papel:

Atue como organizador de resultados cientométricos.

Tarefa:

Converta os resultados abaixo em tabela interpretativa, preservando fidelidade aos dados.

Resultados: [INSERIR RESULTADOS].

Restrições:

Não acrescente valores, autores ou tendências que não estejam nos resultados. Não produza avaliação de qualidade. Use linguagem descritiva e cautelosa. Se algum resultado for ambíguo, marque como "exige revisão".

Formato de saída:

Gere tabela com as colunas: dimensão analisada; resultado observado; indicador ou evidência; interpretação descritiva; cautela metodológica; possível visualização; uso na redação.

Esse *prompt* é particularmente útil na transição entre análise e escrita. Muitos estudos cientométricos acumulam tabelas, gráficos e mapas, mas enfrentam dificuldade para convertê-los em narrativa coerente. A tabela analítica funciona como ponte: organiza resultados sem ainda produzir discussão final. Ela também permite que diferentes membros da equipe revisem interpretações antes da redação.

Prompt 5.8.4 Prompt para transformar saída textual em protocolo

Contexto:

Abaixo está uma resposta textual gerada durante o planejamento de um estudo cientométrico. Preciso convertê-la em um protocolo metodológico estruturado.

Papel:

Atue como metodologista de pesquisa cientométrica.

Tarefa:

Transforme a resposta textual em protocolo, separando decisões já definidas, decisões pendentes, riscos e próximos passos.

Texto de origem: [INSERIR TEXTO].

Restrições:

Não acrescente decisões que não estejam no texto. Quando algo estiver implícito ou incompleto, marque como pendente. Não trate sugestões como decisões finais. Diferencie o que é objetivo, critério, procedimento, indicador, ferramenta e limitação.

Formato de saída:

Apresente as seções: título provisório do estudo; objetivo; perguntas de pesquisa; escopo; bases de dados; estratégia de busca; critérios de inclusão e exclusão; metadados necessários; indicadores; análises previstas; *prompts* a utilizar; validação humana; riscos metodológicos; decisões pendentes; próximos passos.

Esse *prompt* é útil porque muitas interações com IA produzem texto rico, porém pouco estruturado. Transformar a saída em protocolo ajuda a recuperar decisões e lacunas. O protocolo resultante não deve ser aceito automaticamente; ele deve ser revisado e aprovado pela equipe. Sua função é apoiar formalização.

Prompt 5.8.5 Prompt para pedir nível explícito de incerteza

Contexto:

Estou analisando resultados cientométricos e preciso evitar conclusões excessivas.

Tarefa:

Interprete os resultados abaixo, classificando cada afirmação segundo nível de incerteza.

Dados: [INSERIR RESULTADOS].

Formato:

Para cada afirmação, indique se é: dado observado; inferência fortemente sustentada; inferência moderada; hipótese especulativa; ou afirmação não sustentada. Explique a classificação e indique que evidência adicional seria necessária.

Esse *prompt* permite graduar a força das afirmações. Ele é útil em discussões e conclusões, onde frequentemente se passa de padrão observado a explicação causal. A classificação ajuda a manter prudência.

Prompt 5.8.6 Prompt para distinguir dado, inferência e hipótese

Contexto:

Abaixo há um texto interpretativo sobre um mapa científico. Preciso separar os níveis de afirmação.

Tarefa:

Reescreva o texto em três blocos: dados observados; inferências sustentadas pelos dados; hipóteses que exigem investigação adicional.

Restrições:

Não mova uma hipótese para o bloco de dados. Não acrescente novos achados. Use linguagem formal científica.

Esse *prompt* pode ser usado como ferramenta de depuração epistemológica. Ele força a separar aquilo que o mapa mostra daquilo que o pesquisador interpreta. Em cientometria, essa distinção é central, pois mapas e indicadores frequentemente sugerem padrões que exigem contextualização.

Prompt 5.8.7 Prompt para auditoria de afirmações

Contexto:

Preciso auditar afirmações de um relatório cientométrico antes da entrega.

Tarefa:

Para cada afirmação do texto, verifique se ela está sustentada pelos dados fornecidos.

Entrada:

Texto do relatório: [INSERIR TEXTO]. Dados disponíveis: [INSERIR DADOS].

Formato:

Tabela com as colunas: afirmação; evidência correspondente; status (sustentada, parcialmente sustentada, não sustentada, exige referência externa); problema identificado; recomendação.

Esse *prompt* deve ser usado como etapa final de controle. Ele não substitui revisão humana, mas ajuda a identificar afirmações que escaparam do *corpus*. É especialmente útil quando textos passam por várias rodadas de edição e podem acumular generalizações.

5.9 Discussão integrada: engenharia de prompts como competência cientométrica

A engenharia de *prompts* aplicada à cientometria não deve ser reduzida a técnica de obtenção de respostas melhores. Ela deve ser entendida como competência metodológica. Essa competência envolve reconhecer a natureza dos dados científicos, compreender os limites dos indicadores, formular perguntas precisas, desenhar instruções verificáveis, documentar interações e avaliar criticamente as saídas. Em outras palavras, a qualidade do *prompt* depende da qualidade do raciocínio cientométrico que o antecede.

Figura 5.2 - Raciocínio cientométrico

Fonte: Imagem gerada por IA, ChatGPT (2026).

A centralidade dos *prompts* decorre da mediação textual introduzida por modelos de linguagem. Antes da IA generativa, a maior parte das operações cientométricas era mediada por interfaces de busca, planilhas, scripts e softwares de visualização. Essas mediações continuam existindo. O que muda é que agora há uma camada conversacional capaz de

interpretar, reorganizar e redigir. Essa camada pode acelerar o trabalho, mas também pode ocultar decisões. A engenharia de *prompts* torna essa camada mais explícita.

Um dos maiores desafios é impedir que a IA transforme incerteza em narrativa. A cientometria lida com fenômenos complexos: produção científica, circulação de ideias, influência, colaboração, institucionalização de temas e disputas simbólicas. Indicadores são aproximações. Redes são representações. Mapas são modelos. Quando a IA interpreta esses objetos, ela pode produzir uma narrativa demasiadamente coerente, como se todo padrão tivesse explicação clara. A engenharia de *prompts* deve inserir fricção: pedir limitações, alternativas e evidências.

Outro desafio é preservar a diferença entre análise e avaliação. Muitas demandas institucionais querem saber "quem produz mais", "qual área é mais relevante" ou "qual grupo é melhor". A cientometria pode oferecer indicadores para subsidiar avaliação, mas não deve ser confundida com julgamento absoluto de qualidade. *Prompts* devem refletir essa cautela. Em vez de pedir que a IA ranqueie pesquisadores, pode-se pedir que descreva padrões de produção e impacto, indicando limitações e necessidade de avaliação qualitativa (HICKS *et al.*, 2015; WALTMAN, 2016).

A engenharia de *prompts* também precisa considerar desigualdades de cobertura. Bases internacionais tendem a privilegiar periódicos em inglês e áreas com forte cultura de publicação indexada. Estudos sobre ciências humanas, contextos regionais, produção técnica, livros, capítulos e literatura cinzenta exigem cuidado. Um *prompt* que interpreta ausência de registros como ausência de produção comete erro grave. Por isso, restrições devem incluir advertências sobre cobertura da base e natureza do *corpus*.

Prompts podem ajudar a tornar essas limitações visíveis. Um *prompt* crítico pode perguntar: "que tipos de produção científica podem estar sub-representados nesta base?" Outro pode solicitar: "liste interpretações alternativas para baixa contagem de citações". Esse uso crítico da IA é valioso porque expande a reflexão metodológica. Contudo, as respostas ainda precisam ser confrontadas com conhecimento de domínio.

Em termos pedagógicos, a engenharia de *prompts* oferece uma oportunidade de ensinar cientometria. Ao construir um *prompt*, o estudante precisa explicitar objetivo, dados, indicadores e limites. Essa explicitação revela lacunas de compreensão. Se alguém não consegue formular um *prompt* claro para interpretar uma rede de cocitação, talvez ainda não compreenda cocitação suficientemente. Assim, o *prompt* funciona como exercício de método.

A engenharia de *prompts* também aproxima cientometria e ciência aberta. Ao documentar *prompts*, dados, versões e saídas, o pesquisador contribui para transparência. Mesmo quando os modelos não são totalmente reprodutíveis, o registro permite auditoria. Em estudos futuros, repositórios poderão incluir não apenas scripts e dados, mas também *prompts* e logs de interação. Essa prática pode tornar-se componente normal de materiais suplementares (PAGE *et al.*, 2021).

Há, porém, riscos de burocratização. Documentar tudo indiscriminadamente pode tornar o processo pesado. É necessário distinguir *prompts* triviais de *prompts* metodologicamente relevantes. Um *prompt* usado para corrigir vírgulas não tem o mesmo peso de um *prompt* usado para classificar mil resumos. A documentação deve ser proporcional ao impacto da saída no estudo. O princípio é registrar tudo que influencie decisões analíticas ou conclusões.

A qualidade da engenharia de *prompts* também depende de infraestrutura. Sistemas que permitem exportar conversas, fixar versões de modelos, controlar parâmetros e proteger dados favorecem uso científico. Sistemas opacos dificultam auditoria. Pesquisadores e

instituições devem considerar esses fatores ao adotar ferramentas. A escolha da plataforma faz parte do método quando a IA participa da análise.

Um ponto crítico é a dependência de fornecedores. Modelos proprietários podem mudar sem aviso, alterar comportamento e limitar acesso a logs. Modelos abertos oferecem maior controle, mas exigem infraestrutura e competência técnica. A decisão entre sistemas deve considerar privacidade, custo, desempenho, reprodutibilidade e governança. *Prompts* bem documentados reduzem, mas não eliminam, essa dependência.

A engenharia de *prompts* não deve substituir formação estatística, bibliométrica ou sociológica. Ao contrário, exige essas formações. Um pesquisador que não compreende normalização de citações não saberá pedir interpretação adequada. Um usuário que desconhece viés de cobertura não saberá impor restrições. Um gestor que confunde produtividade com qualidade pode formular *prompts* avaliativos inadequados. A IA amplia capacidades de quem sabe perguntar; não corrige automaticamente perguntas mal formuladas.

Essa constatação tem implicações para equipes. Projetos cientométricos assistidos por IA devem reunir competências complementares: conhecimento do domínio, bibliometria, ciência de dados, curadoria informacional, ética e comunicação. A engenharia de *prompts* pode ser compartilhada, revisada e aprimorada coletivamente. Bibliotecários e cientistas da informação têm papel importante, pois dominam estratégias de busca, metadados e organização do conhecimento.

Figura 5.3 - Competência cientométrica

Fonte: Imagem gerada por IA, ChatGPT (2026).

Prompts também podem ser usados para negociação interdisciplinar. Em estudos que envolvem áreas diferentes, a IA pode ajudar a traduzir vocabulários, comparar conceitos e explicitar ambiguidades. Por exemplo, o termo “modelo” pode significar modelo estatístico, modelo teórico, modelo computacional ou organismo modelo. Um *prompt* de curadoria

conceitual pode identificar usos distintos e evitar agrupamentos indevidos. Esse apoio é valioso em campos interdisciplinares.

Na redação de relatórios, a engenharia de *prompts* contribui para adequação de linguagem. Um relatório para gestores precisa ser claro e acionável; um artigo científico exige detalhe metodológico; uma nota técnica requer síntese. *Prompts* redacionais podem adaptar o mesmo resultado a públicos diferentes. Contudo, a adaptação não deve alterar o conteúdo. O *prompt* deve dizer: "preserve todos os limites metodológicos". A clareza não pode ser obtida ao preço de simplificação enganosa.

A engenharia de *prompts* também pode apoiar revisão por pares. Autores podem usar *prompts* críticos para identificar fragilidades antes da submissão. Revisores podem, se permitido pelas normas editoriais e respeitando confidencialidade, usar IA para organizar comentários, nunca para substituir julgamento. Editores podem exigir declaração de uso. Em todos os casos, transparência é indispensável.

No campo da avaliação científica, a cautela deve ser redobrada. Métricas influenciam carreiras, reputações e financiamento. A IA pode tornar relatórios mais rápidos, mas também pode naturalizar interpretações questionáveis. *Prompts* devem evitar rankings simplistas e linguagem meritocrática sem contexto. Devem solicitar normalização, comparação adequada, limites e evidências complementares. A engenharia de *prompts* torna-se, aqui, instrumento de responsabilidade.

Um aspecto promissor é a criação de bibliotecas de *prompts* validadas por comunidades. Associações científicas, grupos de pesquisa e bibliotecas universitárias podem desenvolver modelos para tarefas comuns, acompanhados de notas metodológicas. Esses repositórios poderiam indicar finalidade, dados necessários, limitações, exemplos e formas de validação. A padronização relativa pode melhorar qualidade e reduzir improvisação.

Ainda assim, a engenharia de *prompts* deve permanecer sensível ao contexto. A cientometria é altamente dependente de área, idioma, base, período e finalidade. Um *prompt* para literatura biomédica indexada no PubMed não serve sem adaptação para estudos em educação publicados em periódicos regionais. Um *prompt* para artigos não serve para livros. Um *prompt* para países não serve para instituições sem tratar afiliações. A adaptação é parte do método.

Em síntese, a engenharia de *prompts* torna explícita a pergunta, organiza a interação com a IA, define limites, favorece auditoria e apoia comunicação. Sua adoção responsável pode melhorar eficiência e qualidade da cientometria, desde que permaneça subordinada a dados, métodos e revisão humana. O futuro da cientometria com IA não dependerá apenas de modelos mais poderosos, mas de pesquisadores capazes de formular melhores perguntas, documentar melhores processos e interpretar com maior prudência.

5.10 Conclusão

Este capítulo apresentou a engenharia de *prompts* como componente metodológico da cientometria assistida por inteligência artificial. A análise partiu da ideia de que *prompts* não são comandos ocasionais, mas instrumentos de pesquisa que orientam a relação entre dados, modelos e interpretação. Em estudos cientométricos, essa orientação é decisiva porque os resultados envolvem indicadores, redes, mapas, classificações e narrativas que podem influenciar decisões acadêmicas e institucionais.

Foram discutidos os elementos constitutivos de um bom *prompt* científico: objetivo, contexto, papel, tarefa, dados, restrições, formato de saída, critérios de qualidade, rastreabilidade

e declaração de incerteza. Esses elementos reduzem ambiguidade e tornam as respostas mais verificáveis. Também foi proposta uma estrutura mínima baseada em contexto, papel, tarefa, dados, restrições e formato de saída. Essa estrutura pode ser aplicada em diferentes etapas do estudo, da formulação da pergunta à redação do relatório.

A tipologia de *prompts* exploratórios, analíticos, críticos e redacionais mostrou que cada interação com IA possui função distinta. *Prompts* exploratórios ampliam possibilidades; *prompts* analíticos organizam dados; *prompts* críticos identificam riscos; *prompts* redacionais comunicam resultados. O uso responsável da IA depende de reconhecer essas diferenças e evitar que respostas exploratórias sejam tratadas como conclusões.

As cadeias de *prompts* foram apresentadas como arquitetura operacional para conduzir estudos complexos. Ao decompor o fluxo da pergunta ao relatório, elas permitem inserir pontos de validação, documentar decisões e reduzir opacidade. Em vez de pedir à IA um estudo completo, o pesquisador constrói um processo sequencial, no qual cada etapa é revisada antes de alimentar a próxima. Essa lógica é mais compatível com o método científico.

A documentação de *prompts* em protocolos reprodutíveis foi defendida como exigência de transparência. Mesmo quando a reprodução textual perfeita é limitada, é possível registrar *prompts*, modelos, parâmetros, dados de entrada, saídas e decisões humanas. Esse registro torna o uso de IA auditável e permite avaliar o peso da ferramenta nos resultados. *Prompts* que influenciam categorias, interpretações ou conclusões devem ser tratados como artefatos metodológicos.

Por fim, discutiram-se temperatura, variação, consistência e controle de respostas. Modelos generativos produzem linguagem sob condições probabilísticas; por isso, a estabilidade das respostas precisa ser gerenciada. Temperaturas mais baixas, formatos estruturados, repetição controlada, comparação de saídas, validação humana e registro de parâmetros são estratégias úteis. Contudo, nenhuma delas substitui a revisão crítica. Consistência não garante verdade, e fluência não garante validade.

A mensagem central é que a IA pode fortalecer a cientometria quando usada como assistente metodológico, mas pode fragilizá-la quando tratada como autoridade epistêmica. A engenharia de *prompts* é a prática que permite orientar essa assistência. Ela transforma a interação com modelos de linguagem em procedimento explícito, controlável e revisável. Em uma cientometria comprometida com rigor, transparência e responsabilidade, saber perguntar torna-se tão importante quanto saber medir.

Síntese operacional para equipes de pesquisa

A aplicação da engenharia de *prompts* em estudos cientométricos deve ser convertida em rotina de equipe. Em projetos individuais, é comum que o pesquisador formule *prompts* à medida que encontra dificuldades. Em projetos institucionais, essa improvisação tende a produzir inconsistência, sobretudo quando diferentes pessoas analisam partes do *corpus*, revisam indicadores ou redigem seções do relatório. Uma rotina operacional mínima deve estabelecer quem pode criar *prompts*, quem os revisa, onde eles são armazenados, como são versionados e quais saídas podem influenciar decisões analíticas.

Um fluxo recomendável começa com a definição de papéis humanos. O responsável metodológico valida se o *prompt* está coerente com a pergunta de pesquisa e com os indicadores. O especialista de domínio verifica se termos, categorias e interpretações fazem sentido no campo estudado. O curador de dados confere se a entrada está limpa, anonimizada quando necessário e compatível com a tarefa. O redator científico adapta a saída para o produto

final sem acrescentar inferências indevidas. Essa divisão evita que a IA seja usada como atalho para decisões que exigem competência humana.

Também é útil definir níveis de criticidade das tarefas. Tarefas de baixa criticidade incluem reformular uma explicação didática, sugerir títulos provisórios ou melhorar a coesão de um parágrafo. Tarefas de média criticidade incluem agrupar palavras-chave, sugerir categorias e resumir conjuntos de resumos. Tarefas de alta criticidade incluem interpretar resultados para avaliação institucional, justificar políticas de pesquisa, comparar unidades acadêmicas ou recomendar prioridades estratégicas. Quanto maior a criticidade, maior deve ser a exigência de documentação, dupla revisão e validação independente.

A equipe deve manter uma biblioteca de *prompts* aprovados. Essa biblioteca não deve conter apenas o texto do *prompt*, mas também sua finalidade, o tipo de dado aceito, exemplos de uso, limitações, critérios de validação e histórico de alterações. Um *prompt* para classificar resumos, por exemplo, deve indicar se foi testado com determinada área, qual taxa de revisão humana foi adotada e que tipos de erro foram observados. *Prompts* não testados podem ser usados em exploração, mas não deveriam sustentar conclusões sem validação.

Outro procedimento operacional é a separação entre espaço de experimentação e espaço de produção. No espaço de experimentação, a equipe testa *prompts*, compara respostas e ajusta instruções. No espaço de produção, utiliza versões congeladas, registradas e aplicadas de modo consistente. Essa separação impede que o *prompt* seja modificado durante a análise sem registro. Em estudos com grande *corpus*, mudanças no meio do processo podem criar classificações incompatíveis entre subconjuntos.

Figura 5.4 - Fluxo operacional de pesquisa

Fonte: Imagem gerada por IA, ChatGPT (2026).

A governança de *prompts* deve incluir critérios de descarte. Um *prompt* deve ser abandonado quando produz respostas instáveis, não segue restrições, inventa informações, classifica muitos casos sem justificativa ou exige mais correção do que benefício. A decisão de descartar também deve ser registrada, pois mostra que a equipe não aceitou a IA de modo acrítico. Em alguns casos, a melhor decisão metodológica é não usar IA em determinada etapa.

Para equipes que trabalham com dados institucionais, políticas de segurança são indispensáveis. Antes de enviar dados a qualquer sistema, deve-se avaliar termos de uso,

confidencialidade, armazenamento, possibilidade de treinamento com os dados e exigências legais. Quando houver dúvida, recomenda-se trabalhar com dados agregados, anonimizados ou sintéticos. A engenharia de *prompts* não é apenas técnica textual; ela participa da governança informacional do estudo.

Apêndice A.5 - Matriz sintética para desenho de prompts cientométricos

Quadro 5.3 - Matriz sintética para desenho de *prompts* cientométricos

Componente	Pergunta orientadora	Risco se omitido	Exemplo de controle
Contexto	Qual é o tema, a etapa e o *corpus*?	Resposta genérica ou fora do escopo.	Indicar base, período, área e finalidade.
Papel	Que função o modelo deve desempenhar?	Comentário amplo sem função analítica.	Atuar como revisor metodológico ou curador terminológico.
Tarefa	O que deve ser entregue?	Saída vaga, incompleta ou impossível de auditar.	Usar verbos como classificar, comparar, revisar, extrair.
Dados	Quais evidências podem ser usadas?	Invenção de fatos ou uso indevido de conhecimento externo.	Usar somente tabela, resumos ou *clusters* fornecidos.
Restrições	Quais limites devem ser respeitados?	Causalidade indevida, generalização e falsa autoridade.	Não afirmar causalidade; declarar incerteza.
Formato	Como a resposta será revisada?	Texto bonito, mas difícil de verificar.	Tabela com evidência, interpretação, limitação e validação.

Fonte: Elaborado pelo autor (2026).

A síntese operacional pode ser formulada como princípio: nenhum *prompt* metodologicamente relevante deve ser usado sem objetivo, versão, entrada, saída, validação e responsável. Esse princípio não impede criatividade, mas estabelece responsabilidade. A cientometria com IA deve ser capaz de explicar não apenas o que concluiu, mas também como conversou com os sistemas que ajudaram a produzir suas análises.

Capítulo 6

Formulação da pergunta cientométrica

6.1 Introdução

A formulação da pergunta cientométrica é o ponto em que uma curiosidade intelectual, uma demanda institucional ou uma necessidade de gestão se converte em desenho analítico. Antes da coleta de dados, antes do cálculo de indicadores e antes da construção de mapas, existe uma decisão fundamental: que aspecto da ciência se pretende compreender? Essa decisão define a unidade de observação, a escala de análise, o tipo de evidência aceitável, o modo de interpretação e o alcance das conclusões. Quando a pergunta é imprecisa, os resultados podem ser tecnicamente corretos e, ainda assim, pouco informativos; quando a pergunta é bem formulada, mesmo análises simples podem produzir diagnósticos robustos, úteis e transparentes.

Na história da cientometria, a pergunta de pesquisa sempre esteve associada à possibilidade de observar a ciência por meio de seus registros formais. Os trabalhos de Garfield sobre índices de citação, de Price sobre redes de artigos científicos, de Kessler sobre acoplamento bibliográfico e de Small sobre cocitação mostram que diferentes perguntas exigem diferentes representações da literatura. Perguntar quem publica mais não equivale a perguntar quem influencia intelectualmente uma área; perguntar quais países colaboram não equivale a perguntar quais conceitos estruturam uma frente de pesquisa; perguntar quais temas crescem não equivale a afirmar que esses temas são socialmente relevantes ou cientificamente superiores (GARFIELD, 1955; PRICE, 1965; KESSLER, 1963; SMALL, 1973).

No contexto contemporâneo, a cientometria se tornou uma prática transversal. É utilizada por pesquisadores que desejam situar uma área, por bibliotecários que organizam acervos e serviços de apoio à pesquisa, por gestores que monitoram desempenho institucional, por agências que avaliam políticas científicas e por grupos de pesquisa que buscam oportunidades de colaboração. Essa expansão aumenta a importância da pergunta, pois cada público tende a solicitar respostas diferentes a partir de bases semelhantes. A mesma coleção de artigos pode alimentar uma análise de produtividade, um mapa temático, um diagnóstico de colaboração internacional, uma prospecção de temas emergentes ou um relatório de avaliação institucional.

A inteligência artificial generativa modifica esse cenário de modo relevante. Modelos de linguagem podem ajudar a decompor temas amplos, sugerir dimensões analíticas, comparar perguntas alternativas, identificar unidades de análise e construir protocolos preliminares. Entretanto, a IA não substitui a definição epistemológica e metodológica do estudo. Ela opera sobre instruções, exemplos, dados e restrições fornecidos pelo pesquisador. Se o objetivo é vago, se o escopo é ambíguo ou se a finalidade avaliativa não é explicitada, a resposta gerada tende a mascarar a fragilidade do problema inicial com linguagem fluente. Por isso, a formulação da pergunta é o primeiro campo de aplicação crítica dos *prompts* em cientometria (BROWN *et al.*, 2020; LIU *et al.*, 2023).

Este capítulo assume que a pergunta cientométrica deve ser formulada como uma relação entre objeto, evidência e uso. O objeto indica o fenômeno a ser observado, como produção, impacto, colaboração, circulação, visibilidade ou emergência temática. A evidência define quais registros serão mobilizados, como publicações, citações, referências, palavras-chave, resumos, afiliações, patentes, menções em mídias sociais ou documentos de política. O uso explicita a finalidade do estudo, que pode ser descritiva, comparativa, avaliativa, prospectiva,

estratégica ou reflexiva. A ausência de qualquer um desses elementos compromete a interpretabilidade dos resultados.

O capítulo também sustenta que formular perguntas não é apenas escolher frases interrogativas. Trata-se de construir uma arquitetura de decisões metodológicas. A pergunta principal deve ser desdobrada em perguntas secundárias; cada pergunta secundária deve corresponder a uma unidade de análise; cada unidade deve estar associada a indicadores e técnicas; cada indicador deve ser interpretado à luz de uma limitação conhecida; e cada conclusão deve retornar ao escopo inicial. Em estudos assistidos por IA, essa arquitetura pode ser apoiada por cadeias de *prompts*, desde que os *prompts* sejam documentados e avaliados.

Figura 6.1 - Arquitetura de decisões metodológicas

Fonte: Imagem gerada por IA, ChatGPT (2026).

Um estudo cientométrico pode falhar por excesso de ambição. Perguntas como 'qual é o estado da arte da inteligência artificial na saúde?' são amplas demais para orientar diretamente uma análise. Elas podem ser transformadas em perguntas mais operacionais: como evoluiu a produção sobre IA em diagnóstico por imagem entre 2014 e 2025? Quais países e instituições concentram publicações nessa área? Quais documentos são mais cocitados? Quais termos emergem nos resumos dos últimos cinco anos? Cada uma dessas perguntas exige técnicas distintas, como análise temporal, colaboração institucional, cocitação e coocorrência de termos.

Também pode falhar por estreitamento indevido. Uma pergunta como 'quantos artigos sobre determinado tema foram publicados?' pode ser necessária, mas raramente é suficiente. A contagem de publicações descreve atividade, não explica relevância, qualidade, influência ou diversidade. O pesquisador deve perguntar o que a contagem significa no contexto do problema. Um aumento de publicações pode indicar emergência científica, efeito de políticas de financiamento, mudança terminológica, expansão de bases indexadoras ou simplesmente popularização de uma palavra-chave. A IA pode auxiliar na enumeração dessas hipóteses, mas a validação depende de dados e conhecimento de domínio.

A formulação da pergunta também demanda responsabilidade. O Manifesto de Leiden, ao propor princípios para o uso responsável de métricas de pesquisa, ressalta a necessidade de considerar contexto, missão institucional, diferenças disciplinares e julgamento qualitativo. Em consequência, perguntas avaliativas devem ser formuladas de modo a evitar reducionismos.

Em vez de perguntar 'qual instituição é melhor?', uma formulação cientometricamente mais responsável perguntaria: 'como se distribuem a produção, a colaboração e o impacto normalizado de instituições com missões semelhantes, em um campo e período definidos, considerando limitações de cobertura da base de dados?' (HICKS *et al.*, 2015).

Em síntese, este capítulo defende que a pergunta é o dispositivo que torna a cientometria inteligível. Ela define o que será observado, o que permanecerá fora do estudo, que tipo de comparação será legítima, que métricas serão aceitáveis e que conclusões poderão ser sustentadas. A IA generativa amplia a capacidade de explorar perguntas, mas aumenta a necessidade de explicitar pressupostos. O *prompt* central deste capítulo é, portanto, um *prompt* de desenho de pesquisa: ele não pede respostas finais, mas alternativas de formulação, critérios de escolha e advertências metodológicas.

A literatura da área reforça essa orientação. Zupic e Cater destacam que métodos bibliométricos são particularmente úteis quando articulam desenho de pesquisa, coleta de dados, análise e interpretação, e não quando se reduzem à aplicação mecânica de softwares. Aria e Cuccurullo propõem um fluxo de mapeamento científico que inclui coleta, normalização, análise, visualização e interpretação, o que pressupõe uma pergunta capaz de organizar essas etapas. Haghani, ao discutir o que torna uma análise cientométrica publicável e informativa, enfatiza a necessidade de relevância substantiva, desenho metodológico claro e interpretação que não confunda descrição com explicação. Esses argumentos convergem para uma conclusão: a pergunta é o mecanismo de controle de qualidade do estudo (ZUPIC; CATER, 2015; ARIA; CUCCURULLO, 2017; HAGHANI, 2023).

6.2 Perguntas sobre produção científica

Figura 6.2 - O que podemos investigar

Fonte: Imagem gerada por IA, ChatGPT (2026).

Perguntas sobre produção científica são geralmente as primeiras formuladas em um estudo cientométrico. Elas investigam quanto, quando, onde e por quem a ciência é produzida. Em sua forma mais simples, referem-se ao número de publicações em determinado campo, período ou conjunto institucional. Em sua forma mais elaborada, examinam padrões de crescimento, distribuição disciplinar, concentração de autoria, especialização temática, produtividade institucional e variação geográfica. A aparente simplicidade desse tipo de pergunta não deve ocultar sua complexidade metodológica. Contar publicações é uma operação dependente de base de dados, estratégia de busca, tipo documental, idioma, critérios de inclusão e normalização de nomes.

Uma pergunta de produção científica pode ser descritiva: 'como evoluiu o número de artigos sobre vacinas de RNA mensageiro entre 2010 e 2025?'. Pode ser comparativa: 'quais países apresentaram maior crescimento relativo na produção sobre energia solar fotovoltaica no período 2015-2025?'. Pode ser institucional: 'como se distribui a produção científica de uma universidade por áreas do conhecimento?'. Pode ser temática: 'quais subtemas concentram publicações em agricultura de precisão?'. Em todos os casos, o verbo analítico importa. Evoluir, comparar, distribuir, concentrar e diversificar implicam operações distintas.

Na tradição bibliométrica, estudos de produtividade têm raízes em Lotka, que investigou a distribuição da produção científica entre autores, e em Bradford, que discutiu a dispersão da literatura em periódicos. Tais contribuições mostram que a produção científica não se distribui de modo uniforme. Poucos autores, periódicos, instituições ou países podem concentrar parcelas significativas da produção em determinados campos. A formulação da pergunta deve, portanto, antecipar se o estudo pretende apenas medir volume ou compreender padrões de concentração, dispersão e desigualdade (LOTKA, 1926; BRADFORD, 1934).

A unidade de análise é decisiva. Uma pergunta sobre produção de artigos pode ter como unidade o documento; uma pergunta sobre produtividade de autores tem como unidade o pesquisador; uma pergunta sobre produção institucional exige afiliações; uma pergunta sobre produção nacional requer endereços e regras de contagem por país; uma pergunta sobre produção temática depende de classificação por termos, categorias ou modelos de tópicos. Quando a unidade de análise não é explicitada, a interpretação se torna instável. Por exemplo, um artigo internacional com dez instituições pode ser contado integralmente para cada instituição, fracionado entre elas ou atribuído apenas à instituição correspondente, produzindo resultados distintos.

A IA pode contribuir na fase de formulação ao transformar um tema genérico em perguntas de produção científica mais precisas. Dado o tema 'sustentabilidade urbana', um modelo de linguagem pode sugerir perguntas sobre crescimento anual de publicações, distribuição por áreas, principais periódicos, países mais produtivos, instituições com maior presença e mudanças terminológicas. Entretanto, essas sugestões não devem ser aceitas automaticamente. O pesquisador deve avaliar se cada pergunta é respondível com os dados disponíveis, se a estratégia de busca é capaz de recuperar a literatura pertinente e se a unidade de análise é adequada ao fenômeno observado.

Um erro comum é confundir produção com maturidade científica. Uma área muito produtiva pode ser madura, mas também pode estar em fase de explosão especulativa; uma área pouco produtiva pode ser emergente, periférica, altamente especializada ou subindexada. A pergunta de produção, isoladamente, não permite concluir sobre qualidade, impacto ou relevância social. Ela pode, contudo, indicar dimensões que exigem investigação adicional. Por isso, perguntas sobre produção devem ser formuladas como portas de entrada para análises mais ricas, e não como respostas finais sobre o valor de uma área.

Outro problema é a sensibilidade temporal. A produção científica cresce de forma desigual entre áreas e períodos. Um recorte temporal muito curto pode supervalorizar flutuações episódicas; um recorte muito longo pode misturar fases históricas incompatíveis. Perguntas sobre produção devem explicitar a razão do período escolhido. O intervalo pode ser justificado por eventos científicos, mudanças regulatórias, criação de programas de financiamento, surgimento de tecnologias, crises sanitárias ou disponibilidade de dados. Sem essa justificativa, a série temporal se torna descritiva, mas pouco interpretável.

A delimitação documental também influencia a pergunta. Artigos, revisões, trabalhos de congresso, capítulos e *preprints* têm papéis diferentes na comunicação científica. Em áreas

como ciência da computação, anais de congressos podem ser centrais; em biomedicina, artigos de periódico tendem a concentrar prestígio e indexação; em humanidades, livros e capítulos podem ter peso maior. Uma pergunta sobre produção científica que ignora diferenças disciplinares corre o risco de comparar práticas de publicação incomparáveis. A IA pode auxiliar elaborando alertas disciplinares, mas a decisão final exige conhecimento contextual.

Perguntas de produção podem ser articuladas a indicadores simples e complexos. Entre os simples estão número de publicações por ano, taxa de crescimento anual, número de documentos por autor, número de documentos por instituição e distribuição por periódico. Entre os mais elaborados estão índice de especialização temática, produtividade fracionada, diversidade de áreas e crescimento relativo normalizado. A escolha depende da pergunta. Não há indicador universal; há indicadores adequados a determinados problemas.

Em estudos assistidos por *prompts*, recomenda-se que a pergunta de produção seja acompanhada de um pequeno protocolo. O protocolo deve explicitar tema, base de dados, período, tipos documentais, idioma, campos pesquisados, unidade de contagem, regra de atribuição e limitações esperadas. O *prompt* não deve apenas pedir ‹faça uma pergunta sobre produção›, mas exigir que a IA indique requisitos de dados e riscos de interpretação. Essa exigência transforma o modelo em auxiliar de desenho metodológico, e não em gerador de frases genéricas.

Uma boa pergunta sobre produção científica apresenta três características. Primeiro, delimita o universo documental: produção de quem, sobre o quê, onde indexada e em qual período. Segundo, define a operação analítica: descrever evolução, comparar grupos, identificar concentração, analisar dispersão ou mapear especialização. Terceiro, explicita o uso pretendido: diagnóstico exploratório, relatório institucional, revisão de área, prospecção estratégica ou avaliação de política. Quando essas características estão presentes, a contagem de publicações deixa de ser um número solto e passa a integrar uma interpretação cientométrica.

A formulação de perguntas de produção também deve reconhecer a diferença entre crescimento absoluto e crescimento relativo. Um campo pode apresentar muitos artigos, mas crescimento percentual modesto; outro pode ter poucos artigos, mas crescimento acelerado. A pergunta deve indicar se busca volume, expansão, aceleração, desaceleração ou comparação de trajetórias. Essa distinção é essencial em diagnósticos de temas emergentes, pois áreas pequenas podem parecer explosivas quando analisadas apenas por percentuais. A IA pode sugerir ambas as métricas, mas cabe ao pesquisador decidir como interpretá-las.

Em síntese, perguntas sobre produção científica são indispensáveis, porém insuficientes. Elas fornecem a cartografia inicial do volume e da distribuição da pesquisa, mas não explicam sozinhas influência, colaboração, circulação ou qualidade. Em uma cientometria orientada por IA, o desafio é usar *prompts* para refinar a pergunta, revelar pressupostos de contagem e estabelecer limites interpretativos. O resultado esperado não é uma pergunta mais sofisticada por ornamentação linguística, mas uma pergunta mais operacional, validável e conectada ao objetivo do estudo.

6.3 Perguntas sobre impacto e influência

Perguntas sobre impacto e influência ocupam lugar central na cientometria porque buscam compreender como a produção científica circula, é reconhecida, incorporada ou mobilizada por outros atores. Ao contrário das perguntas de produção, que se concentram no volume de documentos, as perguntas de impacto investigam relações posteriores à publicação: citações recebidas, presença em documentos de política, menções sociais, uso em patentes,

repercussão em diretrizes clínicas, circulação em ambientes digitais ou influência sobre frentes de pesquisa. Essa diversidade exige cautela, pois 'impacto' não é um fenômeno único.

Figura 6.3 - Dimensões de impacto

Fonte: Imagem gerada por IA, ChatGPT (2026).

A pergunta 'quais artigos tiveram maior impacto?' é metodologicamente incompleta. Impacto em qual sentido? Impacto acadêmico medido por citações? Impacto tecnológico medido por patentes? Impacto social inferido por menções em notícias? Impacto político observado em documentos governamentais? Impacto educacional expresso em uso didático? Cada dimensão demanda fontes, indicadores e interpretações distintas. A formulação da pergunta deve, portanto, substituir a noção abstrata de impacto por uma dimensão observável.

A citação é o indicador clássico de influência científica. Garfield propôs os índices de citação como instrumentos para recuperar literatura por associação de ideias; Price demonstrou que citações compõem redes de documentos; Small mostrou que cocitações podem revelar proximidades intelectuais. Contudo, a literatura também indica que citações têm significados variados. Podem expressar reconhecimento, uso metodológico, crítica, obrigação retórica, pertencimento a uma tradição ou simples convenção. Bornmann e Daniel, ao revisar estudos sobre comportamento de citação, mostram que a interpretação de citações requer prudência (GARFIELD, 1955; PRICE, 1965; SMALL, 1973; BORNMANN; DANIEL, 2008).

Perguntas de impacto baseadas em citações devem especificar janela temporal. Artigos recém-publicados tendem a ter menos citações simplesmente porque tiveram menos tempo para circular. Algumas áreas acumulam citações rapidamente; outras têm ciclos longos. Revisões costumam receber mais citações que estudos empíricos; métodos e bases de dados podem se tornar documentos altamente citados. Portanto, perguntar 'quais artigos são mais citados?' sem normalização temporal, disciplinar ou documental pode reproduzir vieses conhecidos. Uma formulação mais robusta perguntaria: 'quais documentos apresentam impacto de citação normalizado por ano de publicação e área temática?' (WALTMAN, 2016; TAHAMTAN; SAFIPOUR AFSHAR; AHAMADZADEH, 2016).

A influência também pode ser observada por redes de cocitação e acoplamento bibliográfico. Quando dois documentos são frequentemente citados juntos, podem representar uma base intelectual compartilhada; quando dois documentos citam referências semelhantes, podem indicar proximidade temática. Perguntas sobre influência intelectual, portanto, não se limitam a rankings de citações. Podem investigar quais obras estruturam um campo, quais escolas teóricas se aproximam, quais documentos conectam comunidades e quais referências funcionam como pontes entre áreas.

A IA pode ajudar a formular perguntas de impacto ao diferenciar indicadores possíveis. Ao receber o objetivo 'avaliar a influência da pesquisa brasileira em biotecnologia', um *prompt* bem estruturado deve solicitar que o modelo apresente alternativas: impacto acadêmico por citações normalizadas, influência em colaboração internacional, presença em patentes, uso em documentos regulatórios, inserção em periódicos de alto impacto relativo e papel em redes de cocitação. Cada alternativa deve ser acompanhada de limitações. Essa comparação ajuda a evitar que o pesquisador reduza influência a uma única métrica.

A formulação de perguntas sobre impacto deve considerar a assimetria entre visibilidade e influência. Um artigo pode ser amplamente mencionado em redes sociais, mas pouco incorporado à literatura científica; outro pode ser altamente citado em uma comunidade técnica restrita, mas invisível ao público geral. Altmetria, Webometria e indicadores de atenção digital ampliam a observação da circulação, mas não substituem citações. Bornmann discute que altmétricas podem apontar para dimensões de atenção e possível impacto mais amplo, embora apresentem limitações de cobertura, estabilidade e interpretação (BORNMANN, 2014).

Perguntas avaliativas sobre impacto exigem especial responsabilidade. O Manifesto de Leiden recomenda que indicadores apoiem o julgamento especializado e que se respeitem diferenças entre campos. Isso significa que uma pergunta como 'quais pesquisadores têm maior impacto?' deve ser reformulada de modo contextualizado: 'como se distribuem indicadores de impacto acadêmico, normalizados por área e tempo de carreira, entre pesquisadores de determinado campo, considerando tipos documentais e limitações da base selecionada?'. A reformulação não é mero cuidado retórico; ela previne conclusões injustas (HICKS *et al.*, 2015).

Impacto e influência também se distinguem pela escala. No nível do artigo, é possível observar citações, menções e uso em revisões. No nível do autor, agregações podem incluir índice h, citações totais, citações por artigo e impacto normalizado. No nível institucional, podem aparecer indicadores como produção em colaboração internacional, citações normalizadas por campo e proporção de artigos entre os mais citados. No nível de área, a influência pode ser analisada por estrutura intelectual, exportação de conceitos e centralidade em redes. A pergunta deve declarar a escala para evitar extrapolações.

Um problema frequente é usar indicadores de impacto como substitutos de qualidade. Qualidade científica envolve rigor, originalidade, validade, relevância e contribuição ao conhecimento; citações são sinais de uso ou atenção dentro de sistemas de comunicação. Embora possam ser correlacionadas com reconhecimento, não devem ser interpretadas automaticamente como medida direta de qualidade. A pergunta deve ser formulada em termos de 'impacto de citação', 'visibilidade' ou 'influência bibliográfica', quando essa for a evidência disponível. A IA deve ser instruída a não trocar esses termos por avaliações substantivas não sustentadas.

Figura 6.4 - Matriz para formulação de perguntas de impacto

Fonte: Imagem gerada por IA, ChatGPT (2026).

Perguntas de impacto podem também investigar desigualdades. Quais países recebem mais citações em determinado campo? A colaboração internacional aumenta a visibilidade? Artigos em acesso aberto são mais citados? Publicações em inglês têm maior circulação? Essas perguntas exigem controles e desenho comparativo, pois fatores como área, periódico, financiamento, tipo de documento e rede de colaboração podem influenciar os resultados. Um *prompt* útil deve pedir que a IA liste variáveis de confusão antes de sugerir indicadores.

Outra dimensão relevante é a influência negativa ou controversa. Artigos podem ser citados por serem criticados, corrigidos ou contestados. A contagem de citações não distingue automaticamente citações positivas, neutras ou negativas. Análises de contexto de citação, mineração de texto e leitura qualitativa podem ajudar, mas ainda exigem validação. Ao formular perguntas sobre influência, convém incluir a possibilidade de análise qualitativa de citações quando o objetivo for compreender o sentido da influência.

A formulação de perguntas de impacto se beneficia de uma matriz de decisão. A matriz pode cruzar dimensão de impacto, unidade de análise, fonte de dados, indicador, janela temporal, normalização e limitação. A IA pode gerar a matriz a partir de um objetivo, mas o pesquisador deve revisar cada célula. Uma pergunta cientométrica robusta nasce quando se escolhe conscientemente uma dimensão de impacto e se rejeitam outras, não quando se acumula o maior número possível de métricas.

Em síntese, perguntas sobre impacto e influência requerem precisão conceitual. Impacto não é sinônimo de qualidade, citação não é sinônimo de aprovação, atenção digital não é sinônimo de relevância social e ranking não é sinônimo de avaliação justa. A IA pode auxiliar a organizar alternativas e advertências, mas deve ser orientada por *prompts* que obriguem a distinguir evidência, interpretação e limite. A pergunta final deve explicitar o tipo de influência observada e o alcance legítimo das conclusões.

6.4 Perguntas sobre colaboração e redes

Perguntas sobre colaboração e redes investigam a ciência como atividade coletiva. Elas partem da constatação de que a produção do conhecimento depende de relações entre autores, grupos, instituições, países, áreas, conceitos, documentos e infraestruturas. A colaboração pode ser observada por coautoria, coafiliação, coparticipação em projetos, compartilhamento de referências, cocitação, acoplamento bibliográfico e coocorrência de termos. A formulação da pergunta deve definir que tipo de relação será tomada como evidência de colaboração ou proximidade, pois nem toda rede cientométrica representa colaboração direta.

A coautoria é a forma mais visível de colaboração bibliográfica. Perguntas como 'quais autores colaboram entre si em determinado campo?', 'quais instituições conectam diferentes comunidades?' ou 'como evoluiu a colaboração internacional em uma área?' dependem de metadados de autoria e afiliação. Contudo, coautoria não esgota a colaboração científica. Pesquisadores podem colaborar sem publicar juntos, e artigos multiautorais podem expressar relações muito distintas, desde trabalho intelectual conjunto até participação técnica, coordenação de consórcio ou contribuição de infraestrutura.

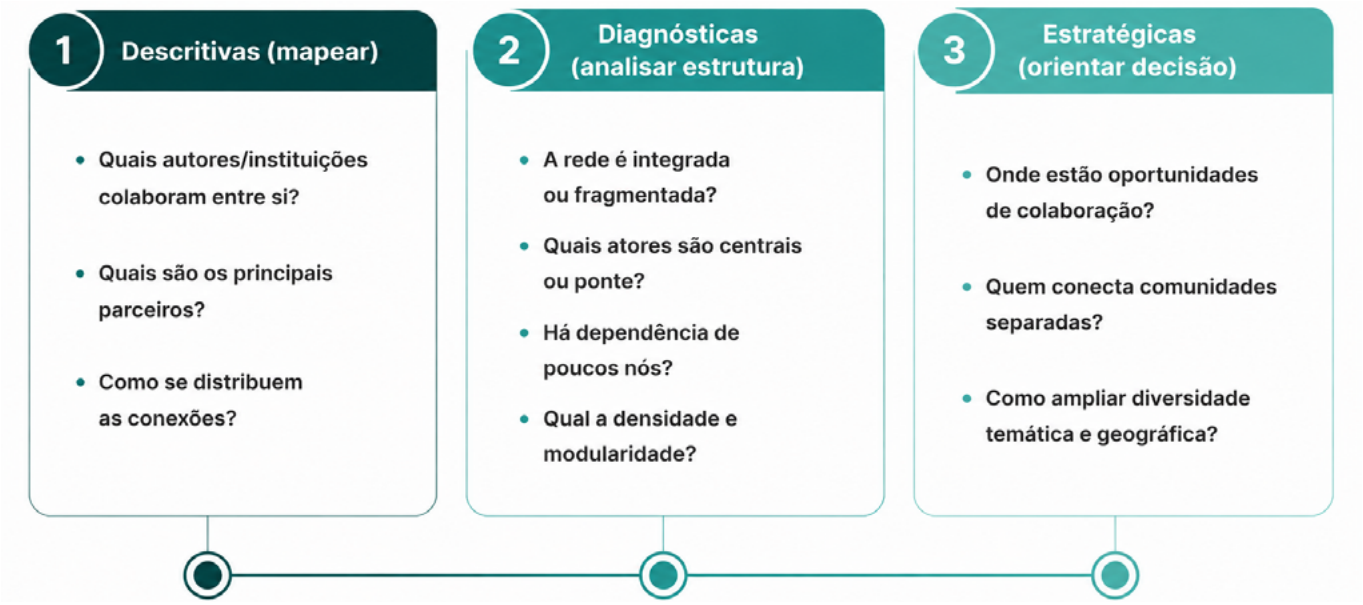
Newman mostrou que redes de colaboração científica apresentam estruturas complexas, com componentes, graus de centralidade e padrões de conectividade. Essas propriedades permitem formular perguntas mais sofisticadas que a simples contagem de coautorias. Em vez de perguntar apenas 'quem colaborou mais?', pode-se perguntar: 'quais atores ocupam posições de intermediação entre comunidades?', 'a rede é fragmentada ou integrada?', 'há dependência de poucos nós centrais?', 'a colaboração internacional amplia a diversidade temática?', ou 'quais instituições conectam subcampos antes separados?' (NEWMAN, 2001).

Em perguntas sobre redes, a unidade de análise e o tipo de aresta são inseparáveis. Se os nós são autores, as arestas podem representar coautoria; se os nós são instituições, as arestas podem representar publicações conjuntas; se os nós são países, as arestas podem representar coautoria internacional; se os nós são documentos, as arestas podem representar citação direta, cocitação ou acoplamento; se os nós são palavras-chave, as arestas podem representar coocorrência. Cada escolha produz um mapa diferente da ciência. Não existe 'a' rede de uma área; existem redes construídas segundo perguntas específicas.

A IA pode apoiar a formulação de perguntas de rede ao sugerir correspondências entre fenômeno e técnica. Se o pesquisador deseja compreender 'comunidades intelectuais', o modelo pode indicar cocitação ou acoplamento bibliográfico; se deseja mapear 'colaboração institucional', pode indicar coautoria por afiliação; se deseja analisar 'estrutura conceitual', pode indicar coocorrência de palavras-chave ou termos de resumo; se deseja observar 'trajetórias de temas', pode propor redes temporais. O *prompt* deve exigir que cada sugestão venha acompanhada de justificativa e limitação.

Um risco recorrente é interpretar proximidade visual como proximidade substantiva sem verificar o método de construção do mapa. Mapas de rede dependem de normalização, algoritmo de layout, limiar de inclusão, método de *cluster*ização e tratamento de dados. Dois nós próximos em uma visualização podem parecer fortemente relacionados, mas essa relação depende da matriz subjacente. Perguntas de rede devem, portanto, incluir a dimensão metodológica: 'que tipo de relação será mapeada, como será normalizada e com que limiares?'.

Figura 6.5 - Perguntas sobre colaboração e redes



Fonte: Imagem gerada por IA, ChatGPT (2026).

Perguntas sobre colaboração podem ser descritivas, diagnósticas ou estratégicas. Uma pergunta descritiva busca identificar parceiros existentes. Uma pergunta diagnóstica analisa integração, centralidade, fragmentação ou dependência. Uma pergunta estratégica busca oportunidades de colaboração, lacunas de conexão e atores ponte. A formulação deve deixar claro se o objetivo é conhecer uma estrutura já existente ou orientar uma decisão futura. O uso estratégico exige cautela adicional, pois métricas de rede podem sugerir oportunidades, mas não avaliam compatibilidade científica, ética, política ou institucional.

Redes de colaboração também devem ser analisadas em termos de assimetria. A colaboração internacional pode elevar visibilidade, mas pode reproduzir desigualdades entre centros e periferias. Instituições de países com maior infraestrutura podem ocupar posições centrais; grupos periféricos podem aparecer como parceiros secundários; autores de maior prestígio podem concentrar conexões. Perguntas responsáveis devem contemplar não apenas a existência de colaboração, mas sua distribuição, reciprocidade e diversidade. A IA pode auxiliar propondo indicadores como grau, centralidade de intermediação, densidade, modularidade e diversidade geográfica, mas a interpretação exige contexto.

Em redes conceituais, a pergunta sobre colaboração é substituída por pergunta sobre associação de temas. Termos que coocorrem em títulos, resumos ou palavras-chave podem indicar proximidade temática, mas não necessariamente relação causal ou integração teórica. A pergunta deve diferenciar 'quais termos aparecem juntos?' de 'quais conceitos estruturam a área?' e de 'quais abordagens teóricas estão integradas?'. O primeiro nível pode ser respondido por coocorrência; os demais exigem interpretação qualitativa e, frequentemente, leitura de documentos representativos.

A temporalidade amplia a força das perguntas de rede. Uma rede agregada de vinte anos pode ocultar mudanças importantes. Perguntas como 'quais instituições se tornaram centrais nos últimos cinco anos?' ou 'quais comunidades temáticas emergiram após determinado evento?' exigem redes longitudinais, janelas móveis ou comparação de períodos. A IA pode ajudar a propor recortes temporais, mas deve ser instruída a justificar a escolha e a reconhecer efeitos de atraso de indexação, crescimento geral da literatura e mudanças terminológicas.

A formulação de perguntas de rede deve considerar o tratamento de nomes. Autores homônimos, variações de afiliação e mudanças institucionais podem distorcer resultados. Antes de perguntar 'quais autores são centrais?', é preciso perguntar se os dados permitem desambiguação confiável. Em instituições, nomes abreviados, traduções e unidades subordinadas exigem normalização. Em países, afiliações múltiplas e contagem fracionada devem ser explicitadas. Sem esses cuidados, a rede pode parecer precisa, mas representar artefatos de metadados (VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017).

Prompts para redes devem pedir ao modelo que separe três níveis: desenho da rede, indicadores de rede e interpretação. O desenho nos define, arestas, pesos, fontes e recortes. Os indicadores incluem grau, força, centralidade, densidade, modularidade e componentes. A interpretação discute o que esses indicadores podem sugerir e o que não podem provar. Essa separação é essencial porque modelos generativos tendem a produzir narrativas coerentes a partir de descrições incompletas, podendo atribuir intencionalidade ou causalidade a padrões relacionais.

Uma pergunta bem formulada sobre colaboração e redes deve, portanto, explicitar: o tipo de rede, a unidade de análise, a relação observada, o período, a fonte de dados, os procedimentos de normalização, os indicadores pretendidos e o objetivo interpretativo. Em vez de 'quem são os principais parceiros?', uma pergunta mais robusta seria: 'como se estruturou a rede de coautoria institucional em pesquisas sobre transição energética no período 2015-2025, considerando contagem fracionada, centralidade de intermediação e comunidades detectadas por modularidade?'.

Em síntese, perguntas sobre colaboração e redes deslocam a cientometria da contagem para a relação. Elas permitem compreender conectividade, comunidades, fronteiras, pontes e padrões de circulação. Contudo, também são vulneráveis a interpretações excessivas, especialmente quando mapas visuais parecem autoexplicativos. Em estudos assistidos por IA, a função do *prompt* é obrigar o modelo a declarar a arquitetura da rede e a limitar a interpretação ao que a evidência relacional sustenta.

6.5 Perguntas sobre emergência de temas

Perguntas sobre emergência de temas estão entre as mais atraentes da cientometria contemporânea, especialmente em ambientes de gestão estratégica, prospecção tecnológica e formulação de políticas científicas. Elas buscam identificar assuntos, conceitos, métodos, aplicações ou combinações de áreas que apresentam crescimento, novidade, centralidade crescente ou potencial de consolidação. Contudo, a emergência é um conceito mais complexo do que o aumento de frequência de palavras. Um tema emergente pode crescer em publicações, aparecer em novos contextos, conectar comunidades antes separadas, receber financiamento recente ou adquirir relevância social em determinado momento histórico.

A pergunta 'quais são os temas emergentes?' deve ser substituída por formulações mais precisas. Emergentes em qual campo? Em qual base? Em qual período? Emergentes por frequência, por citações recentes, por explosão de termos, por conexão interdisciplinar, por presença em patentes, por atenção social ou por prioridade política? Cada critério produz resultados diferentes. Um termo pode crescer rapidamente em resumos, mas não se consolidar como frente de pesquisa; outro pode ser pouco frequente, mas altamente conectado a documentos influentes.

A análise de temas emergentes dialoga com métodos de coocorrência de termos, análise de palavras-chave, mineração de textos, detecção de rajadas, cocitação temporal e mapas temáticos. Chen, ao desenvolver o CiteSpace para detectar tendências emergentes e padrões transitórios, contribuiu para uma abordagem em que a temporalidade e a visualização são centrais. Callon, Courtial e Laville, ao discutirem análise de *co-words*, mostraram que palavras podem funcionar como traços das relações entre problemas científicos e tecnológicos. Essas tradições ajudam a formular perguntas que combinam léxico, rede e tempo (CHEN, 2006; CALLON; COURTIAL; LAVILLE, 1991).

Quadro 6.1 - Categorias de indicadores de emergência temática

Categoria	Indicadores (exemplos)	O que sugerem
Crescimento	Taxa de crescimento anual; crescimento relativo; participação em janelas móveis	Aceleração do volume
Léxico	Frequência recente de termos; *burst* de termos; novos n-grams	Mudança ou adoção terminológica
Citações	Citações recentes (2 a 3 anos); aceleração de citações; referências jovens	Atenção científica recente
Rede	Aumento de centralidade; surgimento de *clusters*; conexões entre comunidades	Reorganização estrutural e integração
Atores	Novos autores e instituições; diversidade de origem geográfica	Entrada de novos atores e difusão
Interdisciplinaridade	Amplitude de áreas; diversidade de categorias; combinações novas	Expansão para outras fronteiras
Outros sinais	Financiamentos recentes; patentes; políticas; mídia; downloads; altmetria	Atenção social, tecnológica ou política
Nenhum sinal isolado confirma emergência: a leitura deve combinar categorias e ser confrontada com o contexto do campo e a cobertura da base.		

Fonte: Elaborado pelo autor (2026).

A IA generativa pode apoiar a emergência temática de duas maneiras. Primeiro, pode ajudar a formular perguntas e hipóteses antes da análise, sugerindo dimensões terminológicas, sinônimos, conceitos relacionados e possíveis mudanças de nomenclatura. Segundo, pode apoiar a interpretação posterior, resumindo *clusters*, propondo rótulos e indicando relações plausíveis entre subtemas. Em ambos os casos, a IA deve operar sobre dados fornecidos e não inventar tendências. Um *prompt* adequado deve exigir que toda afirmação sobre emergência seja vinculada a indicadores ou evidências textuais.

Uma dificuldade central é distinguir emergência de moda terminológica. Certos termos crescem porque são adotados por políticas editoriais, programas de financiamento, discursos institucionais ou mudanças de vocabulário, sem necessariamente indicar transformação cognitiva profunda. A pergunta cientométrica deve contemplar essa possibilidade. Em vez de perguntar apenas se um termo aumentou, pode-se perguntar: 'o crescimento do termo corresponde à formação de um *cluster* consistente, ao aumento de citações recentes e à diversificação de autores e instituições?'. Essa formulação reduz o risco de confundir palavra em alta com campo emergente.

A emergência pode ser temática, metodológica, instrumental ou institucional. Um tema emergente refere-se a novo objeto de pesquisa; uma metodologia emergente envolve técnicas em expansão; uma infraestrutura emergente envolve bases, plataformas ou instrumentos; uma configuração institucional emergente envolve novas redes de atores. A pergunta deve indicar qual dimensão está em foco. Por exemplo, em 'inteligência artificial na saúde', a emergência pode estar em modelos generativos, triagem clínica, diagnóstico por imagem, ética algorítmica, regulação ou integração hospitalar. Cada dimensão exige dados e termos específicos.

O recorte temporal é particularmente sensível. Se o intervalo for longo demais, temas recentes podem ficar diluídos; se for curto demais, picos episódicos podem parecer tendências. Uma solução é combinar janelas: análise histórica ampla para contextualização e análise recente para detecção de aceleração. A pergunta pode assumir a forma: 'quais temas apresentam crescimento relativo acelerado entre 2021 e 2025, quando comparados à trajetória geral do campo entre 2010 e 2025?'. Essa formulação permite distinguir novidade recente de crescimento acumulado.

Perguntas sobre emergência de temas devem também considerar a base de dados. Web of Science, Scopus, PubMed, Dimensions, OpenAlex e bases especializadas possuem coberturas diferentes. Temas emergentes podem aparecer primeiro em *preprints*, anais de congresso, patentes ou literatura cinzenta, antes de se consolidarem em periódicos indexados. Em ciência da computação, conferências podem antecipar tendências; em biomedicina, *preprints* e ensaios clínicos podem indicar movimentos recentes; em políticas públicas, relatórios institucionais podem preceder artigos. A pergunta deve alinhar fonte e fenômeno.

A análise de emergência pode envolver indicadores como taxa de crescimento anual, frequência recente de termos, *burst detection*, idade média das referências, citações recentes, centralidade temporal, novos autores, diversidade institucional, interdisciplinaridade e presença em documentos estratégicos. Nenhum indicador é suficiente isoladamente. A pergunta deve definir se busca sinais fracos, tendências consolidadas ou frentes de pesquisa influentes. Sinais fracos exigem sensibilidade; tendências consolidadas exigem consistência; frentes influentes exigem relações com citações e redes.

A IA pode ser útil na construção de taxonomias preliminares. Dado um conjunto de títulos e resumos, um modelo pode sugerir grupos temáticos e nomeá-los. Contudo, esse processo deve ser validado por amostragem, comparação com palavras-chave originais e leitura de documentos centrais. A pergunta não deve ser 'quais temas a IA encontrou?', mas

‘quais agrupamentos temáticos, derivados dos dados e validados por critérios explícitos, sugerem emergência, consolidação ou declínio?’. Essa distinção evita atribuir autoridade analítica ao modelo.

Outra cautela diz respeito à novidade. Um tema pode ser novo em uma área, mas antigo em outra. A pergunta deve diferenciar novidade local e novidade global. Por exemplo, uma técnica pode ser emergente na educação, mas consolidada na ciência da computação. Um conceito pode emergir em periódicos de gestão, mas já estar presente em sociologia. Perguntas sobre emergência devem, portanto, considerar fronteiras disciplinares e circulação entre campos. A IA pode ajudar a identificar áreas de origem e de adoção, mas precisa de dados comparativos.

A emergência também pode ser interpretada a partir de conexões entre temas. Quando termos de comunidades diferentes começam a coocorrer, pode haver formação de agenda interdisciplinar. Perguntas desse tipo não se concentram apenas em termos mais frequentes, mas em combinações novas. Exemplo: ‘que combinações entre IA generativa, avaliação científica, integridade de pesquisa e bibliometria aparecem com maior frequência nos últimos anos?’. Essa formulação é especialmente pertinente à trilha deste livro, pois a inovação pode surgir da recombinação entre cientometria e modelos generativos.

Prompts para emergência devem exigir que o modelo diferencie evidência quantitativa, interpretação temática e hipótese prospectiva. A evidência quantitativa pode ser aumento de frequência ou centralidade. A interpretação temática pode ser a descrição de um *cluster*. A hipótese prospectiva pode sugerir consolidação futura, mas deve ser marcada como hipótese. Sem essa diferenciação, o texto pode converter uma tendência observada em previsão indevida. Em ciência, tendência não é destino.

Em síntese, perguntas sobre emergência de temas exigem uma combinação de tempo, texto, rede e interpretação. Elas são poderosas para prospecção, mas vulneráveis a modismos, vieses de indexação e excesso de narrativa. A IA pode ampliar a capacidade de explorar vocabulários e hipóteses, mas deve ser controlada por *prompts* que exijam evidências, critérios e níveis de incerteza. Uma pergunta cientométrica sobre emergência só é robusta quando define o que conta como emergência e como essa emergência será distinguida de ruído, moda ou mudança terminológica.

6.6 Perguntas sobre políticas científicas e avaliação institucional

Perguntas sobre políticas científicas e avaliação institucional conectam a cientometria à tomada de decisão. Diferentemente de estudos puramente descritivos, essas perguntas frequentemente envolvem consequências práticas: distribuição de recursos, desenho de programas, avaliação de desempenho, definição de prioridades, identificação de capacidades, monitoramento de internacionalização e análise de retorno de investimentos. Por isso, são metodologicamente sensíveis e eticamente exigentes. Um indicador mal escolhido pode produzir diagnóstico distorcido; uma comparação sem contexto pode induzir decisões injustas.

A avaliação institucional deve começar pela missão. Uma universidade de pesquisa intensiva, um instituto tecnológico, uma faculdade regional, um hospital universitário e um centro de pesquisa aplicado não devem ser avaliados pelos mesmos critérios sem contextualização. O Manifesto de Leiden enfatiza que a avaliação quantitativa deve apoiar avaliação qualitativa especializada e que o desempenho deve ser medido em relação às missões de pesquisa da instituição, do grupo ou do pesquisador. Logo, a pergunta cientométrica deve explicitar a finalidade institucional e o contexto de comparação (HICKS *et al.*, 2015).

Perguntas como 'qual é a melhor universidade?' ou 'qual programa é mais produtivo?' tendem a ser cientometricamente inadequadas quando não definem campo, período, missão, tamanho, recursos, perfil docente, composição disciplinar e indicadores. Uma formulação mais responsável seria: 'como evoluíram, entre 2018 e 2025, a produção, o impacto normalizado, a colaboração internacional e a diversidade temática dos programas de uma instituição, considerando diferenças de área e tamanho do corpo docente?'. A pergunta deixa de buscar um julgamento absoluto e passa a estruturar um diagnóstico contextual.

Em políticas científicas, perguntas podem se concentrar em efeitos de programas. Por exemplo: 'após a criação de um edital de pesquisa em energias renováveis, houve crescimento da produção, diversificação de instituições e aumento de colaboração internacional no tema?'. Essa pergunta exige desenho quase avaliativo, com período anterior e posterior, grupo de comparação, identificação de publicações relacionadas ao edital e cautela na inferência causal. A cientometria pode fornecer sinais, mas raramente prova causalidade sozinha. A IA deve ser instruída a distinguir associação temporal de efeito causal.

Outra classe de perguntas diz respeito a capacidades científicas. Agências e instituições podem perguntar: 'em quais áreas temos massa crítica?', 'quais grupos atuam em temas estratégicos?', 'quais competências estão dispersas?', 'quais redes podem ser fortalecidas?'. Essas perguntas combinam produção, colaboração, temas e impacto. O *prompt* pode ajudar a decompor a demanda em subperguntas: volume por tema, autores ativos, instituições envolvidas, redes de coautoria, documentos influentes, evolução temporal e lacunas. A decomposição evita que uma pergunta estratégica seja respondida por um único ranking.

A avaliação institucional também envolve comparação. Comparar instituições sem normalização é uma das fontes mais comuns de erro. Instituições maiores tendem a produzir mais; áreas biomédicas tendem a citar mais que humanidades; bases internacionais cobrem melhor periódicos em inglês; colaborações internacionais podem aumentar citações. A pergunta deve definir se a comparação será por volume absoluto, produção per capita, impacto normalizado por campo, colaboração relativa, diversidade temática ou indicadores compostos. Cada escolha expressa valores e pressupostos.

A IA pode ser particularmente útil como ferramenta de crítica de perguntas avaliativas. Um *prompt* pode solicitar: 'identifique riscos de injustiça, variáveis de contexto e indicadores inadequados nesta pergunta'. Esse uso crítico é mais importante que o uso produtivo. Modelos de linguagem podem ajudar a revelar pressupostos ocultos, como comparar áreas distintas, ignorar tamanho institucional, usar índice h para grupos, privilegiar periódicos indexados internacionalmente ou confundir visibilidade com qualidade. A IA funciona, nesse caso, como checklist argumentativo, não como avaliadora.

Perguntas sobre políticas científicas também podem abordar desigualdade e inclusão. Como se distribui a colaboração internacional entre regiões? Quais instituições periféricas participam de redes de pesquisa estratégicas? A produção em língua local é invisibilizada por bases internacionais? Mulheres e grupos sub-representados aparecem em posições de liderança autoral? Essas questões exigem dados sensíveis e cuidados éticos, especialmente quando envolvem atributos pessoais. A formulação deve prever anonimização, agregação e justificativa de uso. A IA não deve inferir atributos pessoais a partir de nomes sem base confiável e autorização ética.

A relação entre ciência e sociedade amplia o escopo das perguntas. Políticas científicas podem demandar evidências de impacto social, contribuição para objetivos de desenvolvimento sustentável, presença em diretrizes, transferência tecnológica ou participação em debates públicos. No entanto, indicadores de impacto social são mais difíceis de padronizar que citações.

A pergunta deve definir quais evidências serão aceitas: menções em documentos de política, patentes, guias clínicos, relatórios técnicos, parcerias, bases de inovação ou altmétricas. Sem essa definição, 'impacto social' permanece uma categoria retórica.

Outro desafio é a temporalidade das políticas. Programas científicos podem produzir efeitos em horizontes diferentes. Publicações podem surgir anos após o financiamento; citações demoram ainda mais; formação de redes pode preceder resultados bibliográficos; impactos sociais podem não ser capturados por artigos. Perguntas avaliativas devem alinhar janela de observação ao tipo de efeito esperado. Uma avaliação de curto prazo pode medir mobilização e produção inicial; uma avaliação de médio prazo pode medir colaboração e impacto acadêmico; uma avaliação de longo prazo pode investigar consolidação de áreas e efeitos sociais.

A formulação de perguntas institucionais deve incluir governança dos dados. Quem fornece os dados? Que bases serão usadas? Como serão tratados nomes de instituições, pesquisadores e áreas? Quem validará os resultados? Como serão comunicadas incertezas? A IA pode ajudar a construir um plano de governança, mas não pode assumir responsabilidade institucional. Estudos que afetam reputação e recursos devem ter critérios transparentes, revisão por especialistas e possibilidade de contestação.

Em relatórios de gestão, há tendência a privilegiar indicadores sintéticos. Embora úteis para comunicação, indicadores compostos podem ocultar escolhas arbitrárias de ponderação. A pergunta cientométrica deve evitar começar por um índice único. É melhor iniciar por dimensões: produtividade, impacto, colaboração, formação de redes, diversidade temática, internacionalização e alinhamento estratégico. Somente depois se pode discutir se alguma síntese é apropriada. A IA deve ser instruída a não criar rankings compostos sem metodologia explícita.

Perguntas sobre avaliação institucional também devem contemplar o que não será medido. Ensino, extensão, formação, liderança comunitária, inovação social e contribuição cultural podem ficar fora de bases bibliográficas. Uma pergunta responsável inclui limites: 'este estudo avalia produção científica indexada, não desempenho global da instituição'. Essa frase simples previne usos indevidos. A IA pode ser solicitada a redigir uma seção de escopo e limitações antes de qualquer análise.

Em síntese, perguntas sobre políticas científicas e avaliação institucional exigem a maior cautela entre as categorias discutidas neste capítulo. Elas não são apenas técnicas; são normativas, porque definem quais valores serão quantificados e quais dimensões permanecerão invisíveis. A IA pode apoiar a formulação, a crítica e a documentação, mas deve permanecer subordinada a princípios de avaliação responsável. A pergunta final deve ser justa, contextualizada, transparente e proporcional ao uso pretendido.

Quadro 6.2 - Checklist de responsabilidade

CHECKLIST DE RESPONSABILIDADE *(antes de formular uma pergunta avaliativa com métricas de pesquisa)*
☐ A pergunta está alinhada à missão e aos objetivos institucionais?
☐ Os indicadores são adequados ao campo e comparáveis?
☐ As normalizações consideradas (área, período, tamanho) são adequadas?
☐ Há risco de viés de cobertura (idioma, base, tipo documental)?
☐ O que fica de fora do que será medido?

CHECKLIST DE RESPONSABILIDADE *(antes de formular uma pergunta avaliativa com métricas de pesquisa)*
☐ Os resultados podem reforçar desigualdades ou estigmas?
☐ Quais decisões de valor foram feitas (ponderações, prioridades)?
☐ Quem validará os resultados e poderá contestá-los?
☐ Como as incertezas serão comunicadas?

Registro da verificação

Pergunta avaliada	Responsável	Data

Fonte: Elaborado pelo autor (2026).

6.7 Transformando objetivos vagos em perguntas operacionalizáveis

A transformação de objetivos vagos em perguntas operacionalizáveis é a competência central deste capítulo. Muitos estudos cientométricos começam com formulações amplas: 'mapear a literatura sobre um tema', 'avaliar a produção de uma instituição', 'identificar tendências', 'analisar impacto', 'entender colaborações'. Embora legítimos como intenções iniciais, esses objetivos não bastam para orientar coleta, análise e interpretação. A operacionalização consiste em converter intenção em pergunta respondível por dados específicos, com unidades de análise, indicadores, métodos e limites.

Figura 6.6 - Sete passos para transformar objetivos em perguntas operacionalizáveis

Fonte: Imagem gerada por IA, ChatGPT (2026).

O primeiro passo é identificar o fenômeno principal. O objetivo fala de produção, impacto, colaboração, circulação, emergência, avaliação, estrutura intelectual ou política

científica? Frequentemente, um objetivo contém vários fenômenos. 'Mapear a pesquisa sobre IA na educação' pode envolver evolução da produção, países mais produtivos, redes de colaboração, temas emergentes, documentos influentes e lacunas. A operacionalização exige desdobrar o objetivo em perguntas secundárias, em vez de tentar responder tudo por uma única análise.

O segundo passo é definir a unidade de análise. Perguntas diferentes podem usar a mesma base, mas unidades distintas. Artigos servem para produção e impacto documental; autores para produtividade e colaboração; instituições para diagnóstico organizacional; países para geografia da ciência; periódicos para canais de comunicação; palavras-chave para estrutura temática; referências para base intelectual; citações para influência; conceitos para análise semântica. A pergunta deve nomear a unidade de modo explícito.

O terceiro passo é delimitar o escopo. Escopo inclui tema, período, base de dados, idioma, tipo documental, áreas do conhecimento e critérios de inclusão. A IA pode ajudar a listar possibilidades, mas a escolha deve ser fundamentada. Um escopo amplo aumenta cobertura, mas também ruído; um escopo estreito aumenta precisão, mas pode excluir literatura relevante. A pergunta operacionalizada deve refletir esse equilíbrio. Por exemplo: 'como evoluiu a produção de artigos e revisões sobre aprendizagem federada em saúde, indexados na Scopus e Web of Science, entre 2016 e 2025?'.

O quarto passo é escolher indicadores coerentes. A pergunta deve orientar o indicador, não o contrário. Se a pergunta é sobre crescimento, indicadores temporais são centrais; se é sobre influência, citações normalizadas e redes de cocitação podem ser adequadas; se é sobre colaboração, coautoria e centralidade; se é sobre temas emergentes, coocorrência, *bursts* e crescimento relativo; se é sobre avaliação institucional, métricas normalizadas e contextualizadas. Começar por indicadores disponíveis sem pergunta clara leva a resultados desarticulados.

O quinto passo é prever visualizações. A visualização não é etapa decorativa, mas parte da resposta à pergunta. Séries temporais respondem a evolução; mapas geográficos respondem a distribuição espacial; redes respondem a relações; mapas temáticos respondem a estrutura conceitual; diagramas de Sankey respondem a fluxos; tabelas comparativas respondem a diferenças entre grupos. Uma pergunta operacionalizável pode já indicar o tipo de visualização esperado. Isso facilita a comunicação e reduz a tentação de gerar gráficos sem função analítica.

O sexto passo é explicitar limitações. Toda pergunta cientométrica deve carregar seus limites desde o início. Bases de dados têm cobertura seletiva; citações variam por área; nomes exigem desambiguação; palavras-chave são inconsistentes; altmétricas medem atenção e não necessariamente impacto; mapas dependem de parâmetros; modelos de linguagem podem alucinar ou inferir indevidamente. *Prompts* devem solicitar limitações junto com perguntas, não apenas ao final do relatório.

Uma técnica útil é transformar o objetivo vago em matriz. As colunas podem ser: objetivo inicial, pergunta operacional, unidade de análise, base, período, indicador, método, visualização, decisão interpretativa e limitação. A IA pode preencher uma versão preliminar da matriz. Em seguida, o pesquisador revisa, elimina perguntas inviáveis, combina perguntas complementares e justifica escolhas. Essa matriz funciona como ponte entre *brainstorming* e protocolo.

Outra técnica é a decomposição por verbo. Objetivos vagos contêm verbos como compreender, analisar, mapear, avaliar, identificar e explorar. Cada verbo pode ser traduzido em operação. 'Compreender' pode virar descrever, comparar, explicar ou interpretar; 'mapear'

pode virar construir rede, mapa temático ou distribuição geográfica; 'avaliar' pode virar comparar indicadores contextualizados; 'identificar' pode virar ranquear, classificar, detectar ou selecionar; 'explorar' pode virar gerar hipóteses. O *prompt* deve pedir essa tradução verbal.

A operacionalização também exige perguntas negativas: o que não será respondido? Um estudo sobre produção não responderá qualidade; uma análise de citações não medirá impacto social; um mapa de termos não provará relação causal; uma rede de coautoria não capturará toda colaboração; um recorte em inglês não representará literatura local completa. A pergunta final deve ser acompanhada de cláusulas de exclusão. Essa prática protege o estudo contra usos indevidos e fortalece sua credibilidade.

A IA pode ser usada em três momentos da operacionalização. No momento divergente, o modelo gera alternativas de perguntas. No momento analítico, compara alternativas, aponta requisitos e limitações. No momento convergente, ajuda a redigir a pergunta final e o protocolo. Essa sequência é preferível a pedir uma única pergunta pronta. A cadeia de *prompts* permite que o pesquisador veja o espaço de possibilidades antes de escolher.

A qualidade do *prompt* determina a qualidade das alternativas. Um *prompt* fraco diria: 'crie perguntas bibliométricas sobre sustentabilidade'. Um *prompt* forte diria: 'a partir do tema sustentabilidade urbana, proponha perguntas cientométricas separadas por produção, impacto, colaboração, emergência temática e política científica; para cada pergunta, indique unidade de análise, fonte de dados, período recomendado, indicadores, visualização, limitações e tipo de conclusão possível'. O segundo *prompt* não busca apenas frases; busca desenho metodológico.

A operacionalização deve ser iterativa. A primeira versão da pergunta pode revelar dados indisponíveis; a busca piloto pode mostrar ruído terminológico; a análise inicial pode sugerir subtemas; a revisão por especialistas pode indicar lacunas. Em vez de tratar mudanças como falhas, o pesquisador deve documentá-las. Protocolos reprodutíveis não exigem imobilidade absoluta, mas transparência sobre decisões. A IA pode auxiliar mantendo um histórico de versões da pergunta e justificativas para alterações.

A validação da pergunta deve anteceder a análise completa. Recomenda-se submeter a pergunta a quatro testes: responsabilidade, relevância, coerência e responsabilidade. Responsabilidade pergunta se há dados suficientes. Relevância pergunta se a resposta contribui para conhecimento ou decisão. Coerência pergunta se unidade, indicador e método se alinham. Responsabilidade pergunta se a pergunta evita interpretações injustas, reducionistas ou eticamente frágeis. Um *prompt* de auditoria pode aplicar esses testes.

Quadro 6.3 - Exemplo de matriz de operacionalização

Objetivo inicial	Pergunta operacionalizável	Unidade de análise	Base de dados	Período
Mapear a literatura sobre IA na educação.	Como evoluiu a produção de artigos sobre IA na educação entre 2015 e 2025?	Artigos	• Web of Science • Scopus	2015-2025
Avaliar a produção de uma instituição.	Quais áreas do conhecimento concentram a produção de uma universidade?	Instituições e áreas	Scopus	2016-2025
Analisar impacto.	Quais documentos tiveram maior impacto de citação normalizado por área e ano?	Documentos	Web of Science Core Collection	2015-2024

Indicadores	Métodos	Visualizações	Limitações
• Nº de artigos por ano • Taxa de crescimento (documentos)	• Contagem anual • Normalização por área (opcional)	Gráfico de série temporal	• Cobertura da base • Idioma (inglês dominante) • Tipos documentais
• Nº de artigos por área • Índice de especialização • Índice de diversidade	• Classificação por área (ASJC/CWTS) • Indicadores de especialização	• Tabela e gráfico de barras • Mapa de calor	• Variações de afiliação • Cobertura por área • Políticas institucionais
• Impacto normalizado (NCPP/FNCPP) • Citações recebidas	• Normalização por área e ano • Janela de 3 a 5 anos	• Ranking e tabela • Distribuição de citações	• Janela temporal • Diferenças entre áreas • Autocitações

Fonte: Elaborado pelo autor (2026).

Em síntese, transformar objetivos vagos em perguntas operacionalizáveis é converter linguagem ampla em desenho de pesquisa. Essa transformação exige escolhas explícitas e renúncias conscientes. A IA é valiosa porque amplia alternativas, organiza critérios e revela riscos; é perigosa quando transforma ambiguidade em texto fluente sem esclarecimento metodológico. A pergunta operacionalizada é, portanto, o primeiro produto científico do estudo cientométrico.

6.8 Prompts do capítulo

Os *prompts* a seguir foram concebidos como instrumentos de formulação e validação metodológica. Eles devem ser adaptados ao tema, à área, à disponibilidade de dados e à finalidade do estudo. Em todos os casos, recomenda-se registrar a versão do modelo utilizado, a data de uso, os dados fornecidos, a saída obtida e as decisões humanas tomadas a partir da resposta.

Prompt 6.8.1 Prompt central - transformar tema em perguntas cientométricas

Atue como especialista em Cientometria.

A partir do tema [TEMA], proponha cinco perguntas de pesquisa cientométricas.

Para cada uma, indique: unidade de análise, base de dados recomendada, período temporal, indicadores possíveis, tipo de visualização e limitações metodológicas.

Separe as perguntas nas seguintes categorias: produção científica, impacto e influência, colaboração e redes, emergência de temas e políticas científicas ou avaliação institucional. Não invente dados empíricos.

Quando houver incerteza sobre a melhor fonte ou indicador, explicite a incerteza e proponha alternativas.

Este *prompt* é adequado para a fase divergente do desenho de pesquisa. Sua função não é decidir o estudo, mas criar um conjunto de alternativas comparáveis. O pesquisador deve avaliar quais perguntas são relevantes, respondíveis e proporcionais à finalidade do projeto. A presença obrigatória de unidade de análise, base, período, indicadores, visualização e limitações impede que a IA produza perguntas genéricas sem operacionalização.

Prompt 6.8.2 Prompt para decompor objetivo vago

> Receba o seguinte objetivo preliminar: [OBJETIVO].
>
> Transforme-o em um conjunto de perguntas cientométricas operacionalizáveis.
>
> Para cada pergunta, informe: fenômeno observado, unidade de análise, dados necessários, estratégia de busca preliminar, indicadores, método analítico, possível visualização, tipo de conclusão permitida e limites de interpretação.
>
> Ao final, indique quais perguntas são descritivas, comparativas, avaliativas, prospectivas ou estratégicas.

O *prompt* de decomposição é útil quando o estudo nasce de uma demanda institucional ou de uma intenção ampla. Ele força a tradução de verbos vagos, como mapear, avaliar ou compreender, em operações analíticas. Também ajuda a separar perguntas que podem ser respondidas com dados bibliográficos daquelas que exigem dados qualitativos, administrativos ou contextuais.

Prompt 6.8.3 Prompt para escolher pergunta principal

> Considere as perguntas cientométricas abaixo: [LISTA DE PERGUNTAS].
>
> Avalie cada uma segundo quatro critérios: responsabilidade com dados disponíveis, relevância científica ou institucional, coerência entre unidade de análise e indicador, e risco de interpretação indevida.
>
> Atribua uma avaliação qualitativa para cada critério, justifique a avaliação e recomende uma pergunta principal e até três perguntas secundárias.
>
> Não use rankings numéricos se os critérios forem incomparáveis; explique *trade-offs*.

A escolha da pergunta principal é uma decisão metodológica e estratégica. Este *prompt* ajuda a tornar explícitos os critérios de decisão, mas não substitui o julgamento do pesquisador. Em estudos avaliativos, recomenda-se acrescentar critérios de justiça, comparabilidade e alinhamento à missão institucional.

Prompt 6.8.4 Prompt para refinar pergunta de produção científica

> Refine a seguinte pergunta sobre produção científica: [PERGUNTA].
>
> Verifique se ela especifica tema, base de dados, período, tipo documental, unidade de contagem, regra de atribuição, idioma, área do conhecimento e finalidade analítica.
>
> Reescreva a pergunta em versão operacionalizada e liste as decisões metodológicas que ainda precisam ser tomadas.

Esse *prompt* atua como checklist. Ele é especialmente útil para evitar que a contagem de publicações seja apresentada sem escopo. A versão refinada deve permitir que outro pesquisador compreenda o que será contado e em que condições.

Prompt 6.8.5 Prompt para refinar pergunta de impacto e influência

> Refine a seguinte pergunta sobre impacto ou influência: [PERGUNTA].
>
> Diferencie impacto acadêmico, influência intelectual, visibilidade social, impacto tecnológico e uso em políticas públicas. Indique qual dimensão a pergunta realmente pretende observar, quais indicadores seriam coerentes, que normalizações são necessárias e quais conclusões não poderiam ser sustentadas apenas por esses indicadores.

O *prompt* previne um dos erros mais frequentes em cientometria: usar impacto como termo genérico. Ao obrigar a distinção entre dimensões, o modelo ajuda a delimitar o alcance da análise e a evitar afirmações sobre qualidade, relevância social ou causalidade sem evidência.

Prompt 6.8.6 Prompt para refinar pergunta de colaboração e redes

> Refine a seguinte pergunta sobre colaboração ou redes: [PERGUNTA].
>
> Informe quais serão os nós, quais serão as arestas, se a rede será ponderada ou não, qual será o período, que método de normalização pode ser usado, quais indicadores de rede são adequados e que interpretações devem ser evitadas.
>
> Diferencie rede de coautoria, cocitação, acoplamento bibliográfico e coocorrência de termos.

Esse *prompt* é recomendado antes da construção de mapas. Ele obriga o pesquisador a declarar a arquitetura da rede e, portanto, reduz o risco de interpretar visualizações como se fossem representações naturais da ciência. Toda rede é uma construção metodológica.

Prompt 6.8.7 Prompt para formular pergunta sobre emergência temática

> A partir do tema [TEMA] e do período [PERÍODO], proponha perguntas para detectar temas emergentes.
>
> Para cada pergunta, indique como a emergência será definida: crescimento de publicações, crescimento relativo de termos, *burst* de citações, surgimento de novos *clusters*, conexão interdisciplinar, presença em patentes, atenção social ou documentos de política.
>
> Indique também como diferenciar emergência real de moda terminológica ou ruído de indexação.

A detecção de temas emergentes exige critérios explícitos. Este *prompt* reduz a tendência de chamar qualquer crescimento de emergência. Ele também obriga a considerar sinais complementares, como consistência de *clusters*, diversidade de atores e persistência temporal.

Prompt 6.8.8 Prompt para perguntas de política científica e avaliação institucional

> Transforme a demanda institucional abaixo em perguntas cientométricas responsáveis: [DEMANDA].
>
> Para cada pergunta, explicite a missão ou contexto institucional, a unidade de comparação, os indicadores possíveis, as normalizações necessárias, os riscos de uso indevido e as evidências que devem complementar os dados bibliométricos.
>
> Evite formulações que produzam rankings simplistas ou comparações sem contexto.

O *prompt* é adequado a relatórios de gestão e avaliação. Sua função principal é proteger a análise contra reducionismo métrico. Ele torna explícito que indicadores devem apoiar julgamento qualificado, e não substituir avaliação contextual.

Prompt 6.8.9 Prompt de auditoria da pergunta cientométrica

> Audite a seguinte pergunta cientométrica: [PERGUNTA].
>
> Aplique os critérios de clareza, responsabilidade, relevância, coerência metodológica, adequação ética, comparabilidade, rastreabilidade e risco de interpretação indevida.
>
> Para cada critério, indique pontos fortes, fragilidades e recomendações de reformulação.
>
> Ao final, apresente uma versão revisada da pergunta e uma lista de limitações que devem aparecer no protocolo.

A auditoria deve ser aplicada antes da coleta definitiva. O objetivo é detectar falhas de desenho quando ainda é possível corrigi-las. Em estudos com IA, esse *prompt* também ajuda a documentar a contribuição do modelo e as decisões humanas subsequentes.

Prompt 6.8.10 Prompt para transformar pergunta em protocolo preliminar

> A partir da pergunta cientométrica [PERGUNTA], elabore um protocolo preliminar contendo: justificativa da pergunta, objetivo geral, objetivos específicos, unidade de análise, fontes de dados, estratégia de busca, critérios de inclusão e exclusão, período, indicadores, métodos de análise, visualizações, procedimentos de validação, riscos metodológicos, limitações e produtos esperados.
>
> Não inclua resultados. Separe claramente decisões tomadas de decisões pendentes.

Este *prompt* conecta a formulação da pergunta às etapas seguintes do estudo. Ele deve ser usado depois que a pergunta já foi escolhida e refinada. A distinção entre decisões tomadas e pendentes é essencial para evitar que o protocolo pareça mais maduro do que realmente está.

6.9 Exemplo aplicado

Considere o objetivo vago: 'mapear a pesquisa sobre inteligência artificial na educação superior'. A formulação inicial é ampla e não indica unidade de análise, período, base, tipo de IA, dimensão educacional ou finalidade do mapeamento. Um uso direto desse objetivo poderia produzir uma análise extensa, porém dispersa. Com apoio de *prompts*, o objetivo pode ser transformado em perguntas mais operacionais.

Matriz de transformação do objetivo vago

Pergunta de produção: como evoluiu a produção de artigos e trabalhos de congresso sobre inteligência artificial na educação superior entre 2015 e 2025 nas bases Scopus e Web of Science?

Pergunta de impacto: quais documentos sobre IA na educação superior apresentam maior impacto de citação normalizado por ano de publicação e que temas eles representam?

Pergunta de colaboração: como se estrutura a rede de coautoria institucional nesse tema e quais instituições atuam como pontes entre comunidades?

Pergunta de emergência temática: quais termos e *clusters* temáticos apresentam crescimento acelerado após 2020, especialmente em relação a tutores inteligentes, aprendizagem adaptativa, *analytics* e IA generativa?

Pergunta de política científica: em que medida a produção institucional sobre IA na educação superior se alinha a prioridades de inovação pedagógica e transformação digital definidas em documentos estratégicos?

A matriz mostra que o mesmo tema pode originar estudos diferentes. Se a finalidade for uma revisão de área, as perguntas de produção, impacto e emergência podem ser prioritárias. Se a finalidade for gestão universitária, as perguntas de colaboração e política institucional podem ser mais relevantes. Se a finalidade for prospecção, a emergência temática ganha centralidade. A IA ajuda a explicitar essas alternativas, mas a escolha deve ser guiada por finalidade, dados e responsabilidade interpretativa.

A partir da matriz, uma pergunta principal poderia ser: 'como evoluíram a produção, a colaboração institucional e os temas emergentes da pesquisa sobre inteligência artificial na educação superior, indexada na Scopus e Web of Science, entre 2015 e 2025?'. Essa pergunta combina três dimensões e ainda pode ser ampla. Uma versão mais restrita seria: 'quais temas emergentes associados à IA generativa na educação superior aparecem na literatura indexada entre 2020 e 2025 e como se distribuem por países, instituições e tipos de colaboração?'. A segunda versão é mais focalizada e melhor alinhada a um estudo sobre novidade recente.

O exemplo evidencia uma regra prática: perguntas cientométricas podem ser integradas, mas não devem acumular dimensões sem desenho correspondente. Quanto mais dimensões uma pergunta contém, maior a necessidade de subperguntas, métodos complementares e limites interpretativos. A IA pode sugerir combinações, mas também deve ser solicitada a apontar quando a pergunta está excessivamente ampla.

6.10 Diretrizes práticas para formulação de perguntas cientométricas com IA

Quadro 6.4 - Diretrizes práticas para a formulação de perguntas cientométricas com IA

Nº	Diretriz	Referência rápida
01	**Comece pela finalidade do estudo** Antes de escolher indicadores, defina se o objetivo é descrever uma área, comparar atores, apoiar uma decisão, avaliar uma política, identificar tendências ou construir uma revisão de literatura. A finalidade condiciona o desenho.	**Exemplos de finalidades** • Descrever a evolução de um campo • Comparar países ou instituições • Apoiar decisões estratégicas • Avaliar políticas e programas • Identificar tendências emergentes • Construir revisão de literatura
02	**Especifique o fenômeno observado** Produção, impacto, colaboração, circulação, visibilidade, emergência e avaliação são dimensões diferentes. Uma pergunta pode combiná-las, mas deve saber que as combina.	**Dimensões do fenômeno** Produção · Impacto · Colaboração · Circulação · Visibilidade · Emergência · Avaliação
03	**Declare a unidade de análise** Artigos, autores, periódicos, instituições, países, áreas e conceitos não são intercambiáveis. Cada unidade exige metadados e métodos próprios.	**Unidades de análise** Artigos · Autores · Periódicos · Instituições · Países · Áreas · Conceitos
04	**Justifique o período** O recorte temporal deve dialogar com a dinâmica do campo, a disponibilidade de dados e a finalidade do estudo. Não use períodos arbitrários e sem justificativa.	**Fatores para justificar o período** • Dinâmica do campo • Eventos marcantes • Políticas e programas • Tecnologias emergentes • Disponibilidade de dados
05	**Escolha a fonte de dados conforme a pergunta** A base mais ampla nem sempre é a mais adequada. Cobertura disciplinar, tipos documentais e qualidade dos metadados são decisivos. Entre as opções usuais estão Scopus, Web of Science, Dimensions, PubMed e SciELO.	**Critérios para escolher a fonte** • Cobertura temática e geográfica • Tipos documentais disponíveis • Qualidade e padronização dos metadados • Atualização e histórico • Acesso e permissões • Adequação à pergunta
06	**Associe cada pergunta a indicadores específicos** Indicadores não devem ser escolhidos por disponibilidade ou popularidade, mas pela capacidade de responder à pergunta.	**Exemplos de indicadores por dimensão** • Produção: nº de documentos, taxa de crescimento, distribuição por área, produtividade • Impacto: citações, impacto normalizado, índice h, documentos entre os 10% mais citados • Colaboração: coautoria, centralidade, internacionalização, instituições parceiras • Emergência: *burst* de termos, crescimento relativo, novos tópicos, *clusters* emergentes • Avaliação: normalização por área, benchmarking, participação relativa, alinhamento estratégico

Nº	Diretriz	Referência rápida
07	**Inclua limitações desde a formulação** Toda pergunta deve vir acompanhada dos possíveis vieses de cobertura, indexação, idioma, citação, autoria, afiliação e interpretação.	**Limitações comuns** Cobertura da base · Idioma · Comportamento de citação · Desambiguação de autoria · Afiliação institucional · Seleção e indexação · Interpretação
08	**Use a IA para ampliar e criticar alternativas, não** para decidir sozinho O modelo pode sugerir perguntas, comparar critérios e apontar riscos; a decisão metodológica continua sendo humana.	**Como a IA pode apoiar** • Sugerir variações de perguntas • Comparar critérios explícitos • Apontar riscos e limitações • Listar indicadores e métodos • Questionar suposições e vieses
09	**Documente prompts e versões** Em estudos assistidos por IA, os *prompts* fazem parte do caminho metodológico e devem ser registrados de modo rastreável.	**Registro mínimo recomendado** • *Prompt* utilizado • Modelo e versão • Data da execução • Dados fornecidos • Saída principal gerada • Decisão humana e alterações • Justificativas e observações
10	**Submeta a pergunta ao teste de responsabilidade** Pergunte se a análise pode ser usada de modo injusto, se compara objetos comparáveis e se comunica incertezas de forma explícita.	**Teste de responsabilidade** • Comparabilidade: os objetos e os períodos são comparáveis? A normalização é adequada? • Justiça e ética: a análise pode gerar injustiças, rankings simplistas ou usos indevidos? • Incertezas explícitas: as limitações, os vieses e as incertezas estão claramente comunicados? • Uso apropriado: a pergunta e a análise atendem à finalidade informada e ao público interessado?
Regra de ouro *Perguntas cientométricas bem formuladas geram análises específicas, respondíveis, transparentes e responsáveis. Elas transformam intenções amplas em investigações científicas sólidas e éticas.*		

Fonte: Elaborado pelo autor (2026).

A IA é uma parceira metodológica poderosa quando usada com critério, transparência e supervisão humana. Boas perguntas geram bons caminhos de pesquisa não substituem o julgamento do pesquisador.

6.11 Considerações finais

A formulação da pergunta cientométrica é o momento em que o estudo escolhe seu compromisso epistemológico. Ela define o que será considerado evidência, que dimensões da ciência serão observadas, que comparações serão aceitas e que conclusões poderão ser sustentadas. Em um ambiente de abundância de dados, ferramentas e modelos generativos, a pergunta torna-se ainda mais importante, porque impede que a análise se transforme em acumulação de gráficos, rankings e narrativas sem coerência.

Este capítulo mostrou que perguntas sobre produção científica, impacto e influência, colaboração e redes, emergência de temas, políticas científicas e avaliação institucional possuem lógicas próprias. Cada uma exige unidades de análise, fontes, indicadores e formas de interpretação distintas. A produção descreve atividade; o impacto observa reconhecimento ou circulação; a colaboração revela relações; a emergência combina tempo e novidade; a avaliação institucional conecta métricas a contextos decisórios. Confundir essas dimensões é uma das principais fontes de erro em estudos cientométricos.

A inteligência artificial, quando usada por meio de *prompts* bem estruturados, pode ampliar a capacidade de formular perguntas. Ela pode decompor objetivos vagos, sugerir alternativas, identificar requisitos de dados, antecipar riscos, propor matrizes de decisão e transformar perguntas em protocolos preliminares. Contudo, sua contribuição deve ser compreendida como apoio metodológico. Modelos generativos não conhecem, por si mesmos, a finalidade real do estudo, a qualidade dos dados disponíveis, as implicações políticas de uma avaliação ou as nuances disciplinares que condicionam a interpretação.

A noção central é que *prompts* não devem ser usados apenas para pedir respostas, mas para organizar escolhas. Um bom *prompt* cientométrico explicita contexto, papel, tarefa, dados, restrições, critérios e formato de saída. Na formulação de perguntas, isso significa solicitar que a IA indique unidade de análise, base, período, indicadores, visualizações, limitações e tipo de conclusão permitida. Quanto mais a pergunta é acompanhada de critérios, menor o risco de que a resposta seja genérica ou metodologicamente frágil.

O capítulo também enfatizou a responsabilidade. Perguntas cientométricas não são neutras. Elas selecionam dimensões da ciência, tornam algumas práticas visíveis e deixam outras de fora. Em avaliações institucionais, podem afetar reputações, recursos e estratégias. Em prospecções temáticas, podem orientar prioridades. Em diagnósticos de impacto, podem reforçar desigualdades de idioma, área e indexação. Por isso, formular perguntas é também formular limites éticos e políticos para o uso dos resultados.

A pergunta cientométrica bem construída não promete revelar a verdade total sobre a ciência. Ela delimita uma observação possível, fundamentada em dados específicos, métodos transparentes e interpretação cautelosa. Essa humildade metodológica é condição de rigor. Em vez de perguntar 'qual é a melhor área, instituição ou pesquisador?', a cientometria responsável pergunta 'que evidências, sob quais critérios e com quais limitações, ajudam a compreender determinado fenômeno científico?'.

Nos capítulos seguintes da trilha, a pergunta formulada aqui servirá como base para estratégias de busca, coleta de dados, pré-processamento, indicadores, redes, mapas temáticos, análise temporal, visualização, interpretação e redação de relatórios. Todas essas etapas dependem da clareza inicial. A IA pode acompanhar cada uma delas, mas a primeira tarefa continua sendo humana: decidir que pergunta merece ser feita e de que modo ela pode ser respondida com justiça, transparência e rigor científico.

Apêndice A.6 - Checklist de formulação da pergunta cientométrica

O checklist a seguir pode ser utilizado em oficinas, disciplinas de pós-graduação, laboratórios de bibliometria ou núcleos de inteligência científica. Sua finalidade é permitir que a equipe avalie a maturidade de uma pergunta antes de iniciar a coleta definitiva. O uso do checklist não elimina a necessidade de revisão por especialistas, mas cria uma camada de controle metodológico.

1. A pergunta identifica claramente o tema ou domínio científico? Uma pergunta deve evitar termos genéricos sem delimitação. Quando o tema é amplo, recomenda-se indicar subárea, aplicação, método, população, tecnologia ou problema específico.

2. A pergunta explicita o fenômeno cientométrico observado? É necessário saber se a pergunta trata de produção, impacto, colaboração, estrutura intelectual, circulação, visibilidade, emergência ou avaliação. Fenômenos diferentes não devem ser fundidos sem justificativa.

3. A pergunta define unidade de análise? A unidade deve ser compatível com os dados. Artigos, autores, instituições, países e conceitos exigem procedimentos distintos de limpeza e interpretação.

4. A pergunta define período temporal? O período deve ser justificável em termos de dinâmica científica, disponibilidade de dados, marcos históricos ou finalidade do relatório.

5. A pergunta indica fonte de dados? A escolha da base deve considerar cobertura, atualização, tipos documentais, exportação de metadados, transparência e adequação disciplinar.

6. A pergunta antecipa indicadores? Indicadores devem ser derivados da pergunta. Quando a pergunta é sobre crescimento, usam-se séries temporais; quando é sobre influência, usam-se citações e redes; quando é sobre colaboração, usam-se relações de coautoria ou afiliação.

7. A pergunta prevê limitações? A ausência de limitações na formulação inicial é sinal de fragilidade. Cobertura, idioma, indexação, normalização e viés de citação devem ser mencionados.

8. A pergunta é proporcional ao uso pretendido? Relatórios de gestão, artigos científicos, revisões de literatura e políticas públicas exigem níveis diferentes de validação, detalhamento e cautela.

9. A pergunta evita conclusões indevidas? Se os dados medem publicações, não se deve concluir qualidade total; se medem citações, não se deve concluir impacto social; se medem coautoria, não se deve concluir colaboração informal completa.

10. A pergunta pode ser comunicada a públicos não especialistas? Uma boa pergunta é tecnicamente precisa, mas compreensível. A clareza pública é parte da responsabilidade metodológica.

O checklist pode ser aplicado manualmente ou por meio de IA. Quando aplicado por IA, recomenda-se solicitar que o modelo apresente justificativas e marque incertezas. A equipe deve registrar as recomendações aceitas e rejeitadas, com justificativa. Esse registro é parte da rastreabilidade do estudo.

Seção complementar - modalidades de perguntas e decisões de desenho

A pergunta cientométrica pode assumir modalidades distintas, e cada uma produz um desenho de pesquisa próprio. A pergunta exploratória busca abrir um campo, reconhecer vocabulários, identificar atores e organizar hipóteses iniciais. Ela é comum quando o pesquisador entra em um domínio pouco conhecido ou quando uma instituição deseja compreender uma área emergente. A pergunta confirmatória, por sua vez, testa uma expectativa formulada anteriormente: por exemplo, se determinada política aumentou a colaboração internacional ou se um programa de financiamento ampliou a produção em uma agenda estratégica. Confundir essas modalidades pode comprometer a análise, pois estudos exploratórios toleram maior abertura interpretativa, enquanto estudos confirmatórios exigem critérios prévios mais rígidos.

Outra distinção relevante ocorre entre perguntas descritivas e perguntas explicativas. Perguntas descritivas perguntam o que ocorreu: quantos documentos foram publicados, quais

países participaram, que termos apareceram, que periódicos concentraram publicações. Perguntas explicativas perguntam por que ocorreu. A cientometria pode contribuir para explicações ao revelar padrões, associações e trajetórias, mas raramente explica sozinha. Explicações exigem conhecimento histórico, institucional, econômico, político e disciplinar. Por isso, uma pergunta cientométrica deve evitar prometer explicação causal quando dispõe apenas de indicadores bibliográficos. A IA deve ser instruída a marcar como hipótese toda tentativa de explicar padrões sem evidência adicional.

Perguntas comparativas devem declarar a base de comparação. Comparar países, instituições, programas ou áreas é uma operação sensível, pois diferenças de tamanho, missão, idioma, financiamento e práticas de publicação afetam indicadores. Uma comparação responsável pode usar normalização por área, contagem fracionada, indicadores relativos e contextualização qualitativa. Em vez de perguntar qual país lidera uma área, pode ser mais adequado perguntar como se distribuem produção, impacto normalizado e colaboração internacional entre países com presença relevante em determinado domínio. A pergunta deixa de produzir julgamento absoluto e passa a construir evidência comparativa.

Perguntas longitudinais exigem atenção à continuidade dos dados. Mudanças de indexação, fusões institucionais, alterações de nomes, criação de novas categorias e atualização de bases podem alterar séries históricas. O pesquisador deve perguntar se a variação observada representa mudança real na ciência ou mudança no sistema de registro. A IA pode auxiliar listando possíveis artefatos temporais, mas a verificação exige inspeção dos dados. Uma pergunta longitudinal robusta deve prever análise de sensibilidade, comparação de períodos e, quando possível, validação com eventos históricos do campo.

Perguntas de fronteira investigam relações entre áreas. Elas são particularmente relevantes para interdisciplinaridade, inovação e emergência temática. Podem perguntar como conceitos de uma área são apropriados por outra, quais documentos conectam comunidades, que periódicos funcionam como canais de circulação e que combinações terminológicas aparecem em novos *clusters*. Essas perguntas exigem cuidado porque fronteiras disciplinares são construções institucionais e bibliográficas. Categorias de bases de dados podem não refletir práticas reais de pesquisa. O uso de IA pode ajudar a reconhecer vocabulários transversais, mas a interpretação deve considerar história e sociologia das disciplinas.

Perguntas orientadas por problemas sociais diferem de perguntas orientadas por áreas acadêmicas. Um problema como mudança climática, segurança alimentar ou saúde mental atravessa disciplinas, políticas e setores. A estratégia de busca deve combinar termos científicos, aplicações, populações e tecnologias. A unidade de análise pode incluir artigos, relatórios, patentes e documentos de política. O pesquisador deve decidir se deseja mapear a literatura acadêmica sobre o problema, a resposta científica a uma agenda pública ou a circulação de conhecimento entre ciência e sociedade. Cada opção produz uma pergunta diferente.

Perguntas de lacuna de pesquisa devem ser formuladas com prudência. A ausência de publicações em uma base não significa ausência de conhecimento. Pode indicar literatura em outro idioma, produção local não indexada, conhecimento técnico não publicado, práticas comunitárias ou baixa cobertura da fonte. Uma pergunta mais cuidadosa não afirma diretamente que há lacuna, mas pergunta quais temas aparecem sub-representados no universo documental analisado, considerando os critérios de busca e cobertura. A IA deve ser instruída a não transformar baixa frequência em lacuna substantiva sem validação.

Perguntas sobre liderança científica também exigem precisão. Liderança pode significar volume de produção, centralidade em redes, coordenação de colaborações, influência por citações, pioneirismo temporal, capacidade institucional ou formulação conceitual. Um autor

ou instituição pode liderar em volume, mas não em influência; pode ser central em colaboração, mas periférico em inovação temática. A pergunta deve definir a dimensão de liderança antes de escolher indicadores. Rankings simples tendem a misturar dimensões e produzir interpretações sedutoras, mas frágeis.

Perguntas sobre circulação do conhecimento podem usar evidências variadas. Circulação pode ocorrer por citações acadêmicas, redes de coautoria, leitura, downloads, menções em mídias sociais, documentos de política, patentes, notícias e ensino. Cada forma de circulação tem temporalidade e significado próprios. A formulação deve evitar o termo circulação como sinônimo de impacto. O que circula? Entre quais atores? Por quais canais? Em que período? Com que evidências? A IA pode transformar essas perguntas em matriz de fontes, mas a seleção final deve refletir a finalidade do estudo.

Perguntas sobre visibilidade devem distinguir ser visto de ser usado. Um artigo pode receber muita atenção digital por controvérsia, curiosidade pública ou divulgação institucional, mas não ser incorporado à literatura. Outro pode ser muito usado por especialistas e pouco visível ao público geral. A pergunta deve indicar se visibilidade será entendida como atenção, presença em canais, acessibilidade, circulação em redes ou reconhecimento acadêmico. Essa distinção protege a análise contra a equivalência indevida entre popularidade e contribuição científica.

Perguntas sobre avaliação de pesquisadores individuais são as mais delicadas. Métricas individuais podem ser influenciadas por idade acadêmica, área, gênero, idioma, oportunidades de colaboração, tamanho de grupo e práticas de autoria. Além disso, a contribuição intelectual de um pesquisador não se reduz a publicações e citações. Em muitos contextos, é preferível formular perguntas agregadas por grupos, programas ou áreas, preservando contextualização e evitando julgamentos automatizados. Quando indivíduos forem analisados, os critérios devem ser transparentes, proporcionais e complementares por avaliação qualitativa.

Perguntas sobre periódicos devem diferenciar canal, qualidade editorial, audiência e impacto. O fator de impacto de periódico não deve ser usado como substituto da qualidade de um artigo. Uma pergunta responsável pode investigar quais periódicos concentram a literatura de um campo, quais atuam como canais interdisciplinares ou como a distribuição de periódicos mudou ao longo do tempo. Já perguntas que buscam inferir qualidade de pesquisadores a partir do prestígio médio dos periódicos exigem cautela e justificativa robusta.

Perguntas sobre conceitos devem reconhecer que conceitos não são apenas palavras. Um mesmo conceito pode aparecer por termos diferentes, e um mesmo termo pode ter sentidos distintos em áreas diferentes. A IA pode auxiliar na identificação de sinônimos, traduções e termos relacionados, mas a validação semântica é indispensável. Em estudos de coocorrência, recomenda-se combinar análise automática com leitura de amostras de documentos. A pergunta deve indicar se se pretende mapear termos, conceitos ou problemas; essa diferença altera o desenho.

Perguntas de prospecção tecnológica frequentemente combinam artigos e patentes. A relação entre ciência e tecnologia não é linear: uma área pode ter produção acadêmica intensa e pouca conversão tecnológica, ou o inverso. A pergunta deve definir se o objetivo é mapear base científica de uma tecnologia, identificar fronteiras de aplicação, observar convergência entre literatura e patentes ou detectar atores com capacidade translacional. A IA pode propor termos técnicos e classes de patentes, mas a consulta a especialistas e bases apropriadas é essencial.

Perguntas sobre internacionalização devem diferenciar participação internacional, colaboração internacional e dependência internacional. Ter artigos com autores de outros países não significa necessariamente liderança, autonomia ou equilíbrio colaborativo. A pergunta pode

incluir posição autoral, diversidade de parceiros, reciprocidade, redes regionais e evolução temporal. Também deve considerar políticas institucionais, financiamento e idiomas. Uma análise responsável evita celebrar qualquer coautoria internacional sem examinar sua estrutura.

Perguntas sobre interdisciplinaridade exigem indicadores compatíveis. É possível analisar diversidade de categorias de periódicos, variedade de referências, combinação de termos, composição de equipes ou circulação entre comunidades de citação. Cada indicador representa um aspecto da interdisciplinaridade. A pergunta deve declarar se busca diversidade cognitiva, integração metodológica, colaboração entre áreas ou impacto em múltiplos campos. A IA pode listar dimensões, mas não deve fundi-las sem critério.

Perguntas sobre qualidade metodológica da literatura geralmente ultrapassam a bibliometria clássica. Elas podem exigir leitura de artigos, extração de métodos, avaliação de desenho de pesquisa e análise de evidências. A IA pode apoiar a triagem textual e a sumarização, mas resultados devem ser validados por humanos, especialmente quando se avaliam métodos, riscos de viés ou força de evidência. A pergunta deve diferenciar análise cientométrica de revisão sistemática ou avaliação crítica de literatura.

Perguntas sobre alinhamento a agendas estratégicas devem tornar explícita a agenda usada como referência. Pode ser uma política nacional, um plano institucional, os Objetivos de Desenvolvimento Sustentável, uma agenda tecnológica ou uma prioridade de financiamento. A formulação deve indicar como os temas serão associados à agenda: por palavras-chave, classificação manual, modelo de tópicos, vocabulário controlado ou validação por especialistas. Sem essa etapa, o alinhamento pode se tornar interpretação subjetiva sem rastreabilidade.

Perguntas sobre mapas científicos precisam evitar a ideia de que o mapa é o próprio território. Um mapa de ciência é representação construída a partir de dados e escolhas analíticas. Ele revela padrões sob determinados parâmetros e oculta outros. A pergunta deve indicar o propósito do mapa: exploração, comunicação, comparação, identificação de *clusters* ou apoio à decisão. A IA pode ajudar a escrever legendas e interpretações, mas deve ser instruída a não atribuir aos mapas conclusões que dependem de validação externa.

Por fim, perguntas cientométricas devem ser revisadas em linguagem natural antes de serem convertidas em comandos técnicos. Uma boa pergunta deve poder ser entendida por especialistas do campo, analistas de dados e gestores. Se a pergunta só faz sentido como lista de indicadores, ainda não está madura. Se faz sentido apenas como demanda gerencial, mas não como desenho de pesquisa, também precisa de refinamento. A maturidade está na interseção entre relevância substantiva, viabilidade empírica e rigor metodológico.

Aplicação da pergunta cientométrica em ciclos de pesquisa assistidos por IA

Em um ciclo completo de pesquisa assistido por IA, a pergunta cientométrica funciona como eixo de coordenação entre atividades humanas e atividades automatizadas. No início, ela orienta a exploração conceitual: o pesquisador solicita ao modelo que identifique dimensões, vocabulários e riscos. Em seguida, ela orienta a busca: termos e *strings* são avaliados segundo sua capacidade de responder à pergunta. Depois, orienta a limpeza: nomes, instituições e termos são normalizados conforme a unidade de análise escolhida. Por fim, orienta a interpretação: resultados são julgados pelo que dizem sobre a pergunta, não pelo que parecem sugerir isoladamente.

Esse ciclo evidencia que a IA deve ser usada de modo situado. O mesmo modelo pode apoiar uma revisão de literatura, uma análise institucional ou um diagnóstico de temas emergentes, mas o comportamento desejado muda conforme a pergunta. Em uma pergunta

exploratória, valoriza-se amplitude e diversidade de alternativas. Em uma pergunta avaliativa, valoriza-se cautela, normalização e identificação de riscos. Em uma pergunta redacional, valoriza-se clareza e fidelidade aos resultados. O *prompt* deve indicar essa finalidade para evitar respostas inadequadas ao estágio do estudo.

A pergunta também permite avaliar a suficiência da evidência. Após a coleta piloto, o pesquisador pode perguntar se os dados recuperados realmente respondem ao problema formulado. Uma busca que retorna muitos documentos pode, ainda assim, ser inadequada se mistura sentidos distintos do tema. Uma busca que retorna poucos documentos pode ser adequada se o tema é emergente e o objetivo é prospecção. A suficiência não é apenas quantidade; é relação entre dados e pergunta. A IA pode ajudar a revisar títulos e resumos amostrais para identificar ruído e lacunas.

Durante a análise, a pergunta deve funcionar como critério de seleção de gráficos e tabelas. Relatórios cientométricos frequentemente sofrem de excesso visual: séries temporais, redes, nuvens de palavras, mapas geográficos e rankings são apresentados porque a ferramenta permite produzi-los, não porque respondem à pergunta. Em uma abordagem orientada por *prompts*, cada visualização deve ser justificada: que subpergunta responde, que dado utiliza, que interpretação permite e que limitação possui. Essa disciplina fortalece a coerência do capítulo de resultados.

A IA pode ser solicitada a confrontar resultados com a pergunta inicial. Um *prompt* de retorno pode pedir: 'dado este conjunto de achados, indique quais respondem diretamente à pergunta, quais respondem parcialmente, quais são apenas contextuais e quais extrapolam o escopo'. Esse procedimento reduz o risco de narrativas oportunistas, nas quais o relatório enfatiza achados interessantes, mas desconectados do objetivo. A interpretação cientométrica deve permanecer fiel ao contrato analítico estabelecido pela pergunta.

Outro uso importante é a revisão de linguagem. Perguntas e objetivos devem ser redigidos em termos que não antecipem conclusões. Uma pergunta como 'por que a instituição X perdeu relevância científica?' pressupõe perda antes de demonstrá-la. Uma formulação mais adequada seria: 'como evoluíram os indicadores de produção, colaboração e impacto da instituição X no período definido?'. A IA pode ser instruída a identificar pressupostos embutidos na linguagem e a propor versões neutras. Essa revisão é particularmente importante em estudos avaliativos e comparativos.

A pergunta deve também orientar o nível de granularidade. Uma análise por país pode ser suficiente para observar geografia global da produção, mas insuficiente para políticas institucionais. Uma análise por autor pode ser detalhada demais para um diagnóstico nacional. Uma análise por tema amplo pode ocultar subáreas críticas. A granularidade adequada é aquela que permite responder à pergunta sem produzir excesso de ruído ou perda de significado. A IA pode ajudar a propor escalas alternativas e seus custos interpretativos.

Em estudos com dados sensíveis, a formulação da pergunta deve ser acompanhada por critérios de proteção. Mesmo quando se usam dados bibliográficos públicos, determinadas análises podem produzir exposição indevida de indivíduos ou grupos pequenos. Perguntas sobre desempenho individual, colaboração de minorias, redes periféricas ou produtividade de equipes reduzidas devem considerar agregação, anonimização e finalidade legítima. A IA não deve ser usada para inferir atributos pessoais sensíveis a partir de nomes, imagens ou padrões indiretos. A responsabilidade ética antecede a análise técnica.

A pergunta cientométrica também é uma ferramenta pedagógica. Em disciplinas de pós-graduação, solicitar que estudantes transformem temas amplos em perguntas operacionais ajuda a compreender a diferença entre assunto, problema e método. O uso de *prompts* pode

tornar esse processo visível: os estudantes comparam alternativas geradas, criticam limitações e justificam escolhas. Assim, a IA não substitui a aprendizagem metodológica; ela cria material para discussão crítica. O professor pode avaliar não apenas a pergunta final, mas a qualidade do processo de refinamento.

Em ambientes institucionais, a pergunta serve como contrato de comunicação entre demandantes e analistas. Gestores podem desejar respostas rápidas, mas a equipe técnica deve explicitar o que pode ser respondido com dados cientométricos e o que exige outras evidências. Um bom protocolo registra a pergunta acordada, os indicadores aceitos, as limitações reconhecidas e os usos previstos. Quando esse contrato é inexistente, resultados podem ser reinterpretados para finalidades não previstas. A IA pode auxiliar redigindo termos de escopo e notas de cautela.

Ao final do ciclo, a pergunta deve reaparecer na conclusão. Uma conclusão cientométrica robusta não apenas lista achados; ela responde explicitamente à pergunta, indica o grau de suporte empírico, distingue resultados principais de evidências auxiliares e declara limites. Também pode sugerir perguntas futuras, derivadas de lacunas identificadas. A IA pode apoiar a redação dessa seção, desde que receba os resultados reais e seja instruída a não criar dados, citações ou implicações ausentes.

Portanto, a formulação da pergunta não é uma etapa isolada no início do estudo. Ela acompanha todo o ciclo de pesquisa, da exploração ao relatório. Em trabalhos assistidos por inteligência artificial, essa função coordenadora é ainda mais necessária, porque a facilidade de gerar textos, tabelas e interpretações pode afastar o pesquisador do problema original. A pergunta é o mecanismo que preserva foco, rigor e responsabilidade.

Nota ampliada sobre rastreabilidade, iteração e decisão metodológica

A rastreabilidade da pergunta cientométrica exige que cada mudança de escopo seja registrada em linguagem clara. Quando um termo de busca é removido, quando uma base é substituída, quando o período é alterado ou quando uma unidade de análise é refinada, a decisão modifica o universo observado. Em estudos com inteligência artificial, a iteração por *prompts* acelera essas mudanças e, por isso, deve ser documentada com atenção. A versão final da pergunta deve preservar a memória das alternativas descartadas, pois essas alternativas explicam o caminho metodológico e ajudam avaliadores externos a compreender por que o estudo chegou à configuração apresentada.

A rastreabilidade não deve ser entendida como burocracia suplementar. Ela é parte do próprio argumento científico. Quando um estudo afirma que analisou a produção sobre determinado tema, o leitor precisa saber como esse tema foi operacionalizado: quais termos foram usados, quais sinônimos foram incluídos, quais expressões foram excluídas, que idiomas foram considerados e que bases sustentaram a coleta. Em ambientes de IA, essa exigência se amplia, porque o modelo pode sugerir termos plausíveis que não correspondem a vocabulários consolidados. O pesquisador deve registrar se aceitou, modificou ou rejeitou tais sugestões.

Um procedimento recomendável é manter um diário de formulação da pergunta. Esse diário pode conter a versão inicial do objetivo, os *prompts* utilizados, as respostas obtidas, os critérios de decisão e a versão final adotada. Também pode registrar as razões para exclusões. Por exemplo, se a equipe decide não incluir *preprints*, deve explicar se essa decisão decorre de baixa consistência dos metadados, de foco em literatura revisada por pares ou de limitações práticas de coleta. Se decide incluir anais de congresso, deve justificar a relevância desse tipo documental para a área.

A iteração com IA deve alternar momentos de expansão e de contração. Na expansão, o modelo é solicitado a gerar alternativas de perguntas, unidades de análise e indicadores. Na contração, o pesquisador seleciona as alternativas mais coerentes e pede ao modelo que critique essas escolhas. A sequência é metodologicamente mais segura que pedir diretamente a pergunta final. Ela permite observar o espaço de possibilidades, identificar pressupostos e documentar o raciocínio que levou à delimitação do estudo.

A decisão metodológica também deve considerar o custo analítico. Uma pergunta pode ser interessante, mas exigir dados que a equipe não consegue obter ou validar. Outra pergunta pode ser respondível, mas pouco relevante para o problema institucional. Uma terceira pode ser relevante e respondível, mas eticamente sensível. A seleção da pergunta deve equilibrar mérito científico, viabilidade, responsabilidade e utilidade. A IA pode organizar essa comparação, mas não deve substituir a decisão da equipe, pois a decisão envolve valores, prioridades e compromissos institucionais.

Em avaliações institucionais, recomenda-se que a pergunta seja discutida com atores afetados pela análise antes da coleta final. Departamentos, programas, bibliotecas, pró-reitorias e grupos de pesquisa podem identificar dimensões não capturadas por bases bibliográficas. Esse diálogo não elimina a independência analítica, mas aumenta a pertinência e reduz o risco de interpretações inadequadas. A IA pode auxiliar na preparação de versões da pergunta para diferentes públicos, traduzindo termos técnicos sem alterar o escopo metodológico.

A formulação da pergunta deve produzir uma fronteira analítica clara. Tudo que está dentro da fronteira será observado segundo os métodos escolhidos; tudo que está fora deve ser reconhecido como limitação. Essa fronteira não diminui o valor do estudo. Ao contrário, aumenta sua confiabilidade. Estudos que prometem mapear uma área inteira sem declarar exclusões, critérios e limitações tendem a produzir conclusões amplas demais. A boa cientometria não é aquela que responde tudo, mas aquela que responde bem ao que se propôs a responder.

A in*corpora*ção de IA à formulação de perguntas cientométricas deve, portanto, ser acompanhada de uma ética da explicitação. O pesquisador explicita o que pediu ao modelo, o que recebeu como resposta, o que aceitou, o que rejeitou e por quê. Essa prática permite que leitores, revisores e gestores compreendam a contribuição da IA sem confundi-la com autoridade científica autônoma. O uso transparente de *prompts* fortalece a pesquisa, enquanto o uso invisível ou não documentado pode comprometer a confiança no estudo.

Por fim, a pergunta operacionalizada deve permanecer viva durante a pesquisa, mas não arbitrária. Se os dados piloto indicarem que o escopo é inviável, a pergunta pode ser revisada; se a análise revelar subcampos inesperados, perguntas secundárias podem ser acrescentadas; se uma base apresentar cobertura insuficiente, outra pode ser incorporada. Cada alteração, entretanto, deve ser registrada. A estabilidade do estudo não depende de nunca mudar, mas de mudar com justificativa.

Essa perspectiva torna a formulação da pergunta um processo de aprendizagem metodológica. O pesquisador aprende com a literatura, com os dados, com especialistas, com limitações das bases e com as respostas da IA. O resultado final é mais que uma frase interrogativa: é uma síntese de decisões que organiza todo o estudo. Em uma cientometria assistida por inteligência artificial, essa síntese deve ser clara, auditável e responsável.

Capítulo 7

Estratégias de busca com apoio de IA

7.1 Introdução

Este capítulo integra a trilha "Cientometria com Inteligência Artificial: *Prompts* para Analisar, Mapear e Interpretar a Ciência", cujo propósito é capacitar o leitor a conduzir estudos cientométricos desde a formulação da pergunta até a produção de relatórios, mapas, indicadores e diagnósticos científicos assistidos por IA. Após os capítulos dedicados aos fundamentos da cientometria, à aplicação de inteligência artificial, à engenharia de *prompts* e à formulação da pergunta de pesquisa, a etapa de busca assume posição decisiva. Em estudos cientométricos, a busca bibliográfica não é mera operação preliminar, mas a própria condição de possibilidade do *corpus*. O modo como se nomeia um fenômeno, se traduz um conceito, se escolhe uma base e se combina operadores booleanos determina quais documentos serão visíveis, quais ficarão ausentes e quais padrões poderão ser interpretados posteriormente.

A busca em cientometria difere de uma busca bibliográfica casual porque tem compromisso simultâneo com abrangência, coerência, auditabilidade e adequação à pergunta. Uma consulta feita para localizar um artigo específico pode tolerar aproximações intuitivas; uma estratégia usada para mapear a evolução de um campo precisa ser explícito, replicável e defensável. A inteligência artificial, especialmente modelos de linguagem orientados por *prompts*, pode apoiar esse processo ao expandir vocabulários, sugerir sinônimos, comparar terminologias, simular vieses de busca, revisar *strings* e produzir documentação metodológica. No entanto, a mesma tecnologia pode gerar termos inexistentes, autores fictícios, traduções inadequadas, descritores não registrados ou falsas equivalências conceituais. Por isso, este capítulo não apresenta a IA como substituta da competência informacional, mas como instrumento auxiliar submetido a validação humana, verificação em bases reais e documentação rigorosa.

A literatura sobre busca sistemática (BRAMER *et al.*, 2018; RETHLEFSEN *et al.*, 2021) enfatiza que a construção de estratégias de recuperação é atividade complexa, iterativa e sujeita a erros, especialmente quando há múltiplos conceitos, sintaxes distintas entre bases e necessidade de equilibrar sensibilidade e precisão. Bramer *et al.* (2018) propuseram uma abordagem sistemática para desenvolver buscas completas e eficientes, destacando a importância de combinar termos livres, vocabulários controlados, sintaxe específica da base e tradução entre interfaces. Rethlefsen *et al.* (2021), ao apresentarem o *PRISMA-S*, reforçam que a descrição da busca é requisito de transparência e reprodutibilidade em revisões e sínteses de evidências. Embora a cientometria nem sempre tenha os mesmos objetivos de uma revisão sistemática, ela se beneficia fortemente desses princípios quando precisa justificar a composição do conjunto documental analisado.

No contexto da IA generativa, a questão torna-se ainda mais delicada. Estudos recentes mostram que modelos de linguagem podem gerar consultas booleanas úteis como ponto de partida, mas ainda carecem de sensibilidade, precisão e confiabilidade suficientes para uso autônomo (ADAM *et al.*, 2024). Adam *et al.* (2024) avaliaram um modelo de linguagem treinado para gerar buscas em PubMed e concluíram que as consultas produzidas poderiam ser úteis para escopo inicial e ensino, mas não deveriam ser utilizadas sem escrutínio de bibliotecários e especialistas. Essa conclusão é particularmente relevante para a cientometria com *prompts*: a IA deve produzir hipóteses de busca, não protocolos definitivos. A validação

continua dependente de especialistas, testes em bases, comparação com conjuntos-semente e registro transparente das decisões.

O capítulo está organizado em seis seções. A primeira discute palavras-chave, descritores e termos emergentes. A segunda examina estratégias booleanas para bases bibliográficas. A terceira aborda a tradução terminológica entre áreas do conhecimento. A quarta diferencia busca sensível e busca específica. A quinta propõe procedimentos de teste, refinamento e documentação. A sexta discute como evitar que a IA invente termos, autores ou bases inexistentes. Ao final, são apresentados *prompts* aplicados para expandir vocabulário de busca, criar *strings* booleanas, revisar ruído e lacunas e comparar estratégias em Scopus, Web of Science, PubMed, Dimensions e OpenAlex. O objetivo central é formar uma postura metodológica: usar IA para ampliar a inteligência do processo, sem reduzir a busca a uma sequência opaca de comandos.

7.2 Palavras-chave, descritores e termos emergentes

A escolha de palavras-chave é o primeiro ato interpretativo de uma busca cientométrica. Antes de qualquer operador booleano, antes da exportação de registros e antes da construção de mapas, há uma decisão de nomeação: como denominar o fenômeno que se deseja estudar? Em áreas consolidadas, essa nomeação pode parecer simples, pois existe vocabulário estável, indexação relativamente padronizada e tradição terminológica. Em áreas emergentes, interdisciplinares ou tecnicamente fragmentadas, a nomeação é um problema de pesquisa. O mesmo objeto pode aparecer com denominações distintas em comunidades diferentes; um termo pode ser usado como conceito em uma área e como técnica em outra; uma expressão pode ser recente demais para ter descritor controlado; e uma sigla pode ter significados divergentes conforme o campo disciplinar.

Em estudos cientométricos, palavras-chave funcionam como portas de entrada para a literatura, mas também como filtros de visibilidade. Uma busca sobre "inteligência artificial generativa", por exemplo, pode recuperar documentos indexados com expressões como "*generative AI*", "*large language models*", "*foundation models*", "*text-to-text generation*", "ChatGPT", "*prompt engineering*" e "*natural language generation*". Dependendo do período temporal e da base utilizada, alguns desses termos serão predominantes; outros aparecerão apenas em títulos recentes; outros existirão como palavras-chave dos autores, mas ainda não como descritores controlados. Assim, a expansão terminológica não é simples aumento de lista. Ela requer compreensão histórica, disciplinar e semântica do objeto.

A IA pode apoiar essa etapa ao sugerir termos relacionados, sinônimos, hiperônimos, hipônimos, siglas, traduções e variantes ortográficas. Contudo, essa contribuição deve ser tratada como produção de candidatos, e não como validação. Um modelo de linguagem pode sugerir um termo plausível, mas inexistente na literatura indexada; pode traduzir uma expressão técnica de forma literal, mas não usada pela comunidade; pode propor acrônimos ambíguos; ou pode confundir termos próximos que representam tradições metodológicas diferentes. Por isso, cada candidato sugerido pela IA deve ser verificado em pelo menos uma base bibliográfica, em vocabulários controlados quando existirem, e em amostras de registros recuperados.

Convém distinguir palavras-chave livres, descritores controlados e termos emergentes. Palavras-chave livres são expressões utilizadas em títulos, resumos, palavras-chave dos autores ou metadados textuais. Elas captam linguagem viva, variações terminológicas, neologismos e marcas discursivas de comunidades específicas. Descritores controlados pertencem a tesauros, ontologias ou vocabulários normalizados, como MeSH, DeCS ou Emtree, e permitem recuperar documentos indexados sob uma categoria conceitual mesmo quando o

texto usa variações. Termos emergentes são expressões que começam a ganhar densidade na literatura, mas ainda não foram estabilizadas pela indexação formal. Em cientometria, a combinação desses três tipos de termo é frequentemente necessária, pois a análise da ciência deve capturar tanto a estabilidade conceitual quanto a inovação terminológica.

A tensão entre termo livre e descritor controlado é metodologicamente produtiva. Termos livres são sensíveis à novidade; descritores são consistentes e institucionalizados. Uma busca exclusivamente baseada em descritores pode perder literatura recente, especialmente em campos de rápida mudança. Uma busca exclusivamente baseada em termos livres pode aumentar ruído, recuperar usos metafóricos ou deixar de capturar documentos indexados sob vocabulários técnicos. A IA pode auxiliar comparando essas alternativas: primeiro, listando termos livres prováveis; segundo, verificando quais possuem equivalentes controlados; terceiro, distinguindo termos que devem ser buscados em título e resumo daqueles que devem ser buscados em campos de descritores; quarto, sugerindo testes empíricos para avaliar a contribuição de cada grupo de termos.

Nos estudos cientométricos, a emergência terminológica pode ser objeto de análise, não apenas etapa de busca. Quando uma expressão passa a aparecer com frequência crescente em títulos e palavras-chave, isso pode indicar consolidação de um tema, difusão de uma técnica, mudança discursiva ou reconfiguração de campo. Entretanto, a detecção dessa emergência depende de estratégia de busca suficientemente ampla para não confundir ausência de termo com ausência de literatura. Por exemplo, antes da popularização de "*large language models*", muitos trabalhos relevantes estavam indexados sob "*language models*", "*pre-trained language models*", "*neural language models*" ou "*transformers*". Um estudo que busca apenas o termo mais recente poderá produzir falsa narrativa de surgimento abrupto.

A construção de vocabulário deve começar por um conjunto-semente. Esse conjunto pode ser formado por artigos clássicos, revisões reconhecidas, documentos de referência, autores centrais, conferências relevantes ou registros já conhecidos pelo pesquisador. A IA pode receber títulos, resumos e palavras-chave desse conjunto para extrair candidatos terminológicos, mas a extração precisa ser acompanhada de critérios. O *prompt* deve solicitar separação entre termos centrais, termos periféricos, nomes de métodos, nomes de aplicações, nomes de populações, nomes de bases, siglas e expressões ambíguas. Sem essa separação, a IA tende a misturar níveis analíticos: pode tratar uma técnica como tema, uma área como método, uma instituição como conceito ou um software como objeto de pesquisa.

Uma prática recomendável consiste em construir uma matriz terminológica. Nessa matriz, cada linha representa um conceito da pergunta de pesquisa, e cada coluna registra termos livres, descritores, siglas, traduções, variações ortográficas, termos excluídos, exemplos de documentos recuperados e observações. Em estudos sobre cientometria com IA, por exemplo, uma matriz poderia distinguir os conceitos "cientometria", "bibliometria", "mapeamento científico", "inteligência artificial", "modelos de linguagem" e "*prompts*". Essa distinção evita que a busca se torne uma lista indiferenciada de termos, na qual operadores booleanos combinam elementos sem lógica conceitual clara.

A IA pode ser particularmente útil na identificação de termos fronteiriços. Termos fronteiriços são aqueles que conectam comunidades disciplinares: *"knowledge mapping", "science mapping", "research fronts", "topic modeling", "information retrieval", "evidence synthesis", "research evaluation", "scholarly communication", "AI-assisted reviewing"*. Um estudo cientométrico pode precisar decidir se esses termos pertencem ao escopo central, ao contexto periférico ou a buscas complementares. A decisão não deve ser tomada apenas pela semelhança lexical, mas pela pergunta de pesquisa. Se o objetivo é mapear a literatura sobre estratégias de busca com IA em estudos cientométricos, "*systematic review search strategy*"

pode ser relevante metodologicamente, embora não seja cientometria em sentido estrito. Se o objetivo é medir produção em cientometria, o mesmo termo talvez seja periférico.

A expansão de vocabulário também exige atenção à temporalidade. Termos não nascem todos no mesmo momento. Alguns designam tradições antigas; outros são atualizações de técnicas anteriores; outros surgem como nomes comerciais ou nomes de modelos. Em uma busca longitudinal, o pesquisador deve evitar anacronismo terminológico: usar apenas o termo atual para recuperar literatura passada pode subestimar períodos anteriores; usar termos antigos para mapear literatura recente pode incluir campos que se afastaram do objeto. A IA pode apoiar a construção de uma linha do tempo terminológica, mas essa linha do tempo deve ser confirmada por frequências reais em bases, análise de títulos e leitura de documentos-chave.

Outra questão é a variação linguística. Estudos cientométricos globais tendem a privilegiar a literatura em inglês, não apenas porque as principais bases indexam majoritariamente periódicos internacionais, mas porque estratégias de busca são frequentemente formuladas apenas nessa língua. Em contextos latino-americanos, lusófonos ou ibero-americanos, a inclusão de termos em português e espanhol pode alterar significativamente o *corpus*, sobretudo quando a base inclui SciELO, Redalyc, Latindex, repositórios institucionais ou registros multilíngues. A IA pode sugerir traduções, mas deve ser orientada a distinguir tradução literal, termo técnico usual e equivalente disciplinar. *"Science mapping"*, por exemplo, pode ser traduzido como "mapeamento científico", "mapas da ciência" ou "mapeamento da ciência", expressões que não têm o mesmo uso em todos os contextos.

A busca por termos emergentes deve combinar três movimentos: observação de literatura recente, mineração de vocabulário e validação por especialistas. A literatura recente pode ser localizada por buscas amplas em bases abertas, por alertas de periódicos, por conferências-chave ou por análise de citações recentes. A mineração de vocabulário pode usar ferramentas de extração de n-gramas, análise de palavras-chave, redes de coocorrência e apoio de IA generativa. A validação por especialistas distingue modismos lexicais de conceitos substantivos. Nem todo termo que cresce em frequência representa um novo campo; algumas expressões refletem mudanças editoriais, influência de plataformas, campanhas de financiamento ou tendências de comunicação científica.

Em termos práticos, recomenda-se que a expansão terminológica assistida por IA seja realizada em ciclos. No primeiro ciclo, a IA gera uma lista ampla de candidatos a partir da pergunta e de documentos-semente. No segundo, o pesquisador classifica os candidatos por conceito, tipo e relevância. No terceiro, os candidatos são testados em bases reais, observando número de resultados, títulos recuperados e presença de documentos esperados. No quarto, a IA pode ser usada novamente para analisar os resultados dos testes, apontando termos que produziram ruído excessivo, termos que parecem úteis e lacunas possíveis. No quinto, a lista final é documentada, incluindo termos rejeitados e razões de rejeição. Esse registro é fundamental para transparência, pois mostra que a estratégia de busca foi construída, não simplesmente intuída.

A lógica de documentação é especialmente importante quando a IA participa do processo. Sem registro do *prompt*, da resposta, dos termos aceitos, dos termos recusados e das evidências usadas na decisão, a busca torna-se difícil de auditar. Em estudos cientométricos, isso afeta a credibilidade dos indicadores subsequentes. Um mapa de coocorrência, uma rede de cocitação ou uma curva temporal podem parecer sofisticados, mas serão metodologicamente frágeis se o *corpus* tiver sido produzido por uma busca opaca. A qualidade da visualização não compensa a fragilidade da recuperação documental. Por isso, a primeira recomendação

do capítulo é tratar a matriz terminológica como artefato científico: ela deve ser preservada, versionada e, quando possível, disponibilizada como material suplementar.

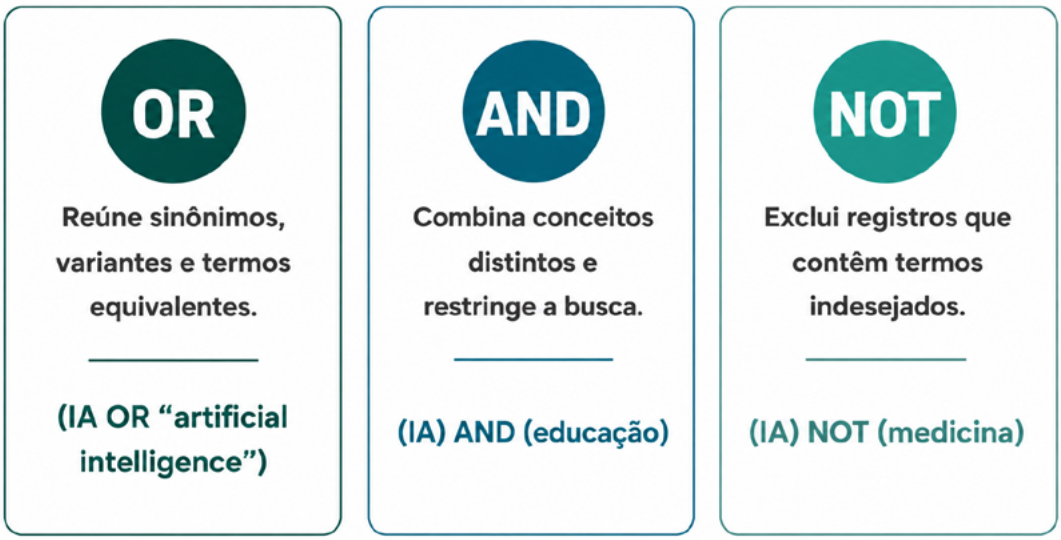
O pesquisador também deve diferenciar termos de inclusão e termos de exclusão. Termos de inclusão ampliam o escopo; termos de exclusão removem usos indesejados. A IA pode sugerir exclusões, mas o uso do operador NOT requer prudência. Excluir "*education*" de uma busca sobre "*AI prompts*" pode eliminar ruído de ensino, mas também remover estudos relevantes de formação científica. Excluir "*clinical*" de uma busca sobre "*search strategy*" pode retirar documentos de saúde que contêm métodos úteis de busca. Em cientometria, é frequentemente preferível recuperar ruído e filtrar posteriormente do que excluir inadvertidamente segmentos inteiros de literatura. O operador NOT deve ser documentado e justificado com exemplos de ruído real.

Em síntese, palavras-chave, descritores e termos emergentes não são apenas insumos técnicos. Eles expressam uma teoria implícita sobre o objeto analisado. Quando o pesquisador escolhe termos, ele define fronteiras de campo, relações disciplinares, temporalidades e objetos de visibilidade. A IA amplia a capacidade de explorar essas possibilidades, mas não elimina a necessidade de julgamento conceitual. A pergunta decisiva não é "quais termos a IA sugeriu?", mas "quais termos são conceitualmente justificáveis, empiricamente verificáveis e adequados à base, ao período e à unidade de análise do estudo?".

7.3 Estratégias booleanas para bases bibliográficas

A estratégia booleana é a forma operacional pela qual a matriz conceitual se transforma em consulta executável. Ela combina termos por meio de operadores lógicos, campos, parênteses, aspas, truncamentos, proximidade e filtros. Em sua forma mais simples, o operador OR reúne sinônimos ou variantes de um mesmo conceito; AND combina conceitos distintos; NOT exclui registros que contenham termos indesejados. Contudo, na prática, uma estratégia booleana de qualidade exige muito mais do que conhecer três operadores. Ela exige compreender a arquitetura de cada base, a sintaxe de seus campos, o comportamento do mecanismo de busca, a existência de vocabulário controlado, os limites de exportação e a forma como a base interpreta expressões, plural, acentos, hifens e caracteres especiais.

Figura 7.1 - Operadores booleanos



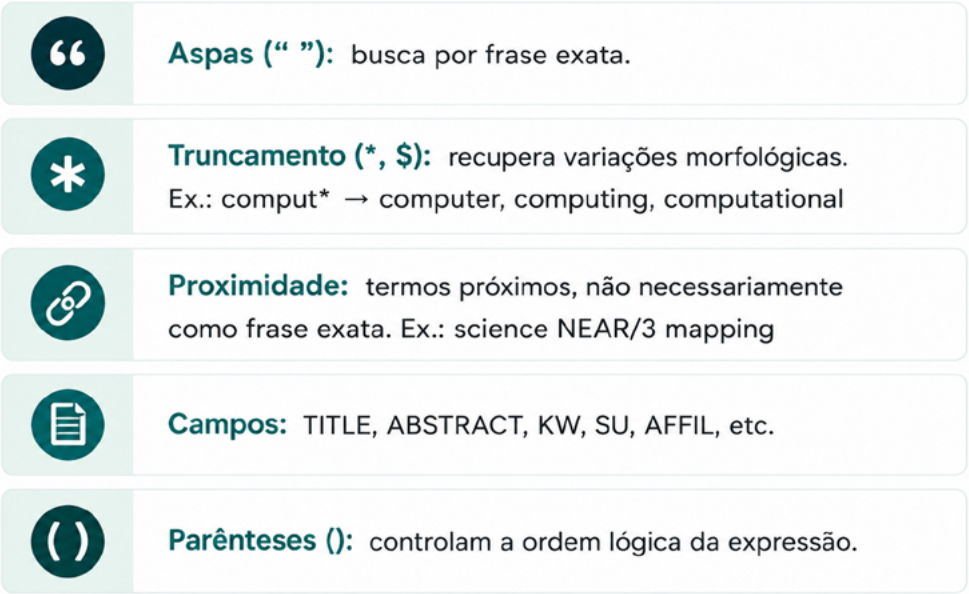
Fonte: Imagem gerada por IA, ChatGPT (2026).

Em cientometria, a *string* booleana deve ser derivada da pergunta e da unidade de análise. Se a pergunta busca mapear a produção sobre "uso de *prompts* em estudos cientométricos", a *string* precisa combinar um bloco de termos sobre cientometria, um bloco sobre IA/modelos de linguagem e um bloco sobre *prompts*. Se a pergunta é mais ampla, por exemplo "como a IA tem sido usada na análise da ciência?", talvez o bloco de *prompts* não deva ser obrigatório, pois excluiria trabalhos de classificação automática, extração de tópicos e mineração de texto. Essa diferença ilustra um princípio central: operadores booleanos não são neutros; eles materializam escolhas conceituais. Um AND adicional pode aumentar precisão, mas reduzir sensibilidade; um OR adicional pode ampliar sensibilidade, mas aumentar ruído.

A anatomia de uma *string* robusta começa com blocos conceituais. Cada bloco deve agrupar termos equivalentes ou próximos por OR. Em seguida, os blocos são combinados por AND. Por exemplo: (*scientometric** OR *bibliometric** OR *"science mapping"* OR "*research evaluation*") AND (*"artificial intelligence"* OR "*machine learning*" OR "*large language model**" OR "*generative AI*") AND (*prompt** OR "*prompt engineering*"). Essa forma é apenas ilustrativa. Ela precisa ser adaptada aos campos, aos operadores de proximidade e ao vocabulário da base. Em Scopus, pode-se usar TITLE-ABS-KEY; em Web of Science, TS; em PubMed, [Title/Abstract] e MeSH; em Dimensions, a interface tem comportamento próprio; em OpenAlex, filtros e *endpoints* diferem de uma *string* textual clássica.

A IA pode gerar uma *string* inicial, mas deve ser instruída a explicitar a lógica de cada bloco. Um erro comum em *prompts* genéricos é pedir "crie uma *string* de busca sobre o tema X". A resposta tende a misturar campos, operadores e termos sem explicar a relação entre eles. Um *prompt* metodologicamente adequado deve solicitar: definição dos conceitos centrais; lista de termos por conceito; justificativa de inclusão; identificação de termos ambíguos; proposta de *string* para cada base; observação de sintaxe; e advertência sobre validação. A *string* gerada deve ser vista como rascunho. O pesquisador precisa executá-la, verificar contagens, ler títulos, identificar ruído e revisar.

Figura 7.2 - Componentes sintáticos



Fonte: Imagem gerada por IA, ChatGPT (2026).

A adaptação entre bases é uma das tarefas mais propensas a erro. Web of Science e Scopus usam sintaxes diferentes para campos de título, resumo e palavras-chave. PubMed combina automaticamente termos com MeSH em algumas situações, mas o uso de tags específicas altera o comportamento da busca. Dimensions e OpenAlex incorporam outros modelos de dados, com filtros, APIs e categorias que nem sempre equivalem aos campos tradicionais. Google Scholar tem limitações de reprodutibilidade e exportação. A IA pode auxiliar traduzindo uma estratégia, mas a tradução deve ser testada na interface real, pois pequenos detalhes sintáticos produzem resultados diferentes; no caso do PubMed, inclusive recursos de relevância como o Best Match dependem da infraestrutura própria da base (FIORINI *et al.*, 2018). Um parêntese ausente, aspas mal posicionadas ou truncamento não suportado pode modificar profundamente o *corpus*.

O uso de aspas merece atenção especial. Aspas forçam busca por frase exata, o que aumenta precisão quando a expressão é estável, como *"science mapping"* ou "*large language model*". Entretanto, a busca por frase exata pode perder variações relevantes, como "*mapping of science", "scientific mapping", "large-scale language model*" ou expressões com hifenização. Truncamentos, por sua vez, ampliam variações morfológicas, mas podem recuperar termos inesperados. *"comput*"* recupera *compute*, *computer, computational e computing,* mas também pode recuperar formas pouco relevantes dependendo da base. Em português, truncamentos podem ser afetados por plural, gênero e acentos. A IA pode sugerir truncamentos, mas o pesquisador deve testá-los e observar exemplos de registros recuperados.

Operadores de proximidade são úteis quando termos precisam ocorrer próximos, mas não necessariamente como frase exata. Em algumas bases, NEAR, W/n, PRE/n ou ADJ/n permitem recuperar expressões em que as palavras aparecem a certa distância. Uma busca por *science NEAR/3 mapping* pode recuperar "*mapping of science*" e "*science knowledge mapping*", dependendo da sintaxe. Porém, cada base implementa proximidade de modo distinto, e algumas não a suportam. Para estudos cientométricos, proximidade pode ser relevante quando termos isolados são ambíguos. "*mapping*" isolado é amplo; *"science mapping"* é específico; "*mapping scientific collaboration*" pode estar dentro do escopo. O uso inteligente de proximidade pode reduzir ruído sem impor frase exata excessivamente rígida.

A definição de campos de busca é outro componente crítico. Buscar em "título" aumenta precisão, mas reduz sensibilidade. Buscar em "título, resumo e palavras-chave" amplia a recuperação e costuma ser adequado para mapeamentos temáticos. Buscar em texto completo, quando disponível, pode recuperar muitos documentos que apenas mencionam o termo de passagem, aumentando ruído. Em cientometria, a escolha do campo deve ser justificada pela pergunta. Para mapear uma área em que o termo é central, título, resumo e palavras-chave podem bastar. Para mapear a adoção de uma técnica que aparece nos métodos, o resumo pode ser insuficiente. Quando bases oferecem texto completo, é tentador ampliar a busca, mas a comparabilidade com bases que não oferecem esse campo pode ser prejudicada.

A busca por autores, instituições ou periódicos exige outra lógica. Quando a unidade de análise é autor, o problema de desambiguação é central. Nomes podem variar, autores diferentes podem compartilhar iniciais e identificadores como ORCID nem sempre estão completos. Quando a unidade é instituição, mudanças de nome, fusões, traduções e afiliações múltiplas complicam a busca. Quando a unidade é periódica, mudanças de título, ISSN impresso e eletrônico, indexação irregular e escopo editorial devem ser considerados. A IA pode ajudar a listar variantes de nomes institucionais ou periódicos, mas essas variantes precisam ser verificadas em registros reais e, quando possível, em identificadores persistentes.

Filtros temporais e documentais também devem ser tratados com cautela. Limitar a busca a artigos pode excluir anais de conferências, relevantes em áreas de computação

e engenharia. Incluir todos os tipos de documento pode trazer editoriais, cartas, erratas e materiais não comparáveis. Limitar idiomas pode reduzir ruído, mas introduzir viés linguístico. Limitar períodos pode ser necessário para análise temporal, mas precisa refletir a história do tema. A IA pode sugerir filtros, mas o pesquisador deve explicitar sua justificativa. Em estudos sobre IA generativa, por exemplo, uma janela a partir de 2018 pode ser defensável por causa da consolidação de modelos de linguagem pré-treinados; uma janela a partir de 2022 pode capturar o impacto de ferramentas conversacionais, mas perder antecedentes técnicos.

A construção de *strings* em bases bibliométricas deve considerar também a exportação. Uma estratégia que recupera milhões de registros pode ser conceitualmente ampla, mas operacionalmente inviável. Scopus, Web of Science e Dimensions têm limites de exportação, formatos diferentes e campos com qualidade variável. OpenAlex permite acesso em larga escala por API, mas exige cuidado com paginação, filtros e normalização. A IA pode ajudar a planejar a coleta, propondo divisão por anos, áreas, tipos documentais ou consultas segmentadas, mas essas divisões podem introduzir duplicatas ou lacunas. Por isso, a estratégia booleana deve ser pensada junto com o plano de coleta e deduplicação.

Um aspecto frequentemente negligenciado é a legibilidade da *string*. Estratégias muito longas podem ser necessárias, mas se tornam difíceis de auditar. Recomenda-se organizar a *string* em blocos comentados durante a construção, ainda que a base exija uma versão em linha única para execução. O pesquisador pode manter uma versão documentada, com comentários sobre cada bloco, e uma versão executável, adaptada à sintaxe da base. A IA pode transformar a versão documentada em versão executável, mas a correspondência deve ser conferida. Esse procedimento reduz erros de parênteses, repetições, operadores ausentes e inconsistências entre bases.

A estratégia booleana não deve ser confundida com estratégia de busca completa. Ela é componente central, mas não único. Uma busca completa pode incluir identificação de documentos-semente, busca direta por autores, análise de referências citadas, busca por documentos que citam os estudos centrais, consulta a bases complementares e rastreamento de termos emergentes. Em cientometria, essas etapas podem ter dupla função: melhorar a recuperação e produzir evidências sobre a estrutura do campo. A IA pode sugerir formas de busca complementar, mas cada técnica deve ser registrada separadamente para que o *corpus* final seja interpretável.

Um checklist pode ajudar a revisar uma *string* e garantir que os principais conceitos estão contemplados.

Quadro 7.1 - Checklist de revisão da estratégia

CHECKLIST DE REVISÃO DA ESTRATÉGIA *(Use antes de executar a busca definitiva)*
☐ Todos os conceitos centrais estão representados?
☐ Há termos excessivamente amplos ou redundantes?
☐ Os operadores e parênteses estão corretos?
☐ A sintaxe corresponde à base selecionada?
☐ A *string* recupera documentos-semente conhecidos?
☐ O nível de ruído é aceitável para o objetivo?

CHECKLIST DE REVISÃO DA ESTRATÉGIA *(Use antes de executar a busca definitiva)*
☐ Os filtros e campos são justificáveis?
☐ A busca pode ser reproduzida e documentada?

Registro da revisão

Base e versão da string	Responsável	Data

Fonte: Elaborado pelo autor (2026).

Essas perguntas podem ser convertidas em *prompt* de auditoria. Entretanto, a IA só poderá revisar adequadamente se receber informações suficientes: pergunta de pesquisa, base, campo, período, *string*, objetivos de sensibilidade ou precisão, lista de documentos esperados e exemplos de ruído.

Um procedimento útil é criar três versões da *string*: ampla, intermediária e restrita. A versão ampla privilegia sensibilidade e é usada para explorar vocabulário e identificar documentos-semente adicionais. A versão intermediária busca equilíbrio entre abrangência e pertinência, sendo candidata ao estudo principal. A versão restrita prioriza precisão e pode ser usada para análises exploratórias, validação temática ou comunicação executiva. A IA pode ajudar a gerar e comparar essas versões, mas a escolha final deve resultar de testes empíricos. Em muitos estudos cientométricos, a versão intermediária é mais adequada, pois corpora excessivamente amplos dificultam interpretação e corpora muito restritos empobrecem mapas e indicadores.

A noção de "base bibliográfica" também precisa ser compreendida criticamente. Bases não são espelhos neutros da ciência. Elas têm políticas de indexação, cobertura geográfica, linguística, disciplinar e temporal. Web of Science historicamente enfatiza seletividade; Scopus amplia cobertura em muitos campos; Dimensions integra diferentes objetos de pesquisa; OpenAlex oferece abertura e escala, mas com desafios de qualidade de metadados; PubMed tem forte orientação biomédica e vocabulário controlado robusto. Uma mesma *string* pode produzir *corpora* distintos em cada base, não apenas por sintaxe, mas por cobertura (HOOK; PORTER; HERZOG, 2018; SINGH *et al.*, 2021; VISSER; VAN ECK; WALTMAN, 2021). A comparação entre bases deve, portanto, fazer parte da interpretação metodológica, não apenas da coleta.

Em síntese, estratégias booleanas são instrumentos de delimitação epistêmica. Elas traduzem uma pergunta em regras formais de recuperação, mas também revelam decisões sobre linguagem, campo, período e fonte. A IA pode acelerar a produção de versões iniciais e auxiliar na revisão, mas não substitui a experiência do pesquisador, do bibliotecário ou do especialista em informação. Uma estratégia booleana cientometricamente defensável, conforme princípios de busca estruturada e relato transparente (BRAMER *et al.*, 2018; RETHLEFSEN *et al.*, 2021), é aquela que pode ser explicada, testada, corrigida, documentada e vinculada aos objetivos do estudo.

7.4 Tradução terminológica entre áreas do conhecimento

A tradução terminológica entre áreas é uma das tarefas mais complexas em estudos cientométricos interdisciplinares. Em linguagem comum, traduzir pode significar passar uma palavra de uma língua para outra. Em cientometria, traduzir também significa converter conceitos entre tradições disciplinares, comunidades metodológicas, bases de indexação e regimes de comunicação científica. O mesmo fenômeno pode ser descrito por vocabulários distintos em ciência da informação, ciência da computação, saúde, educação, administração, comunicação e estudos sociais da ciência. Quando um estudo pretende mapear um tema transversal, a busca precisa reconhecer essas diferenças para não reduzir o campo a uma única comunidade discursiva.

Considere o tema "uso de IA para apoiar buscas bibliográficas". Na saúde, a literatura pode falar em *"systematic review search strategy", "information retrieval", "evidence synthesis", "screening automation"* e *"MeSH term identification"*. Na computação, pode aparecer como *"query expansion", "neural information retrieval", "semantic search", "language model for search", "retrieval augmented generation"* ou *"Boolean query generation"*. Na ciência da informação, podem surgir *"bibliographic retrieval", "controlled vocabulary", "indexing", "search filters"* e *"knowledge organization"*. Em administração, o mesmo tema pode ser associado a *"systematic literature review"* e *"literature sampling"*. Uma busca que usa apenas uma dessas linguagens produzirá um mapa parcial.

A IA pode ser útil justamente porque modelos de linguagem foram treinados em grande diversidade textual e conseguem sugerir equivalências prováveis entre áreas. Porém, essa capacidade semântica é acompanhada de risco: a IA pode aproximar termos por semelhança superficial e ignorar diferenças conceituais. "*Semantic search*" e "*systematic search*" parecem próximos, mas pertencem a tradições distintas; "*topic modeling*" e *"science mapping"* podem dialogar, mas não são sinônimos; "*bibliometric analysis*" e "*systematic literature review*" podem coexistir em um artigo, mas têm objetivos diferentes. A tradução terminológica deve, portanto, operar por comparação contextual, não por equivalência automática.

Um método prático consiste em construir mapas de conceitos por comunidade. Primeiro, identifica-se o núcleo conceitual do estudo. Depois, pede-se à IA que liste como esse núcleo é nomeado em diferentes áreas, solicitando sempre justificativa e exemplos de contextos de uso. Em seguida, o pesquisador testa os termos em bases ou amostras documentais. O resultado pode ser uma tabela com colunas: área, termo provável, tipo de termo, relação com o conceito central, risco de ambiguidade, decisão de inclusão e observações. Essa tabela ajuda a distinguir termos nucleares, termos ponte e termos contextuais. Termos nucleares compõem a *string* principal; termos ponte podem ser usados em buscas complementares; termos contextuais ajudam na interpretação, mas talvez não devam entrar na busca principal.

A tradução entre áreas também envolve níveis de granularidade. Um termo amplo como *"artificial intelligence"* pode ser adequado para recuperar literatura geral, mas insuficiente para estudos centrados em *prompts*. Termos específicos como "*prompt engineering*", "*few-shot prompting*", "*chain-of-thought prompting*" ou "*large language models*" capturam subcampos, mas podem excluir literatura anterior. Em outra direção, termos como "*machine learning*" podem ser amplos demais e recuperar estudos sem relação com IA generativa. A decisão depende da pergunta. Se o objetivo é mapear estratégias de busca assistidas por IA em sentido amplo, *machine learning* pode entrar. Se o foco é *prompts* em modelos generativos, o termo deve ser usado com filtros conceituais mais fortes.

A dimensão linguística adiciona outra camada. Traduções literais podem falhar porque comunidades acadêmicas adotam termos híbridos, anglicismos ou expressões específicas.

Em português, "engenharia de *prompts*" convive com "*prompt engineering*", "comando", "instrução", "solicitação" e "*prompt*". Em espanhol, "*ingeniería de prompts*" e "*prompts*" podem aparecer lado a lado. Em bases internacionais, títulos em português podem ter resumos em inglês ou palavras-chave bilíngues. A IA pode sugerir versões multilíngues, mas o pesquisador deve verificar uso real. Uma estratégia multilíngue deve evitar tanto excesso de termos raros quanto exclusão de expressões locais relevantes.

A tradução terminológica pode ser validada por análise de coocorrência. Depois de executar buscas piloto, o pesquisador pode extrair palavras-chave e termos frequentes dos registros recuperados. Se os termos sugeridos pela IA aparecem com frequência e em contextos relevantes, ganham plausibilidade. Se aparecem apenas em registros periféricos ou não aparecem, devem ser revisados. Essa prática transforma a IA em participante de um ciclo empírico: ela sugere, a base responde, o pesquisador interpreta, e a estratégia é ajustada. A busca deixa de ser ato único e passa a ser processo iterativo de aproximação entre linguagem teórica e linguagem indexada.

Em campos emergentes, a tradução terminológica deve lidar com marcas de produto, nomes de modelos e nomes de plataformas. "ChatGPT" pode recuperar literatura sobre IA generativa, mas também introduzir viés de marca e período. "GPT" pode ser ambíguo ou específico demais. "*large language model*" é mais conceitual, mas ainda recente. "*foundation model*" amplia para modelos multimodais e infraestrutura. Uma estratégia cientométrica precisa decidir se marcas e modelos serão termos de busca, objetos de análise ou ambos. Quando incluídos na busca, podem aumentar sensibilidade para literatura recente; quando usados apenas na análise, podem servir para classificar subtemas dentro de um *corpus* mais amplo.

A transição entre descritores controlados e termos livres exige atenção particular. Um vocabulário controlado pode indexar um conceito sob rótulo diferente daquele usado pelos autores. A IA pode ajudar a identificar descritores candidatos, mas não deve inventá-los. O procedimento adequado é consultar diretamente o tesauro da base. No caso de PubMed, por exemplo, MeSH deve ser verificado no MeSH Database; no caso de DeCS, em sua interface própria; no caso de Emtree, conforme acesso disponível. Se a IA sugere um descritor inexistente, o erro pode parecer pequeno, mas compromete a execução da busca. Em bases que não usam vocabulário controlado tradicional, a tradução passa por campos de tópico, categorias, conceitos ou classificações automatizadas, cada uma com pressupostos próprios.

A tradução entre bases é distinta da tradução entre áreas, mas ambas se cruzam. Uma expressão pode ser relevante conceitualmente, mas ter baixo desempenho em uma base porque o campo não está disponível, porque a indexação é inconsistente ou porque a base interpreta caracteres de modo diferente. Em OpenAlex, por exemplo, o pesquisador pode explorar conceitos, tópicos e filtros estruturados; em Web of Science, categorias de área e campos TS; em Scopus, TITLE-ABS-KEY e códigos de área; em PubMed, MeSH e campos específicos. A IA pode comparar essas possibilidades, desde que o *prompt* solicite explicitamente: "não apenas traduza os termos, mas traduza a estratégia para a lógica da base".

A interdisciplinaridade também exige evitar capturas assimétricas. Quando um tema cruza áreas, a base dominante pode representar melhor uma delas. Estudos sobre IA e saúde são fortemente visíveis em PubMed; estudos sobre IA e computação aparecem em conferências e bases técnicas; estudos de ciência da informação aparecem em periódicos de biblioteconomia e *scientometrics*; estudos de gestão podem estar em Scopus e Web of Science, mas com vocabulário diferente. Uma busca que combina todas as áreas em uma única *string* pode tornar-se volumosa e pouco precisa. Em alguns casos, é metodologicamente melhor construir subestratégias por comunidade e depois integrar corpora com deduplicação e documentação clara.

A IA pode apoiar essa integração propondo "famílias terminológicas". Uma família terminológica agrupa expressões que representam o mesmo núcleo conceitual em comunidades diferentes, sem tratá-las como sinônimos perfeitos. Por exemplo, a família "recuperação de literatura" pode conter "*literature search*", "*bibliographic search*", "*information retrieval*", "*search strategy*", "*query formulation*" e "*database searching*". Cada termo recebe anotação sobre área de uso, campo preferencial, risco e decisão. Esse tipo de anotação evita a ilusão de que basta empilhar termos com OR. O OR entre termos heterogêneos precisa ser sustentado por uma justificativa conceitual.

Por fim, a tradução terminológica deve ser incorporada ao relatório metodológico. O leitor precisa entender por que certos termos foram incluídos e outros excluídos. Essa transparência é essencial quando a IA participou da expansão. O relatório não precisa reproduzir todas as interações com o modelo, mas deve descrever o procedimento: documentos-semente utilizados, *prompt* aplicado, critérios de seleção, bases consultadas para validação, testes realizados e versão final da matriz. Quando a busca é multilíngue ou interdisciplinar, convém explicitar as comunidades cobertas e as limitações. Nenhuma estratégia recupera a totalidade da ciência; o objetivo é construir *corpus* adequado, justificado e interpretável.

7.5 Busca sensível versus busca específica

A distinção entre busca sensível e busca específica é central para qualquer estratégia bibliográfica, mas assume implicações particulares em cientometria. Uma busca sensível busca recuperar a maior proporção possível de documentos relevantes, mesmo ao custo de incluir muitos registros irrelevantes. Uma busca específica procura recuperar registros mais precisamente alinhados ao tema, mesmo ao custo de perder parte da literatura relevante. Na linguagem da recuperação da informação, essa tensão se aproxima da relação entre revocação e precisão. Em estudos cientométricos, o equilíbrio entre essas dimensões determina o tamanho do *corpus*, a estabilidade dos indicadores, a qualidade dos mapas e a credibilidade da interpretação.

Não existe nível ideal universal de sensibilidade ou especificidade. A escolha depende da finalidade do estudo. Quando o objetivo é revisão sistemática de evidências clínicas, alta sensibilidade costuma ser prioridade, pois perder estudos relevantes pode enviesar conclusões sobre intervenções. Quando o objetivo é mapeamento cientométrico de um campo amplo, busca muito sensível pode gerar *corpus* heterogêneo demais, dificultando a interpretação temática. Quando o objetivo é identificar área emergente, sensibilidade é importante para capturar termos recentes e variantes. Quando o objetivo é comparar produtividade institucional em tema muito delimitado, especificidade pode ser mais relevante para evitar atribuições indevidas.

A IA pode ajudar o pesquisador a explicitar esse compromisso. Um *prompt* bem formulado deve pedir ao modelo que proponha versões ampla, equilibrada e restrita da estratégia, indicando os efeitos esperados de cada uma. A versão ampla incluirá mais sinônimos, termos relacionados, variantes e campos. A versão restrita usará frases exatas, campos centrais, operadores de proximidade e talvez exclusões. A versão equilibrada buscará recuperar documentos centrais sem capturar todo o entorno temático. O pesquisador então testa as versões, compara contagens e analisa amostras de resultados. Essa prática torna visível o trade-off, evitando que a escolha da *string* pareça arbitrária.

Uma forma robusta de avaliar sensibilidade é usar conjunto-semente ou conjunto de referência. Esse conjunto contém documentos que, por conhecimento prévio ou curadoria especializada, devem ser recuperados pela busca. Se a *string* não recupera artigos centrais, a

estratégia provavelmente tem lacunas. O conjunto-semente pode incluir revisões relevantes, artigos clássicos, estudos recentes conhecidos, trabalhos de autores fundamentais e documentos de diferentes comunidades disciplinares. A IA pode ajudar a organizar o conjunto, extrair termos dos registros e perguntar por que alguns não foram recuperados. Contudo, a seleção do conjunto-semente deve ser humana e justificável.

A avaliação de especificidade pode ser feita por amostragem de ruído. Após executar a busca, seleciona-se amostra de registros recuperados, preferencialmente estratificada por ano, base ou bloco terminológico. O pesquisador classifica os registros como relevantes, parcialmente relevantes ou irrelevantes. A IA pode apoiar a classificação preliminar, desde que receba critérios explícitos e que uma amostra seja validada por humanos. O resultado permite identificar termos que produzem ruído. Por exemplo, "*mapping*" pode recuperar cartografia geográfica; "*impact*" pode recuperar impacto ambiental ou social sem relação com impacto científico; "*prompt*" pode recuperar *prompts* educacionais ou estímulos psicológicos. Esses achados orientam ajustes na *string*.

O número de resultados não deve ser interpretado isoladamente. Uma busca com muitos resultados não é necessariamente boa; uma busca com poucos resultados não é necessariamente ruim. O número precisa ser avaliado em relação ao campo, ao período, à base e à pergunta. Um tema emergente pode ter poucos registros e ainda assim ser relevante; uma expressão ampla pode recuperar milhares de documentos com baixa pertinência. A IA pode comparar contagens entre versões, mas não deve concluir sozinha qual é melhor. O pesquisador deve examinar qualidade, diversidade, cobertura e adequação aos objetivos.

Em estudos cientométricos, há risco de confundir *corpus* grande com estudo robusto. Mapas gerados a partir de grandes volumes podem parecer mais "científicos", mas o tamanho do *corpus* não garante validade. Um *corpus* amplo demais pode produzir *clusters* dominados por temas periféricos, indicadores inflados e relações de coocorrência pouco interpretáveis. Por outro lado, *corpus* restrito demais pode apagar subtemas, países, idiomas ou comunidades minoritárias. O equilíbrio exige julgamento substantivo. A IA pode ser usada para simular implicações: "Que tipos de documentos esta versão pode perder?"; "Que ruídos esta versão pode incluir?"; "Como a inclusão deste termo altera o escopo conceitual?".

A busca sensível é especialmente importante para identificar a emergência de temas. Termos emergentes podem aparecer primeiro em conferências, *preprints*, palavras-chave de autores, títulos informais ou combinações terminológicas instáveis. Uma estratégia restrita pode detectar o tema apenas quando ele já está consolidado. No entanto, a sensibilidade ampliada aumenta a necessidade de filtragem posterior. Para temas emergentes, recomenda-se separar a busca de descoberta da busca de análise. A busca de descoberta pode ser ampla e exploratória, com apoio de IA para minerar termos. A busca de análise deve ser refinada, documentada e adequada à geração de indicadores. Confundir essas etapas leva a *corpus* instável.

A busca específica é importante quando o objetivo é avaliação institucional, comparação entre países ou análise de desempenho. Nesses casos, incluir documentos fora do escopo pode produzir diagnósticos injustos ou enganosos. Se uma instituição é avaliada por sua produção em "IA aplicada à cientometria", a *string* precisa evitar capturar toda a literatura de IA institucional. Ainda assim, especificidade excessiva pode excluir grupos que usam termos diferentes. A solução é combinar critérios: *string* temática, filtros institucionais, leitura de amostras, validação por especialistas e, quando possível, classificação posterior do *corpus*. A IA pode apoiar a construção de regras de classificação, mas não deve decidir avaliação sem revisão humana.

A literatura de busca sistemática (BRAMER *et al.*, 2018; GUSENBAUER; HADDAWAY, 2020; MACFARLANE; RUSSELL-ROSE; SHOKRANEH, 2022) oferece conceitos úteis para a cientometria, como sensibilidade, precisão, especificidade, número necessário para leitura e avaliação por *relative recall*. Embora esses conceitos tenham sido desenvolvidos principalmente para revisões de evidência, eles podem inspirar procedimentos de teste em mapeamentos científicos. Em vez de assumir que uma busca é adequada, o pesquisador pode medir quantos documentos do conjunto-semente foram recuperados, estimar a proporção de ruído em amostras e documentar as mudanças entre versões. Isso aproxima a cientometria de práticas mais transparentes de recuperação da informação.

A IA pode ainda apoiar a análise de lacunas. Uma lacuna de busca ocorre quando a estratégia não recupera documentos que deveriam estar no *corpus*. As causas podem ser termo ausente, campo inadequado, truncamento insuficiente, frase exata restritiva, base com cobertura limitada ou erro sintático. Ao receber documentos-semente não recuperados, a IA pode comparar títulos, resumos e palavras-chave desses documentos com a *string* usada, sugerindo termos ou campos que explicam a ausência. Essa função é valiosa, mas deve ser verificada. O modelo pode sugerir incluir termos que ampliam demais a busca; por isso, cada sugestão deve ser testada empiricamente.

O ruído também pode ser analisado com apoio de IA. Ao receber uma amostra de registros irrelevantes, o modelo pode identificar padrões: termos ambíguos, áreas indesejadas, usos metafóricos, homônimos, nomes de técnicas com múltiplos significados. O pesquisador pode então decidir se usa exclusões, operadores de proximidade, campos mais restritos ou refinamento por termos adicionais. Em geral, deve-se preferir ajustes positivos, como aumentar especificidade de blocos, antes de recorrer a exclusões amplas. Excluir termos pode parecer eficiente, mas frequentemente cria cegueira metodológica.

O equilíbrio entre sensibilidade e especificidade também depende da etapa do estudo. No início, busca ampla ajuda a compreender o campo. Na fase de coleta principal, busca equilibrada gera *corpus* analisável. Na fase de validação, busca específica pode testar subtemas. Na fase de interpretação, buscas complementares podem esclarecer *clusters* ou anomalias. A IA pode ser usada em cada etapa com *prompts* distintos. *Prompts* exploratórios devem ampliar possibilidades; *prompts* analíticos devem organizar alternativas; *prompts* críticos devem identificar riscos; *prompts* redacionais devem documentar escolhas. Confundir esses modos de *prompt* leva a respostas úteis em aparência, mas inadequadas ao momento metodológico.

Um aspecto ético e político da busca sensível versus específica é a inclusão de literaturas periféricas. Estratégias muito específicas, baseadas em termos dominantes em inglês e bases seletivas, podem excluir produção de países, idiomas e periódicos menos centrais. Estratégias muito sensíveis podem recuperar essa diversidade, mas misturá-la com ruído e dificultar análise. O pesquisador deve decidir explicitamente se seu estudo pretende representar a literatura internacional *mainstream*, a literatura regional, a produção aberta, a literatura de conferências ou um conjunto combinado. A IA pode ajudar a listar bases e termos alternativos, mas a decisão é epistemológica e deve ser transparente.

Em síntese, uma busca cientométrica de qualidade não é simplesmente a mais ampla nem a mais precisa. É aquela cujo equilíbrio entre sensibilidade e especificidade corresponde à pergunta, ao objeto, à unidade de análise, ao período e ao produto esperado. A IA pode tornar esse equilíbrio mais consciente ao gerar versões, simular riscos, analisar amostras e auxiliar na documentação. Contudo, a decisão final deve ser sustentada por testes, não por confiança na fluência da resposta gerada.

7.6 Teste, refinamento e documentação da estratégia de busca

A estratégia de busca raramente nasce pronta. Ela deve ser testada, refinada e documentada em ciclos. Esse processo é fundamental em estudos cientométricos porque o *corpus* resultante será base para indicadores, redes, mapas, séries temporais e diagnósticos. Se a busca é frágil, todo o encadeamento analítico herda a fragilidade. Por isso, a busca deve ser tratada como protocolo metodológico, não como etapa operacional invisível. A IA pode participar de cada ciclo, mas sua participação precisa ser registrada e submetida à verificação.

O primeiro teste consiste em executar a *string* em uma base e verificar se ela roda sem erro. Erros sintáticos podem ser evidentes, como parênteses não fechados, operadores não reconhecidos ou campos inexistentes. Outros são silenciosos: a base pode ignorar aspas, interpretar truncamento de modo inesperado, aplicar stemming automático ou converter termos para vocabulário controlado. O pesquisador deve observar mensagens da interface, contagem de resultados e exemplos de registros. A IA pode revisar a *string* antes da execução, mas não substitui o teste real. Cada base é uma infraestrutura concreta, e a busca deve ser validada nela.

O segundo teste é a recuperação de documentos-semente. A *string* deve recuperar documentos previamente considerados relevantes. Quando não recupera, é necessário investigar. O documento pode estar fora da base, fora do período, classificado como tipo documental excluído, indexado sem resumo, escrito em outro idioma ou nomeado com termo diferente. A IA pode ajudar a diagnosticar causas ao comparar a *string* com metadados do documento não recuperado. Porém, incluir todos os termos que aparecem em documentos-semente pode causar superajuste. A estratégia deve ser ajustada para capturar padrões legítimos, não para forçar recuperação de cada item conhecido.

O terceiro teste é a análise de amostra de resultados. Seleciona-se um conjunto de registros recuperados, preferencialmente incluindo os primeiros resultados, resultados intermediários, registros de diferentes anos e registros de diferentes blocos terminológicos. O pesquisador classifica a pertinência e identifica ruído. A IA pode auxiliar gerando uma tabela de classificação preliminar, mas a decisão deve ser auditável. É recomendável registrar critérios de relevância: o que torna um documento central, periférico ou irrelevante? Essa definição evita que ajustes posteriores sejam guiados apenas por impressão.

O quarto teste é a comparação entre versões da *string*. Versões amplas, intermediárias e restritas devem ser executadas e comparadas em termos de contagem, recuperação de documentos-semente, proporção de ruído em amostra, cobertura temporal e diversidade de fontes. Uma tabela comparativa permite justificar a escolha final. A IA pode produzir a tabela, interpretar diferenças e sugerir próximos testes. Contudo, as contagens devem vir das bases, não da imaginação do modelo. O *prompt* deve instruir a IA a trabalhar apenas com dados fornecidos pelo pesquisador. Se a IA não tiver contagens reais, deve declarar que não pode comparar empiricamente.

O quinto teste é a tradução entre bases. Uma estratégia construída em Scopus não pode ser simplesmente copiada para Web of Science, PubMed, Dimensions ou OpenAlex. A tradução deve adaptar campos, operadores, vocabulário controlado e limites. Após traduzir, o pesquisador executa as estratégias e compara resultados. Diferenças são esperadas, pois as bases têm coberturas distintas. A documentação deve registrar não apenas a *string*, mas a interface, a data da busca, a cobertura da base quando conhecida, os filtros aplicados e o número de registros recuperados. Sem esses dados, a replicação é limitada.

A documentação deve seguir princípios de completude e rastreabilidade. Um leitor deve ser capaz de saber o que foi buscado, onde, quando, como e com quais restrições.

O *PRISMA-S* (RETHLEFSEN *et al.*, 2021) oferece referência útil ao enfatizar itens de relato da busca, como bases consultadas, datas, estratégias completas, limites, filtros, processos de atualização e registro de resultados. Embora muitos estudos cientométricos não sejam revisões sistemáticas, eles podem adotar a lógica de transparência. Do mesmo modo, o PRESS Checklist (MCGOWAN *et al.*, 2016) inspira revisão por pares de estratégias eletrônicas, especialmente quanto a tradução da pergunta, operadores booleanos, cabeçalhos de assunto, sintaxe, limites e erros.

Em estudos com IA, a documentação deve incluir uma camada adicional: registro dos *prompts*. Não é necessário publicar conversas extensas quando isso compromete legibilidade ou contém material irrelevante, mas é recomendável documentar *prompts* metodologicamente decisivos: expansão de vocabulário, geração de *string*, revisão de ruído, tradução entre bases e auditoria de riscos. Para cada *prompt*, deve-se registrar data, ferramenta ou modelo quando possível, versão se disponível, entrada fornecida, saída aproveitada, decisões humanas e validações realizadas. Isso transforma o *prompt* em artefato reprodutível, ainda que a resposta do modelo possa variar.

A variação das respostas de IA é problema para reprodutibilidade. Mesmo com o mesmo *prompt*, modelos podem gerar respostas diferentes em momentos distintos, especialmente quando há atualização de sistemas ou parâmetros de aleatoriedade. Por isso, a reprodutibilidade não deve depender da capacidade de obter a mesma resposta, mas da documentação das decisões tomadas a partir da resposta. O pesquisador deve registrar a saída utilizada e a justificativa da seleção. Quando possível, pode-se configurar baixa temperatura para tarefas analíticas e solicitar formato estruturado. Ainda assim, a busca final deve ser documentada como produto humano validado, não como resultado automático.

O refinamento deve ser orientado por evidências. Se a amostra de resultados mostra muitos registros sobre educação que usam "*prompt*" como estímulo pedagógico, o pesquisador pode ajustar a *string* aproximando "*prompt*" de "*large language model*" por operador de proximidade. Se documentos-semente sobre "*query expansion*" não aparecem, pode ser necessário incluir termos de recuperação da informação. Se muitos registros são sobre mapas geográficos, *"science mapping"* deve ser buscado como frase ou em combinação com "*bibliometric*". Cada ajuste deve responder a problema observado. Ajustes intuitivos não documentados fragilizam a estratégia.

O uso de filtros por área pode ser útil, mas deve ser prudente. Bases como Scopus e Web of Science permitem limitar por áreas ou categorias. Isso pode reduzir ruído, mas também excluir literatura interdisciplinar. Um estudo sobre IA em cientometria pode aparecer em ciência da computação, informação, gestão, medicina, educação ou estudos sociais da ciência. Limitar a "*Library and Information Science*" pode produzir *corpus* mais específico, mas perder aplicações em outras áreas. A IA pode sugerir filtros disciplinares, mas o pesquisador deve testar com e sem filtros e justificar a opção.

A deduplicação faz parte da documentação da busca, embora ocorra depois da recuperação. Quando múltiplas bases são usadas, os mesmos documentos podem aparecer várias vezes. O relatório deve indicar quantos registros foram recuperados em cada base, quantos foram removidos como duplicatas e quais critérios foram usados para identificar duplicatas. DOI, título, ano e autores são campos comuns, mas erros de metadados exigem revisão. A IA pode ajudar a identificar duplicatas prováveis em pequenas amostras, mas grandes corpora exigem ferramentas adequadas e critérios consistentes.

A documentação deve incluir termos rejeitados. Essa prática é pouco comum, mas valiosa. Termos rejeitados mostram que a estratégia foi analisada criticamente. Por exemplo,

um termo pode ter sido rejeitado por ambiguidade excessiva, baixa ocorrência, redundância, inadequação conceitual ou ausência em bases. Registrar rejeições é especialmente importante quando a IA sugeriu muitos candidatos. Sem esse registro, leitores não saberão se lacunas resultam de desconhecimento ou decisão metodológica.

Outra prática recomendável é preservar versões da *string*. A versão 1 pode ser a proposta inicial; a versão 2 incorpora termos de documentos-semente; a versão 3 ajusta ruído; a versão 4 traduz para bases; a versão final é executada para coleta. Cada versão deve ter data, motivo da alteração e resultado do teste. Essa abordagem aproxima a busca de um processo de versionamento científico. Em projetos colaborativos, planilhas, repositórios ou cadernos eletrônicos podem registrar essas versões. A IA pode transformar notas de refinamento em protocolo narrativo, mas não deve apagar a história das decisões.

A atualização da busca também precisa ser prevista. Em campos dinâmicos, especialmente IA, novas publicações e termos surgem rapidamente. Um estudo cientométrico pode precisar de busca de atualização antes da submissão do artigo ou fechamento do relatório. A documentação deve distinguir busca inicial, busca de atualização e eventuais complementações. Se a estratégia mudou entre momentos, isso deve ser explicado. A IA pode sugerir termos recentes para atualização, mas eles precisam ser testados sem comprometer comparabilidade temporal. Introduzir termo novo apenas no final da série pode alterar artificialmente a curva de crescimento.

A validação por pares da estratégia é desejável. Um bibliotecário, cientometrista ou especialista no tema pode revisar a matriz terminológica e a *string*. Essa revisão pode usar checklists inspirados em PRESS. A IA pode simular uma revisão crítica, apontando possíveis problemas, mas a revisão humana continua necessária quando o estudo tem consequências avaliativas ou será publicado. O ideal é combinar revisão automatizada, revisão por especialista em busca e revisão por especialista de domínio. Cada revisor percebe riscos diferentes.

Em síntese, testar, refinar e documentar a busca é transformar uma operação técnica em método científico. A IA amplia a capacidade de gerar alternativas e revisar coerência, mas não torna dispensáveis os princípios de transparência, replicabilidade e validação. Uma estratégia de busca bem documentada permite que leitores compreendam as fronteiras do *corpus* e interpretem resultados com prudência. Uma estratégia mal documentada torna mapas e indicadores pouco confiáveis, ainda que visualmente sofisticados.

7.7 Como evitar que a IA invente termos, autores ou bases inexistentes

Modelos de linguagem são treinados para produzir texto plausível e operar por padrões probabilísticos de geração textual (VASWANI *et al.*, 2017; BROWN *et al.*, 2020; LIU *et al.*, 2023). Essa característica explica parte de sua utilidade, mas também de seus riscos. Ao apoiar estratégias de busca, a IA pode sugerir termos inexistentes, autores fictícios, bases que não existem, descritores que não pertencem a nenhum vocabulário controlado, DOIs falsos ou funcionalidades inexistentes em plataformas reais. Em atividades de busca, esse risco é particularmente sério porque uma sugestão aparentemente técnica pode ser incorporada à estratégia sem verificação. A fluência do modelo pode criar falsa autoridade metodológica.

A prevenção começa pela formulação do *prompt*. O usuário deve instruir explicitamente a IA a distinguir três categorias: termos confirmados, termos prováveis e termos a verificar. Deve pedir que o modelo não invente descritores, autores, bases, operadores ou campos; que indique incerteza quando não puder confirmar; e que recomende validação em fontes específicas. Em vez de perguntar "quais descritores devo usar?", é melhor perguntar "quais

candidatos a descritores devo verificar em MeSH, DeCS ou Emtree?". A mudança de linguagem desloca a IA de autoridade para assistente de triagem.

Outra medida é exigir rastreabilidade. Quando a IA sugere um termo, deve informar por que ele é relevante, em qual conceito se encaixa e qual risco apresenta. Quando sugere um autor, deve indicar se o autor é central, periférico ou apenas exemplo provável. Quando sugere uma base, deve distinguir base bibliográfica, motor de busca acadêmico, repositório, índice de citações, catálogo ou plataforma comercial. Essa classificação reduz a chance de aceitar nomes plausíveis sem função definida. Ainda assim, a classificação é apenas hipótese; a verificação externa continua necessária.

A verificação de bases é simples, mas frequentemente esquecida. Antes de incluir uma base em protocolo, o pesquisador deve confirmar sua existência, escopo, acesso, cobertura, campos exportáveis, limites de uso e adequação ao tema. PubMed, Scopus, Web of Science, Dimensions e OpenAlex têm escopos e infraestruturas diferentes. Bases regionais como SciELO e repositórios institucionais podem ser relevantes conforme a pergunta. A IA pode listar opções, mas não deve ser a fonte final sobre cobertura. O pesquisador deve consultar documentação oficial, artigos metodológicos ou estudos comparativos.

A verificação de autores também é fundamental. Em cientometria, nomes de autores podem ser usados para documentos-semente, validação de campo ou busca complementar. A IA pode sugerir autores reconhecidos, mas também pode misturar pessoas, atribuir temas incorretos ou inventar referências. Para evitar isso, autores sugeridos devem ser confirmados por bases como ORCID, Scopus Author ID, Web of Science ResearcherID, OpenAlex Authors, Google Scholar ou registros bibliográficos. Quando o autor é usado apenas como exemplo, o risco é menor; quando sua produção compõe *corpus* ou diagnóstico, a validação é indispensável.

Descritores controlados exigem verificação direta. A IA pode conhecer muitos termos MeSH, mas também pode gerar rótulos plausíveis que não existem. O pesquisador deve consultar a interface do vocabulário e registrar o descritor exato, identificador quando disponível, data de consulta e decisão de explosão ou não explosão hierárquica. Em PubMed, o uso de MeSH pode recuperar documentos indexados sob conceito, mas documentos recentes ainda não indexados podem ser perdidos; por isso, termos livres e descritores devem ser combinados. A IA pode ajudar a explicar essa lógica, mas não substituir a consulta ao vocabulário.

Outro risco é a invenção de sintaxe. Modelos podem sugerir operadores não suportados pela base ou misturar sintaxes. Uma *string* que combina TITLE-ABS-KEY com MeSH ou TS com [tiab] revela confusão entre plataformas. O pesquisador deve sempre pedir *strings* separadas por base, com nota sobre sintaxe. Mesmo assim, deve testar cada uma. Quando a IA não tiver certeza sobre a sintaxe, deve ser instruída a declarar incerteza. *Prompts* que recompensam respostas completas a qualquer custo aumentam risco de alucinação operacional.

A invenção de termos pode ser mais sutil. O modelo pode sugerir expressões que existem linguisticamente, mas não são usadas como termos acadêmicos relevantes. Por exemplo, uma expressão como "*AI-enhanced scientometric prompting*" pode ser compreensível, mas talvez não apareça na literatura. Incluir esse termo em uma *string* pode não causar ruído, mas também não contribuirá para recuperação. Em alguns casos, termos inventados reduzem eficiência; em outros, mascaram ausência de literatura. A verificação em bases reais mostra se o termo tem ocorrência significativa ou se deve ser tratado apenas como conceito descritivo do pesquisador.

A falsa autoridade aparece quando a IA apresenta justificativas convincentes para uma decisão incorreta. Ela pode afirmar que uma base cobre determinado tipo documental

sem mencionar limitações; pode dizer que um termo é amplamente utilizado quando é raro; pode recomendar excluir um termo relevante por parecer ambíguo; pode sugerir que uma *string* é "otimizada" sem teste. Para reduzir esse risco, os *prompts* devem exigir que o modelo separe dado, inferência e recomendação. "Dado" é algo fornecido pelo usuário ou verificado externamente. "Inferência" é hipótese baseada em padrões. "Recomendação" é proposta a ser testada. Essa separação fortalece a auditoria.

A perda de rastreabilidade ocorre quando decisões passam a ser incorporadas sem registro de origem. Se a IA sugere dez termos e o pesquisador escolhe quatro, é necessário saber por quê. Se uma *string* foi alterada após uma resposta do modelo, a alteração deve ser registrada. Se termos foram descartados, as razões devem ser preservadas. Sem rastreabilidade, não se pode avaliar se a IA melhorou a busca ou introduziu viés. A documentação não precisa ser burocrática; pode ser uma tabela simples com colunas: sugestão da IA, decisão humana, evidência, ação e observação.

Em ambientes institucionais, também há questões de confidencialidade. Estratégias de busca podem envolver temas sensíveis, diagnósticos institucionais, áreas estratégicas ou dados internos. O uso de IA deve respeitar políticas de privacidade e propriedade intelectual. Quando dados não públicos são usados, deve-se evitar inseri-los em ferramentas sem garantias adequadas. A busca cientométrica pode parecer atividade bibliográfica neutra, mas em contextos de avaliação e planejamento ela pode revelar prioridades, fragilidades e interesses estratégicos. O pesquisador deve considerar a governança do uso de IA.

Uma técnica prática para reduzir invenções é exigir que a IA produza listas com status. Por exemplo: "confirmado em vocabulário controlado", "termo livre plausível", "sigla ambígua", "necessita verificação", "não recomendado sem teste". Outra técnica é solicitar que a IA gere perguntas de verificação em vez de respostas finais: "Este termo aparece em títulos ou resumos?", "Há descritor controlado equivalente?", "A base aceita truncamento neste campo?", "O termo recupera documentos-semente?". Essas perguntas orientam o pesquisador a executar validação empírica.

A busca com apoio de IA deve adotar o princípio da desconfiança produtiva. Desconfiar não significa rejeitar a ferramenta, mas reconhecer que sua utilidade depende de validação. A IA é eficaz para ampliar perspectivas, sugerir candidatos e organizar raciocínios; é inadequada como única fonte de verdade sobre bases, autores e termos. Em cientometria, onde resultados podem informar avaliação institucional, políticas científicas e decisões de financiamento, a responsabilidade pelo *corpus* não pode ser atribuída ao modelo. O pesquisador deve assumir autoria metodológica.

Em síntese, evitar invenções da IA requer *prompts* restritivos, verificação externa, documentação de decisões e revisão humana, uma vez que alucinações e imprecisões referenciais em modelos conversacionais já foram documentadas em estudos recentes (CHELLI *et al.*, 2024). A melhor estratégia não é pedir que a IA "não alucine", pois essa instrução isolada é insuficiente. É estruturar o processo para que alucinações sejam detectáveis: separar candidatos de confirmações, exigir incerteza, testar em bases reais, registrar evidências e submeter decisões a revisão. Assim, a IA torna-se componente de um sistema de qualidade, e não fonte opaca de autoridade.

7.8 Prompts do capítulo

Os *prompts* a seguir foram concebidos como modelos reutilizáveis e adaptáveis. Eles não devem ser copiados de forma mecânica. O pesquisador deve preencher campos, inserir contexto, informar base, período, critérios de inclusão e objetivos do estudo. Em todos os

casos, recomenda-se salvar o *prompt*, a resposta, os ajustes humanos e os resultados dos testes. A utilidade científica do *prompt* depende de sua integração ao protocolo de busca.

Prompt 7.8.1 Expandir vocabulário de busca

Atue como especialista em cientometria, recuperação da informação e construção de estratégias de busca bibliográfica.

Meu tema é: [TEMA]. Minha pergunta preliminar é: [PERGUNTA].

Meu período de interesse é: [PERÍODO]. Minhas bases prováveis são: [BASES].

Tarefa: proponha uma matriz terminológica para a busca. Organize a resposta em tabela com as colunas: conceito central; termos livres em inglês; termos livres em português; termos livres em espanhol, quando pertinente; siglas; termos técnicos relacionados; descritores controlados candidatos, se existirem; termos emergentes; termos ambíguos; termos que devem ser verificados antes de uso; justificativa de inclusão; risco de ruído.

Restrições: não trate candidatos como descritores confirmados. Quando não tiver certeza, marque "verificar em base ou tesauro". Não invente autores, bases ou descritores. Separe termos nucleares de termos periféricos. Ao final, indique quais termos devem compor a *string* principal e quais devem ser usados apenas em busca exploratória.

Prompt 7.8.2 Criar string booleana

Atue como especialista em busca bibliográfica para estudos cientométricos. A partir da matriz terminológica abaixo, crie uma estratégia booleana para [BASE], considerando os campos disponíveis e a sintaxe própria da base.

Matriz terminológica: [INSERIR MATRIZ]

Pergunta de pesquisa: [INSERIR PERGUNTA]

Unidade de análise: [ARTIGO/AUTOR/PERIÓDICO/INSTITUIÇÃO/TEMA]

Objetivo da busca: [ALTA SENSIBILIDADE/EQUILÍBRIO/ALTA ESPECIFICIDADE]

Período: [ANOS]

Tipos documentais: [TIPOS]

Idiomas: [IDIOMAS]

Tarefa: construa três versões da estratégia: ampla, equilibrada e restrita. Para cada versão, apresente: blocos conceituais; *string* executável; explicação do uso de OR, AND, aspas, truncamento e proximidade; riscos de ruído; riscos de perda; campos usados; pontos que exigem teste na interface real.

Restrições: não misture sintaxe de bases diferentes. Se não souber a sintaxe exata, declare incerteza e sugira verificação na documentação da base. Não inclua operador NOT sem justificar com exemplos de ruído provável.

Prompt 7.8.3 Revisar ruído e lacunas

Atue como revisor metodológico de estratégia de busca. Abaixo estão: minha pergunta, a *string* usada, a base consultada, a contagem de resultados, uma lista de documentos-semente que deveriam ser recuperados e uma amostra de registros irrelevantes recuperados.

Pergunta: [PERGUNTA]

Base e data da busca: [BASE, DATA]

String: [STRING]

Contagem de resultados: [N]

Documentos-semente recuperados: [LISTA]

Documentos-semente não recuperados: [LISTA COM TÍTULO, RESUMO E PALAVRAS-CHAVE]

Amostra de ruído: [LISTA COM TÍTULO, RESUMO E PALAVRAS-CHAVE]

Tarefa: identifique possíveis lacunas da estratégia, termos que produziram ruído, termos ausentes, problemas de campo, problemas de sintaxe e ajustes possíveis. Organize a resposta em tabela com: problema observado; evidência; ajuste sugerido; efeito esperado sobre sensibilidade; efeito esperado sobre especificidade; risco do ajuste; teste recomendado.

Restrições: não proponha exclusões amplas sem demonstrar risco de perda. Diferencie dado fornecido, inferência e hipótese. Não afirme que a estratégia melhorou sem execução empírica.

Prompt 7.8.4 Comparar estratégias em Scopus, Web of Science, PubMed, Dimensions e OpenAlex

Atue como especialista em bases bibliográficas e cientometria. Quero comparar estratégias de busca para o tema [TEMA] nas bases Scopus, Web of Science, PubMed, Dimensions e OpenAlex.

Pergunta de pesquisa: [PERGUNTA]

Matriz terminológica: [MATRIZ]

Período: [ANOS]

Tipos documentais: [TIPOS]

Objetivo: [OBJETIVO]

Tarefa: para cada base, indique: campos recomendados; forma de adaptar a *string*; limitações de sintaxe; presença ou ausência de vocabulário controlado; cuidados com exportação; riscos de cobertura; metadados críticos para análise cientométrica; como documentar a busca. Em seguida, proponha uma estratégia de comparação de resultados entre bases, incluindo contagem, deduplicação, avaliação de documentos-semente e análise de diferenças de cobertura.

Restrições: não invente funcionalidades. Quando houver dúvida sobre sintaxe, marque como "verificar na documentação oficial". Não conclua que uma base é superior sem relacionar a conclusão à pergunta de pesquisa e à unidade de análise.

7.9 Quadro metodológico: da ideia inicial à estratégia documentada

A elaboração de uma estratégia de busca com apoio de IA pode ser organizada em oito etapas. A primeira é delimitar a pergunta cientométrica. A segunda é decompor a pergunta em conceitos. A terceira é construir matriz terminológica com termos livres, descritores e variantes. A quarta é gerar *strings* preliminares por base. A quinta é executar testes piloto. A sexta é analisar recuperação de documentos-semente e ruído. A sétima é refinar a estratégia. A oitava é documentar a versão final e suas justificativas. Em cada etapa, a IA pode apoiar tarefas específicas, mas a responsabilidade metodológica permanece humana.

Quadro 7.2 - Funções da IA nas etapas da busca cientométrica

Etapa	Contribuição possível da IA	Validação necessária
Delimitação conceitual	Sugerir conceitos centrais e periféricos; identificar ambiguidades.	Revisão por especialista no tema e alinhamento com a pergunta.
Expansão terminológica	Gerar sinônimos, siglas, traduções, termos emergentes e termos ponte.	Verificação em bases, vocabulários controlados e documentos-semente.
Construção booleana	Organizar blocos, propor versões ampla, equilibrada e restrita.	Teste na interface real da base e revisão de sintaxe.
Tradução entre bases	Adaptar campos e operadores conforme a plataforma.	Conferência na documentação oficial e execução piloto.
Análise de ruído	Agrupar padrões de irrelevância em amostras recuperadas.	Classificação humana de pertinência e avaliação de risco de exclusão.
Análise de lacunas	Comparar documentos-semente não recuperados com a *string*.	Testes empíricos de novos termos e prevenção de superajuste.
Documentação	Converter registros e notas em protocolo narrativo.	Conferência das datas, *strings*, contagens e decisões humanas.

Quadro 7.3 - Diferenças gerais entre bases para planejamento de busca

Base	Uso típico em cientometria	Cuidados na busca
Scopus	Ampla cobertura multidisciplinar, metadados de autores, afiliações e citações.	Sintaxe TITLE-ABS-KEY; limites de exportação; diferenças de cobertura por área.
Web of Science	Análises de citação, categorias, periódicos seletivos e séries históricas.	Uso de TS; atenção a índices incluídos; cobertura seletiva e variações por coleção.

Base	Uso típico em cientometria	Cuidados na busca
PubMed	Literatura biomédica, MeSH e buscas sensíveis para saúde.	Combinar MeSH e termos livres; documentos recentes podem não estar indexados com MeSH.
Dimensions	Integração de publicações, *grants*, patentes e outros objetos.	Verificar filtros, tipos documentais e exportação; interpretar cobertura conforme finalidade.
OpenAlex	Fonte aberta em larga escala, útil para APIs e grafos de conhecimento científico.	Avaliar qualidade de metadados, filtros, desambiguação e adequação à pergunta.

7.10 Exemplo aplicado: busca sobre prompts em estudos cientométricos

Suponha que o pesquisador deseje mapear a literatura sobre uso de *prompts* e modelos de linguagem em estudos cientométricos. Uma abordagem apressada poderia buscar apenas *"prompt engineering"* AND *"scientometrics"*. Essa estratégia provavelmente seria específica demais e perderia estudos que usam expressões como *"large language models"*, *"generative AI"*, *"bibliometric analysis"*, *"science mapping"* e *"research evaluation"*. Uma abordagem excessivamente ampla, como *"artificial intelligence"* AND *"bibliometrics"*, recuperaria literatura sobre algoritmos de classificação, redes neurais, aprendizado de máquina e aplicações que talvez não envolvam *prompts*. O desafio é construir blocos coerentes.

Um primeiro bloco poderia representar a análise da ciência: *scientometric**, *bibliometric**, *"science mapping"*, *"research evaluation"*, *"bibliometric mapping"*. Um segundo bloco poderia representar IA generativa e modelos de linguagem: *"large language model*"*, LLM, *"generative AI"*, "ChatGPT", *"foundation model*"*, *"natural language generation"*. Um terceiro bloco poderia representar *prompts*: *prompt**, *"prompt engineering"*, *"few-shot prompting"*, *"chain-of-thought"*. A decisão de tornar o terceiro bloco obrigatório dependeria da pergunta. Se o foco é *prompts*, sim. Se o foco é IA generativa aplicada à cientometria em geral, talvez o terceiro bloco deva ser opcional ou usado para análise interna.

A estratégia ampla poderia combinar os dois primeiros blocos e usar o terceiro como termo de enriquecimento. A estratégia equilibrada poderia combinar os três blocos, mas incluir variantes de bibliometria e modelos de linguagem. A estratégia restrita poderia exigir *"prompt engineering"* ou *prompt** próximo de *"large language model*"*. O pesquisador executaria as versões, observaria contagens, verificaria documentos-semente e analisaria ruído. A IA poderia ajudar a organizar os resultados em tabela, mas a escolha da versão final dependeria de evidências. Esse exemplo mostra que o *prompt* central do capítulo não é apenas um comando para gerar *string*; é instrumento para raciocinar sobre escopo.

7.11 Critérios de qualidade para estratégias de busca assistidas por IA

Uma estratégia de busca assistida por IA pode ser considerada metodologicamente madura quando satisfaz pelo menos dez critérios. Primeiro, a pergunta de pesquisa está explícita. Segundo, os conceitos foram separados antes da combinação booleana. Terceiro, há matriz terminológica com termos aceitos e rejeitados. Quarto, as *strings* foram adaptadas por base, sem mistura de sintaxes. Quinto, houve teste com documentos-semente. Sexto, houve análise de amostra de ruído. Sétimo, as versões da *string* foram preservadas. Oitavo, a participação da IA foi registrada. Nono, as decisões humanas foram justificadas. Décimo, a estratégia final é suficientemente clara para ser reproduzida ou criticada por outro pesquisador.

Esses critérios deslocam a avaliação da busca do plano estético para o plano científico. Uma *string* longa e tecnicamente impressionante pode ser ruim; uma *string* curta pode ser adequada se a pergunta for bem delimitada. Um *prompt* complexo pode produzir resposta útil; um *prompt* simples pode ser suficiente se o pesquisador fornecer dados claros. O que importa é a relação entre pergunta, *corpus*, método e validação. A IA não deve ser julgada por sua fluência, mas pela contribuição verificável que oferece ao processo de busca.

7.12 Considerações finais

A busca bibliográfica é uma das etapas mais decisivas da cientometria, pois define o universo de documentos a partir do qual serão produzidos indicadores, mapas, redes e interpretações. Com o apoio da IA, essa etapa pode tornar-se mais exploratória, mais reflexiva e mais documentada. Modelos de linguagem ajudam a ampliar vocabulários, identificar variantes, traduzir terminologias, revisar *strings*, organizar evidências e redigir protocolos. No entanto, a IA também introduz riscos de alucinação, falsa autoridade e perda de rastreabilidade. A solução não é rejeitar a IA, mas incorporá-la em um fluxo de validação.

A principal tese deste capítulo é que estratégias de busca com IA devem ser tratadas como protocolos auditáveis, em consonância com diretrizes de relato de busca e revisão de estratégias (MCGOWAN *et al.*, 2016; RETHLEFSEN *et al.*, 2021). O *prompt* não substitui a *string*; a *string* não substitui o teste; o teste não substitui a interpretação; e a interpretação não substitui a documentação. Cada etapa tem função própria. Quando essas funções são confundidas, a busca torna-se opaca. Quando são integradas, a IA fortalece a prática cientométrica ao ampliar alternativas e tornar mais explícitas as decisões do pesquisador.

Em uma ciência cada vez mais mediada por bases de dados, algoritmos de busca e modelos generativos, a competência de formular estratégias de recuperação torna-se competência epistemológica. Saber buscar é saber delimitar o que será considerado evidência da ciência. Por isso, a cientometria com IA não deve começar pela visualização de mapas, mas pela pergunta: que literatura estamos tornando visível, por quais termos, em quais bases e sob quais critérios? A resposta a essa pergunta define a qualidade do estudo.

Apêndice A.7.1 - Checklist operacional para busca com IA

O checklist a seguir substitui formulações repetidas por uma matriz editorial sintética. Cada item explicita o que deve ser verificado, qual evidência deve ser preservada e qual ação de controle deve ser registrada antes da coleta definitiva.

Item	Verificação	Evidência e ação de controle
1. Pergunta e escopo	Confirmar objeto, período, unidade de análise, finalidade do mapeamento e limites do *corpus*.	Registrar pergunta final, versões descartadas e justificativa de escopo.
2. Vocabulário	Conferir termos livres, descritores candidatos, siglas, variantes linguísticas e termos emergentes.	Manter matriz terminológica com termos aceitos, rejeitados e pendentes de teste.
3. Sintaxe	Testar operadores, campos, aspas, truncamentos e proximidade na interface real da base.	Salvar *string* executada, data, base, mensagem de erro, ajustes e versão final.

Item	Verificação	Evidência e ação de controle
4. Documentos-semente	Verificar se estudos centrais conhecidos são recuperados pela estratégia.	Listar documentos recuperados e não recuperados; explicar ausências sem superajustar a busca.
5. Ruído	Classificar amostra de registros irrelevantes e identificar padrões de ambiguidade.	Registrar critérios de pertinência, exemplos de ruído e ajustes proporcionais.
6. Lacunas	Comparar documentos não recuperados com a matriz terminológica e com os campos pesquisados.	Testar novos termos ou campos e documentar risco de ampliar excessivamente o *corpus*.
7. Bases	Registrar diferenças de cobertura, filtros, campos exportáveis e tipos documentais.	Justificar a escolha de cada fonte e indicar limitações de comparabilidade.
8. IA	Documentar *prompts*, respostas aproveitadas, decisões humanas e evidências de validação.	Preservar versão do *prompt*, saída utilizada, responsável pela revisão e decisão final.
9. Versões	Preservar mudanças na *string* e justificar cada refinamento.	Manter histórico com versão, motivo da mudança, contagens e efeitos observados.
10. Relato	Apresentar a estratégia final de modo completo, claro e criticável.	Incluir *string* completa, data da busca, base, filtros, contagens, deduplicação e limitações.

Apêndice A.7.2 - Exemplos de decisões de busca e justificativas metodológicas

Exemplo 1: quando um termo é conceitualmente correto, mas operacionalmente problemático. Em estudos sobre ciência aberta, a expressão "*open science*" é central, mas termos como "*open access*", "*open data*", "*open peer review*" e "*citizen science*" podem representar dimensões específicas ou campos parcialmente independentes. A IA tende a reunir todos esses termos como sinônimos ampliados, porém uma estratégia cientométrica precisa decidir se deseja mapear o guarda-chuva conceitual ou suas práticas componentes. A justificativa metodológica deve explicitar se a inclusão de subtermos tem função de sensibilidade ou de análise temática. Caso contrário, o *corpus* pode misturar literatura sobre políticas de acesso aberto com estudos sobre participação pública, produzindo mapas conceitualmente instáveis.

Exemplo 2: quando uma expressão emergente ainda não possui estabilidade lexical. Termos como "*AI-assisted research*", "*AI-augmented science*", "*generative AI in research*" e "*LLM-assisted literature review*" podem aparecer simultaneamente em um campo recente. Um *prompt* pode sugerir todos como relevantes, mas o pesquisador deve testar quais aparecem em título, resumo ou palavras-chave e quais são apenas formulações plausíveis. A decisão de incluir termos raros deve considerar se eles recuperam documentos que não seriam capturados por expressões mais consolidadas. Se não recuperam documentos adicionais, podem ser mantidos apenas como observação analítica, não como parte da *string* final.

Exemplo 3: quando uma sigla é útil, mas ambígua. "LLM" é amplamente usada para *large language model*, mas pode ter outros significados em áreas específicas. "AI" é ainda mais ambígua e pode aparecer em contextos não relacionados a *artificial intelligence* quando a base não diferencia maiúsculas, abreviações ou campos. O uso de siglas deve ser combinado com termos por extenso ou operadores de proximidade. A IA pode listar significados possíveis de uma sigla, mas a verificação deve ocorrer em amostras reais. Se a sigla produz ruído excessivo,

pode ser re*string*ida ao campo de título/resumo ou combinada com termos contextuais como "*language model*" ou "*generative*".

Exemplo 4: quando a busca deve capturar literatura de conferências. Em ciência da computação, muitos avanços metodológicos aparecem primeiro em anais de conferências. Uma estratégia limitada a artigos de periódicos pode subestimar a dinâmica do campo. Em estudos sobre *prompts* e modelos de linguagem, conferências como NeurIPS, ACL, EMNLP, SIGIR, CHI e ICLR podem ser relevantes conforme a pergunta. A IA pode lembrar essa característica disciplinar, mas a base escolhida precisa efetivamente indexar esses documentos. O relatório deve informar se procedimentos foram incluídos e como a inclusão afeta a comparação com áreas em que periódicos são a principal forma de comunicação.

Exemplo 5: quando a base não é adequada ao objeto. PubMed é excelente para literatura biomédica, mas não representa de forma abrangente ciência da informação, administração ou computação. Scopus e Web of Science são úteis para análises multidisciplinares, mas têm diferenças de cobertura e critérios de indexação. Dimensions e OpenAlex ampliam possibilidades de coleta, mas exigem atenção à qualidade de metadados e aos filtros. A IA pode sugerir várias bases, mas a escolha deve ser justificada em função da pergunta. Um estudo cientométrico não precisa usar todas as bases; precisa usar bases coerentes com o objeto e reconhecer limitações.

Exemplo 6: quando o operador NOT parece resolver ruído, mas cria perda invisível. Suponha que uma busca sobre *prompts* em IA recupere muitos estudos educacionais que usam "*prompt*" como estímulo de aprendizagem. Excluir "*education*" pode reduzir ruído, mas também eliminar estudos sobre formação de pesquisadores, ensino de busca bibliográfica ou uso de IA na pós-graduação. Uma alternativa mais segura é exigir proximidade entre *prompt* e termos como "*large language model*", "ChatGPT" ou "*generative AI*". O ajuste positivo delimita melhor o contexto sem excluir uma área inteira. A IA deve ser orientada a sugerir primeiro estratégias de especificação, e apenas depois exclusões.

Exemplo 7: quando há diferença entre busca para descoberta e busca para mensuração. Em uma etapa exploratória, o pesquisador pode usar termos amplos, bases abertas e consultas sem filtros rígidos para descobrir vocabulário e documentos-semente. Na etapa de mensuração, a busca deve ser estabilizada, documentada e executada de forma replicável. A IA é especialmente útil na primeira etapa, porque amplia hipóteses. Na segunda, sua função principal é revisão, organização e documentação. Misturar descoberta e mensuração pode levar a *corpora* mutáveis, nos quais novos termos são adicionados sem controle e curvas temporais deixam de representar uma estratégia constante.

Exemplo 8: quando a IA sugere autores como atalho para a literatura. A busca por autores centrais pode ser útil para identificar documentos-semente, mas não deve substituir a estratégia temática. Autores mudam de tema, colaboram em áreas distintas e podem ter homônimos. Além disso, uma busca centrada em autores tende a reforçar a visão canônica do campo, invisibilizando grupos periféricos. A IA pode indicar nomes prováveis, mas o pesquisador deve verificá-los em bases e decidir se serão usados apenas para validação, para busca complementar ou para análise de redes. O protocolo deve evitar que nomes sugeridos pela IA definam silenciosamente o *corpus*.

Exemplo 9: quando a tradução multilíngue altera o escopo. Incluir termos em português e espanhol pode recuperar literatura regional importante, mas também pode exigir bases que indexem periódicos locais. Se a busca for executada apenas em bases internacionais com baixa cobertura regional, a inclusão linguística pode ter pouco efeito. Por outro lado, bases regionais podem exigir sintaxe diferente e campos menos padronizados.

A IA pode sugerir traduções, mas a validação deve observar uso real em títulos e resumos. O relatório deve explicar se a busca é internacional em inglês, multilíngue ou orientada a determinada região científica.

Exemplo 10: quando contagens entre bases parecem contraditórias. Uma *string* pode recuperar mais registros em Dimensions do que em Web of Science, ou mais em Scopus do que em PubMed, não porque uma base seja "melhor" em abstrato, mas porque seus critérios de cobertura, tipos documentais, áreas e metadados diferem. A IA pode ajudar a interpretar essas diferenças, mas não deve transformá-las em ranking simplista. A comparação deve perguntar: qual base melhor representa a unidade de análise? Qual fornece campos necessários para redes e indicadores? Qual tem cobertura adequada ao período? Qual permite exportação em escala compatível com o estudo?

Exemplo 11: quando a documentação salva a interpretação. Se um mapa temático mostra um *cluster* inesperado, a primeira pergunta deve ser se esse *cluster* representa característica real do campo ou artefato da busca. A documentação da *string*, dos termos incluídos e dos ajustes de ruído permite responder. Sem documentação, interpretações *pós-hoc* podem atribuir significado científico a um resultado produzido por termo ambíguo. Com documentação, o pesquisador pode rastrear a origem do *cluster*, testar buscas alternativas e explicar limitações. Portanto, documentar a busca não é tarefa administrativa; é condição para interpretação responsável dos mapas.

Exemplo 12: quando o *prompt* precisa ser ajustado ao nível de maturidade do estudo. No início, *prompts* abertos como "liste termos relacionados" são aceitáveis. Na fase de construção da *string*, *prompts* devem exigir blocos conceituais, campos e restrições. Na revisão, devem solicitar evidências, riscos e testes. Na redação, devem transformar decisões em texto metodológico. O mesmo modelo de IA pode participar de todos esses momentos, mas o desenho do *prompt* deve mudar. Essa adaptação impede que a IA responda com generalidades quando o estudo já precisa de precisão operacional, ou que imponha rigidez quando o objetivo ainda é explorar possibilidades.

Capítulo 8

Coleta de dados bibliográficos

8.1 Introdução

A coleta de dados bibliográficos é a porta de entrada empírica de um estudo cientométrico. Antes de calcular indicadores, construir redes ou interpretar mapas, o pesquisador precisa definir de onde virão os registros que representarão a literatura analisada. Essa decisão não é neutra. A escolha entre Web of Science, Scopus, Dimensions, Crossref, OpenAlex, Lens, PubMed, SciELO, bases especializadas, repositórios institucionais ou catálogos internos produz diferentes retratos da ciência. Cada fonte expressa uma política de indexação, uma arquitetura de metadados, um modo de tratar documentos, uma forma de representar autores e instituições, uma cobertura disciplinar e geográfica e uma disponibilidade distinta de referências citadas e citações recebidas.

Na cientometria tradicional, a coleta já era uma etapa crítica. Em ambientes assistidos por inteligência artificial, torna-se ainda mais sensível, pois modelos generativos podem auxiliar na organização do processo, mas também podem reforçar lacunas, legitimar erros e converter decisões frágeis em textos aparentemente convincentes. Um *prompt* pode sugerir que uma base é adequada sem reconhecer sua cobertura limitada em determinado campo; pode propor uma rotina de API incompatível com os termos de uso da fonte; pode presumir que todos os registros têm afiliações completas; ou pode transformar critérios de inclusão mal definidos em um protocolo formalmente elegante, porém empiricamente insuficiente. A coleta assistida por IA deve, portanto, ser orientada por protocolos, evidências e documentação.

A literatura recente sobre comparação de fontes bibliográficas evidencia que bases distintas não são intercambiáveis. Visser, van Eck e Waltman (2021) compararam grandes fontes multidisciplinares e mostraram que cobertura documental, completude de links de citação e possibilidades de filtragem variam de maneira substantiva. Martín-Martín *et al.* (2021), ao compararem Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science e COCI por meio de citações, demonstraram diferenças relevantes entre áreas, tipos de documento e fontes. Esses estudos reforçam a necessidade de uma decisão informada sobre a fonte de dados, especialmente quando os resultados serão usados para avaliação, planejamento institucional ou formulação de políticas científicas.

A coleta não se reduz, contudo, à escolha da base. Ela envolve a transformação de uma pergunta cientométrica em estratégia de recuperação, a seleção de campos necessários, o controle de versões, a exportação de registros, a identificação de duplicatas, a preservação dos arquivos originais, a descrição de filtros aplicados e a documentação de limitações. Em estudos de redes, por exemplo, a ausência de referências citadas compromete análises de cocitação e acoplamento bibliográfico. Em estudos de colaboração institucional, a incompletude de afiliações impede atribuições confiáveis a organizações e países. Em estudos de temas emergentes, a ausência de resumos e palavras-chave reduz a qualidade da classificação textual. Em estudos de impacto, a origem dos dados de citação modifica os resultados.

A inteligência artificial pode apoiar a coleta ao organizar esse conjunto de decisões. Um modelo de linguagem pode ajudar a decompor a pergunta de pesquisa em campos necessários; sugerir fontes candidatas; construir uma matriz de comparação entre bases; transformar critérios de inclusão e exclusão em etapas operacionais; redigir um plano de coleta; identificar metadados ausentes em uma amostra; propor pseudocódigo para uso de API; ou produzir

um relatório de rastreabilidade. Todavia, a IA não possui, por si só, acesso automático às bases nem garantia de atualização, completude ou legalidade dos procedimentos sugeridos. O pesquisador deve verificar a documentação oficial, respeitar licenças, registrar datas de acesso e validar resultados por amostragem.

Este capítulo defende que a coleta de dados bibliográficos em cientometria com IA deve ser concebida como um processo auditável. Auditabilidade significa que outro pesquisador deve ser capaz de compreender quais fontes foram usadas, porque foram escolhidas, como a busca foi executada, quais campos foram exportados, que filtros foram aplicados, quais registros foram excluídos, como duplicatas foram tratadas, que limitações permanecem e quais *prompts* apoiaram a tomada de decisão. A presença de IA no processo não reduz a necessidade de documentação; ao contrário, amplia essa necessidade, pois introduz etapas conversacionais que podem influenciar escolhas metodológicas.

A estrutura deste capítulo acompanha essa concepção. Primeiro, discutem-se as fontes de dados. Em seguida, examinam-se os metadados essenciais. Depois, analisam-se formatos de exportação e interoperabilidade. A quarta parte apresenta APIs como meios de coleta programável. A quinta parte diferencia coleta manual, semiautomática e automatizada. A sexta parte propõe formas de registro reprodutível. Ao final, os *prompts* do capítulo são apresentados como modelos comentados, de modo que o leitor possa adaptá-los a diferentes projetos sem confundir apoio textual com validação empírica.

8.2 Fontes de dados: bases comerciais, abertas e institucionais

A fonte de dados é a infraestrutura que torna visível uma fração da produção científica. Em cientometria, essa fração costuma ser tratada como *corpus* analítico, mas ela nunca corresponde à totalidade da ciência produzida. A ciência publicada, indexada, recuperável e exportável é resultado de convenções editoriais, políticas de cobertura, critérios comerciais, práticas de registro de DOI, padrões de afiliação, idiomas de publicação e acordos de interoperabilidade. Por isso, a seleção da fonte deve ser justificada a partir da pergunta de pesquisa, e não apenas pela disponibilidade institucional ou pela familiaridade do pesquisador com determinada plataforma.

As bases comerciais, como Web of Science e Scopus, historicamente desempenham papel central em estudos cientométricos por oferecerem curadoria, campos estruturados, filtros disciplinares, dados de citação e mecanismos de exportação relativamente estáveis (BIRKLE *et al.*, 2020; VISSER; VAN ECK; WALTMAN, 2021). Elas são frequentemente empregadas em avaliações de produção, impacto e colaboração, bem como em estudos que exigem consistência na delimitação de periódicos e categorias. Entretanto, seu uso implica reconhecer limitações de cobertura, especialmente em áreas menos internacionalizadas, periódicos locais, produção em línguas diferentes do inglês, livros, capítulos, anais e literatura regional. Além disso, o acesso depende de assinatura e as possibilidades de redistribuição dos dados podem ser restritas.

As bases abertas e infraestruturas de metadados, como Crossref, OpenAlex, Lens, OpenCitations e PubMed em determinados contextos, ampliam a reprodutibilidade e a transparência da coleta (HENDRICKS *et al.*, 2020; HEIBI; PERONI; SHOTTON, 2019; CULBERT *et al.*, 2025). Elas reduzem barreiras de acesso, permitem consultas programáticas e favorecem a revisão por pares do processo de coleta. No entanto, abertura não equivale automaticamente a completude. Uma fonte aberta pode ter excelente cobertura de determinados identificadores e baixa completude de afiliações; pode oferecer muitos registros, mas com variações na qualidade de resumos, referências ou classificações

temáticas; pode atualizar-se de forma dinâmica, exigindo registro de data, versão e parâmetros de consulta. A cientometria com IA deve tratar essas fontes como recursos poderosos, mas não como espelhos integrais da ciência.

As bases institucionais compreendem repositórios de universidades, sistemas internos de gestão de pesquisa, currículos, catálogos de produção, plataformas de gestão de projetos, bases de programas de pós-graduação, repositórios de teses e dissertações e sistemas nacionais de informação científica. Elas são indispensáveis quando a pergunta se refere à produção de uma instituição, departamento, rede de pesquisa, laboratório ou política pública específica. A vantagem principal é a proximidade com a realidade administrativa e a possibilidade de incorporar documentos não indexados em bases internacionais. A limitação principal é a heterogeneidade: campos preenchidos de forma manual, nomes institucionais variáveis, ausência de DOI, registros duplicados, metadados incompletos e critérios de validação distintos.

A escolha da fonte deve começar pela pergunta cientométrica. Se o objetivo é comparar impacto de publicações internacionais em periódicos indexados, bases comerciais ou OpenAlex podem ser apropriadas. Se o objetivo é analisar a produção regional em saúde coletiva, SciELO, LILACS, PubMed e repositórios nacionais podem ser necessários. Se a pergunta envolve patentes, *grants*, ensaios clínicos ou políticas públicas, Dimensions, Lens, bases patentárias e documentos governamentais podem compor um ecossistema mais pertinente. Se a pergunta envolve literatura local ou produção técnico-científica institucional, sistemas internos e repositórios tornam-se fontes primárias. A pergunta orienta a fonte; a fonte não deve reconfigurar silenciosamente a pergunta.

Em estudos assistidos por IA, é comum pedir ao modelo que recomende bases. Essa prática pode ser útil, desde que o *prompt* solicite justificativas, limites e condições de uso. Um *prompt* adequado não deve perguntar simplesmente 'qual base devo usar?'. Deve informar tema, período, idioma, unidade de análise, tipo de documento, finalidade do estudo e necessidade de campos específicos. A saída esperada deve comparar alternativas, indicar metadados disponíveis, apontar limitações e recomendar verificações. Sem essas restrições, a IA tende a oferecer respostas genéricas, reproduzindo hierarquias de prestígio de bases sem examinar o ajuste ao problema.

A comparação entre fontes deve considerar ao menos seis dimensões: cobertura, granularidade, completude, exportabilidade, atualidade e reprodutibilidade. Cobertura refere-se ao conjunto de documentos, áreas, idiomas e períodos incluídos. Granularidade refere-se ao nível de detalhe dos campos: autores, afiliações, identificadores, referências, financiamento, conceitos, assuntos, resumos. Completude diz respeito ao preenchimento efetivo desses campos. Exportabilidade envolve formatos, limites de download e disponibilidade de API. Atualidade trata da frequência de atualização. Reprodutibilidade refere-se à possibilidade de outros pesquisadores reconstruírem a coleta com os mesmos parâmetros.

A cobertura disciplinar é um ponto crítico. Áreas biomédicas tendem a estar fortemente representadas em PubMed e em bases comerciais, enquanto ciências humanas e sociais frequentemente dependem de livros, capítulos, periódicos locais e documentos em línguas diversas. Estudos em ciência da computação podem exigir atenção a anais de conferências, pois parte expressiva da comunicação relevante ocorre nesse formato. Engenharias podem demandar integração com patentes e normas técnicas. Educação, administração e políticas públicas podem demandar literatura cinzenta e repositórios institucionais. Uma coleta cientométrica responsável explicita essas diferenças para evitar que ausência na base seja confundida com ausência de produção científica.

Outro aspecto é a cobertura temporal. Algumas bases apresentam maior consistência a partir de determinados anos, especialmente em campos como afiliação de autores, referências citadas, identificadores persistentes e resumos. Uma série temporal longa pode incluir períodos com metadados menos completos. Isso afeta análises longitudinais de colaboração, internacionalização e impacto. Em vez de apenas definir um intervalo amplo, o pesquisador deve examinar a qualidade dos campos ao longo do tempo. A IA pode ajudar a gerar um plano de inspeção, mas não pode declarar a completude temporal sem consulta efetiva aos dados.

A cobertura geográfica e linguística também merece atenção. Estudos globais frequentemente privilegiam fontes que indexam periódicos internacionais, mas esse recorte pode sub-representar países periféricos, produção em língua nacional, periódicos emergentes e publicações de relevância local. A escolha de bases abertas não elimina automaticamente esse viés, pois a disponibilidade de metadados depende de infraestrutura editorial, registro de identificadores e integração com agregadores. Em projetos de política científica, essas lacunas devem ser tratadas como parte dos resultados metodológicos, e não como detalhe técnico. A pergunta não é apenas 'quantas publicações existem?', mas 'quais publicações se tornam visíveis por meio desta fonte?'.

A integração de múltiplas fontes pode aumentar a cobertura, mas cria problemas de deduplicação, harmonização e compatibilidade. Um mesmo artigo pode aparecer com títulos ligeiramente diferentes, variações de nome de autor, múltiplos identificadores, erros de ano ou diferenças de periódico. Combinar Scopus, Web of Science, Crossref, OpenAlex e repositórios institucionais exige regras explícitas de priorização. O DOI pode ser uma chave forte, mas nem sempre está presente ou correto. Título, ano, periódico e autores podem auxiliar, mas exigem tolerância a erros. A IA pode propor regras de deduplicação, porém a implementação deve ser testada em amostras verificadas.

Em bases comerciais, a exportação costuma ser mediada por interfaces gráficas e limites de registro por lote. Isso exige atenção à segmentação da coleta. Dividir por anos, áreas, tipos documentais ou intervalos de citação pode ser necessário, mas essa divisão deve ser registrada. Se o pesquisador exporta 2.000 registros por vez e depois combina arquivos, deve preservar os arquivos originais, nomeá-los de forma informativa e registrar o total obtido em cada etapa. A perda de um lote ou a duplicação de outro pode alterar indicadores de produtividade, redes e distribuições temporais. A coleta é uma cadeia de custódia documental.

Em bases abertas, a coleta via API permite automatizar consultas, mas também introduz decisões técnicas: paginação, cursor, filtros, campos selecionados, limites de taxa, tratamento de erros, tentativas repetidas, datas de atualização e armazenamento de respostas brutas. Uma coleta programável não é automaticamente reprodutível se o código não registra parâmetros, versões e logs. Em bases dinâmicas, repetir a mesma consulta meses depois pode produzir resultados diferentes. Portanto, além de registrar o código, é recomendável preservar *snapshots*, arquivos *JSON* originais ou listas de identificadores recuperados.

Quadro 8.1 - Tipos de fontes

Tipo de fonte	Exemplos	Potencial para cientometria	Cuidados metodológicos
Comercial multidisciplinar	Web of Science, Scopus	Curadoria, filtros, citações, categorias, exportações padronizadas.	Acesso por assinatura, limites de exportação, cobertura seletiva, restrições de redistribuição.

Tipo de fonte	Exemplos	Potencial para cientometria	Cuidados metodológicos
Aberta ou comunitária	Crossref, OpenAlex, Lens, OpenCitations, PubMed	Reprodutibilidade, APIs, acesso amplo, integração com identificadores persistentes.	Completude variável, atualização dinâmica, diferenças de qualidade por campo e área.
Especializada ou regional	SciELO, LILACS, ERIC, IEEE Xplore, ACM Digital Library	Adequação temática, cobertura de comunidades específicas, anais e produção regional.	Sintaxes e formatos próprios; necessidade de tradução e harmonização.
Institucional	Repositórios, sistemas de pesquisa, currículos, catálogos internos	Visibilidade de produção local, relatórios institucionais, integração com políticas de gestão.	Heterogeneidade, duplicatas, preenchimento manual e necessidade de validação.
Complementar	Patentes, *grants*, ensaios clínicos, dados de pesquisa, documentos de política	Análise de circulação, inovação, financiamento e impacto além do artigo.	Integração complexa; diferenças conceituais entre objetos documentais.

Fonte: Elaborado pelo autor (2026).

A decisão entre fonte única e múltiplas fontes deve considerar a finalidade do estudo. Para um artigo metodológico que compara bases, múltiplas fontes são o objeto da análise. Para um diagnóstico institucional, pode ser necessário combinar fontes internacionais e internas. Para um mapa de uma área consolidada, uma fonte única pode ser suficiente se suas limitações forem compatíveis com a pergunta. Para uma avaliação de desempenho, a escolha deve ser ainda mais cautelosa, pois métricas baseadas em fontes diferentes podem produzir hierarquias divergentes. A transparência sobre a fonte é condição mínima de interpretação responsável.

Na prática, a IA pode apoiar uma matriz de decisão. O pesquisador pode solicitar que o modelo compare alternativas segundo critérios definidos, mas deve fornecer informações verificadas sobre campos necessários. Por exemplo: 'preciso de referências citadas para acoplamento bibliográfico', 'preciso de afiliações completas para análise por instituição', 'preciso de resumos para classificação temática', 'preciso de identificadores ORCID para desambiguação'. A resposta da IA deve ser tratada como hipótese operacional a ser confirmada na documentação da base e em uma coleta piloto. Essa distinção entre sugestão e validação precisa ser explícita no protocolo.

A coleta bibliográfica também deve respeitar questões legais e éticas. Bases comerciais têm termos de uso que podem restringir redistribuição de dados, mineração em massa ou uso automatizado. APIs abertas têm políticas de uso justo, requisitos de identificação de usuário, limites de requisição e condições de citação. Repositórios institucionais podem conter dados pessoais sensíveis ou metadados administrativos não destinados à divulgação ampla. Estudos cientométricos não são isentos de responsabilidade ética apenas porque lidam com documentos públicos. A extração, o armazenamento e o compartilhamento devem obedecer às normas da fonte e às políticas da instituição do pesquisador.

Por fim, a escolha da fonte deve ser apresentada no relatório como parte da interpretação. Um mapa temático baseado em uma base com pouca literatura regional deve ser interpretado como mapa daquela fonte, não como mapa universal do campo. Uma análise de colaboração baseada em afiliações incompletas deve reconhecer essa limitação. Uma rede de citação

baseada em referências abertas deve explicitar a cobertura de citações disponíveis. Essa postura não enfraquece o estudo; ao contrário, aumenta sua credibilidade, porque torna visíveis as condições de produção dos resultados.

8.3 Metadados essenciais: título, resumo, palavras-chave, autores, afiliações, referências e citações

Metadados são os elementos descritivos que tornam os documentos analisáveis. Em cientometria, os metadados não são acessórios: eles constituem a matéria-prima dos indicadores, redes e mapas. Um título permite recuperar termos e identificar temas; um resumo amplia a análise textual; palavras-chave favorecem coocorrências e classificação; nomes de autores sustentam produtividade e coautoria; afiliações permitem análise institucional e geográfica; referências citadas viabilizam cocitação e acoplamento; citações recebidas alimentam indicadores de impacto; identificadores persistentes conectam registros entre sistemas. Quando esses campos estão ausentes ou inconsistentes, a capacidade interpretativa do estudo é reduzida.

A primeira decisão é distinguir metadados mínimos, desejáveis e específicos. Metadados mínimos são aqueles sem os quais o estudo perde sua unidade básica: título, ano, tipo de documento, fonte, autores e identificador quando disponível. Metadados desejáveis incluem resumo, palavras-chave, afiliações, país, idioma, área, referências, DOI, ORCID, ROR, financiador e acesso aberto. Metadados específicos dependem da pergunta: ensaios clínicos podem exigir identificador de registro; estudos de inovação podem requerer vínculos com patentes; análises de políticas científicas podem incluir financiamento; estudos de circulação social podem integrar altmetria. A coleta deve começar por uma lista explícita de campos necessários.

O título é frequentemente tratado como campo simples, mas possui grande importância. Ele é usado em deduplicação, mineração de termos, classificação temática e verificação manual. Títulos podem aparecer em caixa alta, com caracteres especiais, sem subtítulos ou com tradução. Em bases multilíngues, um mesmo documento pode ter título original e título em inglês. Ao coletar dados, é recomendável preservar o título tal como exportado e, se necessário, criar um campo normalizado separado para análise. A IA pode auxiliar na tradução ou padronização, mas o texto original deve ser mantido para rastreabilidade.

O resumo é um dos campos mais valiosos para análise textual. Ele fornece contexto semântico mais rico que o título e permite classificar métodos, objetos, populações, tecnologias, abordagens teóricas e aplicações. No entanto, resumos nem sempre estão disponíveis, especialmente em registros antigos, livros, anais, bases com restrições de exportação ou metadados derivados de DOI. A ausência de resumo deve ser quantificada antes de se aplicar classificação automática. Se apenas parte do *corpus* tem resumos, análises temáticas baseadas nesse campo podem representar melhor os documentos mais recentes ou os periódicos com metadados mais completos.

Palavras-chave podem ser autorais, indexadas ou extraídas automaticamente. Palavras-chave autorais refletem escolhas dos autores e podem expressar vocabulários emergentes, modismos ou preferências disciplinares. Palavras-chave indexadas, como descritores controlados, favorecem consistência e recuperação. Termos gerados automaticamente, por sua vez, podem ampliar cobertura, mas introduzem decisões algorítmicas que precisam ser compreendidas. Em um estudo cientométrico, é importante separar esses tipos quando a base permite. Combinar indiscriminadamente palavras-chave de origens distintas pode produzir redes de coocorrência artificiais ou superestimar a centralidade de termos controlados.

Autores são unidades centrais para análises de produtividade, colaboração e trajetória científica. No entanto, nomes de autores apresentam problemas conhecidos: homônimos, mudanças de sobrenome, abreviações, transliteração, ordem variável de nomes, inconsistências em acentos e ausência de identificadores. Identificadores como ORCID ajudam, mas não estão presentes em todos os registros e podem estar incorretos ou incompletos. A coleta deve preservar o nome conforme a fonte, mas o estudo precisa prever procedimentos de desambiguação. A IA pode sugerir possíveis agrupamentos, porém decisões sobre identidade autoral exigem validação por evidências externas quando o resultado for sensível.

Afiliações são essenciais para mapear instituições, países e colaboração internacional. Porém, esse campo é um dos mais problemáticos. Um registro pode conter endereço institucional completo, apenas organização, múltiplas unidades, departamento, hospital, universidade, instituto, cidade e país. A ligação entre autor e afiliação nem sempre é completa em registros antigos. Estudos sobre qualidade de dados mostram que a completude de vínculos autor-afiliação varia ao longo do tempo e entre índices. Assim, antes de produzir rankings institucionais ou mapas geográficos, o pesquisador deve verificar se a base informa afiliações no nível necessário e se os autores estão corretamente conectados às instituições.

Referências citadas constituem a base de análises de cocitação, acoplamento bibliográfico, genealogias intelectuais e influência cognitiva. Sem referências, é possível fazer estudos de produção e coautoria, mas não análises robustas da estrutura intelectual baseada em relações bibliográficas. A cobertura de referências varia entre bases e períodos. Algumas fontes incluem listas completas; outras incluem apenas metadados parciais; outras disponibilizam referências apenas quando abertas pelos editores. A coleta deve registrar não apenas se há referências, mas quantas referências por documento foram recuperadas, qual proporção de documentos possui referências e quais tipos documentais apresentam lacunas.

Citações recebidas são frequentemente usadas como indicadores de impacto, mas dependem da fonte. O mesmo artigo pode ter contagens diferentes em Scopus, Web of Science, Dimensions, Google Scholar, OpenAlex ou COCI. Essas diferenças derivam da cobertura de documentos citantes, do tratamento de versões, da inclusão de literatura cinzenta, da atualização temporal e da qualidade dos links de citação. Portanto, a coleta deve registrar a data da extração das citações e a fonte usada. Em análises avaliativas, recomenda-se evitar a mistura de contagens de diferentes fontes sem justificativa, pois a comparabilidade pode ser comprometida.

Identificadores persistentes são elementos de integração. DOI, PMID, PMCID, arXiv ID, ISBN, ISSN, ORCID, ROR, GRID, Wikidata ID e identificadores internos de bases permitem vincular registros, deduplicar, enriquecer metadados e construir *pipelines*. O DOI é central para artigos, mas nem todo documento o possui. ORCID identifica autores, mas sua presença é desigual. ROR identifica organizações e pode apoiar normalização institucional. Identificadores de periódicos ajudam a tratar mudanças de título e variações de ISSN. A coleta deve incluir todos os identificadores disponíveis, mesmo que a análise inicial não os utilize, pois eles podem ser decisivos em etapas posteriores.

Quadro 8.2 - Metadados da pesquisa bibliográfica

Campo	Uso cientométrico principal	Risco quando ausente ou incompleto	Verificação recomendada
Título	Busca, deduplicação, análise textual e verificação manual.	Perda de contexto, duplicação difícil, erros de classificação.	Checar caracteres, idioma, subtítulos e variações.

Campo	Uso cientométrico principal	Risco quando ausente ou incompleto	Verificação recomendada
Resumo	Classificação temática, extração de métodos e objetos de estudo.	Viés de análise textual para registros mais completos.	Calcular percentual de registros com resumo.
Palavras-chave	Coocorrência, mapa temático, emergência de termos.	Rede artificial ou pobre, mistura de vocabulários distintos.	Separar autorais, indexadas e automáticas.
Autores	Produtividade, coautoria, trajetórias e redes.	Homônimos, fragmentação de perfis e rankings incorretos.	Usar ORCID quando disponível e validar nomes ambíguos.
Afiliações	Instituições, países, colaboração internacional.	Mapas institucionais distorcidos e atribuição incorreta.	Verificar vínculo autor-afilição e normalizar organizações.
Referências	Cocitação, acoplamento bibliográfico e estrutura intelectual.	Impossibilidade de mapear relações bibliográficas robustas.	Medir cobertura por documento e por período.
Citações recebidas	Impacto, influência e circulação acadêmica.	Indicadores dependentes da fonte e da data de coleta.	Registrar fonte, data e janela temporal.
Identificadores	Integração, deduplicação e enriquecimento.	Dificuldade de combinar fontes e rastrear registros.	Preservar DOI, PMID, ORCID, ROR, ISSN e IDs internos.

Fonte: Elaborado pelo autor (2026).

O tipo de documento é um campo interpretativo. Artigos, revisões, editoriais, cartas, anais, capítulos, livros, *preprints*, erratas e documentos de dados têm papéis distintos na comunicação científica. Misturar todos os tipos sem controle pode distorcer indicadores de produtividade e impacto. Revisões tendem a receber mais citações; anais são centrais em algumas áreas; editoriais podem inflar produção sem equivaler a pesquisa original; *preprints* podem duplicar artigos publicados. A coleta deve registrar tipos documentais e definir, no protocolo, se todos serão incluídos ou se algum será analisado separadamente.

Campos de fonte, como periódico, conferência, editora, volume, número e páginas, permitem análises de canais de comunicação. Em estudos de dispersão, leis bibliométricas e núcleo de periódicos, esses campos são essenciais. Porém, nomes de periódicos podem aparecer com abreviações, mudanças históricas e erros. ISSN e eISSN ajudam, mas também podem apresentar lacunas. Em anais de congressos, o nome do evento pode variar por edição. A IA pode apoiar a padronização de nomes, mas a normalização deve ser documentada e, quando possível, baseada em identificadores reconhecidos.

Campos de financiamento e agradecimentos permitem análises de políticas científicas, redes de financiamento e relações entre investimento e produção. Nem todas as bases exportam esses campos, e sua completude depende de práticas editoriais. Em estudos de avaliação de programas, a ausência de dados de financiamento pode impedir conclusões sobre retorno de investimento. Quando disponíveis, identificadores de financiadores e números de projetos devem ser preservados. Em bases institucionais, esses dados podem ser mais completos que em bases internacionais, mas exigem tratamento de confidencialidade e padronização.

Metadados de acesso aberto, licença e disponibilidade de texto completo podem apoiar análises sobre circulação, abertura e democratização do conhecimento. *Open access*

status, repositórios, licenças *Creative Commons* e links para PDFs são campos relevantes para estudos de ciência aberta. Contudo, o status de acesso aberto pode mudar ao longo do tempo e variar entre fontes. A coleta deve registrar a data da consulta e a origem do status. A IA pode auxiliar na interpretação de padrões de abertura, mas não deve inferir licença ou disponibilidade quando o campo não foi coletado.

A completude de metadados deve ser tratada quantitativamente (DELGADO-QUIRÓS; ORTEGA, 2024). Antes de calcular indicadores finais, é recomendável produzir uma tabela de preenchimento por campo: número de registros, percentual com DOI, percentual com resumo, percentual com afiliação, percentual com referências, percentual com palavras-chave, percentual com identificadores de autor, percentual com país. Essa inspeção é etapa de controle de qualidade. Se o campo de afiliação está preenchido em apenas 40% dos registros, qualquer mapa institucional deve ser interpretado com cautela. Se referências estão ausentes em metade do *corpus*, redes de cocitação serão parciais.

A inteligência artificial pode ser usada para construir checklist*s* de metadados a partir da pergunta. Por exemplo, se o pesquisador informa que deseja mapear colaboração internacional em inteligência artificial aplicada à saúde, a IA pode sugerir que campos de afiliação, país, autor, ano, tipo documental e área são indispensáveis. Se o objetivo é analisar frentes de pesquisa, a IA pode indicar títulos, resumos, palavras-chave, referências e citações. Essa utilidade depende da qualidade do *prompt* e da validação do pesquisador. A decisão final deve resultar de uma matriz entre pergunta, análise pretendida e campos efetivamente disponíveis.

Outro uso relevante da IA é a inspeção de amostras de metadados. Ao fornecer uma pequena amostra exportada, o pesquisador pode pedir que o modelo identifique campos ausentes, inconsistências aparentes, variações de nomes, problemas de codificação e possíveis riscos para análises planejadas. Esse procedimento não substitui scripts de validação, mas ajuda a transformar dados brutos em perguntas de qualidade. A IA pode apontar que há registros sem DOI, que algumas afiliações estão truncadas, que palavras-chave aparecem em idiomas mistos ou que referências foram exportadas em campo único não estruturado.

A relação entre metadados e interpretação precisa ser destacada. Uma análise cientométrica não interpreta documentos diretamente; interpreta representações documentais mediadas por metadados. Se essas representações são incompletas, a interpretação também será limitada. Por isso, a seção metodológica de um estudo deve informar os campos utilizados, a taxa de completude dos principais campos e os procedimentos de limpeza. A coleta de dados bibliográficos, portanto, é uma etapa de construção do objeto de pesquisa, e não uma operação passiva de captura.

8.4 Exportação de registros e formatos de arquivo

A exportação de registros é o momento em que os dados deixam a interface da fonte e passam a compor o acervo empírico do pesquisador. Essa transição parece simples, mas é uma das etapas em que mais se perdem campos, relações e informações contextuais. A escolha do formato de arquivo influencia a capacidade de importar dados em softwares bibliométricos, scripts de análise, planilhas e bancos relacionais. Um mesmo conjunto de registros pode resultar em análises diferentes se exportado em formato completo, BibTeX, RIS, CSV, plain text, JSON ou XML. Por isso, a exportação deve ser planejada antes da coleta em larga escala.

Os formatos bibliográficos clássicos, como BibTeX e RIS, são amplamente compatíveis com gerenciadores de referências e ferramentas de análise. Eles preservam campos básicos como título, autores, ano, fonte e DOI, mas podem perder relações complexas, afiliações

estruturadas ou referências citadas, dependendo da base e da configuração de exportação. Em bibliometrix, por exemplo, a função de conversão depende de fonte e formato específicos (ARIA; CUCCURULLO, 2017). Em VOSviewer e CiteSpace, determinados campos são necessários para tipos particulares de mapa (VAN ECK; WALTMAN, 2010; CHEN, 2006). O pesquisador deve consultar previamente quais formatos são aceitos pela ferramenta analítica escolhida.

Arquivos CSV e Excel são úteis para inspeção, limpeza e análise exploratória, mas podem apresentar problemas de codificação, separadores, campos multivalorados e quebras de linha. Autores, palavras-chave e referências podem aparecer em uma única célula separados por ponto e vírgula, vírgula ou delimitadores próprios. Se o delimitador também aparece no conteúdo, a estrutura pode ser corrompida. Planilhas ainda impõem riscos de transformação automática de campos, como conversão de identificadores em notação científica ou alteração de datas. Por essa razão, arquivos originais não devem ser editados diretamente em planilhas; é melhor criar cópias de trabalho.

Formatos JSON e XML preservam estruturas mais complexas e são comuns em APIs. Eles permitem representar listas de autores, objetos de afiliação, identificadores, referências e relações de maneira hierárquica. Essa riqueza, entretanto, exige scripts de extração e transformação. Um erro na normalização pode achatar indevidamente relações, por exemplo, associando todos os autores a todas as instituições ou perdendo a ordem autoral. Na coleta via API, é recomendável armazenar a resposta bruta e gerar, separadamente, tabelas analíticas derivadas. Assim, se houver erro na transformação, os dados originais permanecem disponíveis.

Plain text é utilizado por algumas bases para exportações compatíveis com softwares bibliométricos. Pode preservar campos codificados por *tags*, mas também pode sofrer com truncamento, problemas de acentuação e limites de lote. O pesquisador deve verificar se o arquivo contém todos os registros esperados e se campos como referências citadas foram incluídos. Em bases comerciais, a exportação completa de referências pode exigir a seleção de opções específicas. Uma coleta apressada que exporta apenas registro básico pode inviabilizar análises relacionais planejadas.

A exportação deve ser precedida por uma coleta piloto. A amostra piloto, com 20 a 100 registros, permite verificar campos disponíveis, codificação, estrutura de autores, afiliações, palavras-chave, referências e compatibilidade com ferramentas. A coleta piloto é especialmente importante quando se pretende combinar fontes ou usar IA para classificar textos. Antes de solicitar ao modelo uma análise de milhares de resumos, é preciso verificar se os resumos foram de fato exportados e se estão legíveis. Antes de construir uma rede de acoplamento, é preciso confirmar se as referências foram recuperadas em formato utilizável.

A nomenclatura dos arquivos é parte da metodologia. Arquivos como 'export1.csv' ou 'dados_final_final.xlsx' dificultam reprodutibilidade. Recomenda-se nomear arquivos com fonte, tema, data, recorte e lote: por exemplo, 'scopus_ia_saude_2014-2024_lote01_2026-05-30.ris'. Esse padrão permite reconstruir a sequência de coleta. Quando uma busca é segmentada por anos, cada lote deve registrar intervalo e número de registros. Um arquivo de manifesto deve listar todos os arquivos coletados, suas datas, consultas, número de registros e observações.

A preservação dos arquivos originais é regra fundamental. Após exportar, o pesquisador deve manter uma pasta de dados brutos, sem edições manuais, e uma pasta de dados processados. Alterações, deduplicações, normalizações e enriquecimentos devem ocorrer em cópias, preferencialmente por scripts. Essa prática permite auditar decisões, recuperar informações perdidas e responder a questionamentos de revisão. Em estudos assistidos por IA, os arquivos originais são ainda mais importantes, pois ajudam a verificar se uma interpretação gerada por modelo corresponde aos dados efetivamente coletados.

A conversão entre formatos deve ser documentada. Se um arquivo *BibTeX* foi convertido em *data frame*, se um JSON foi achatado em CSV, se um conjunto de RIS foi importado em bibliometrix ou se uma planilha foi transformada em grafo, cada transformação deve ser registrada. O ideal é que scripts de conversão sejam preservados junto aos dados. Quando a conversão é feita por interface gráfica, o pesquisador deve registrar versão do software, opções selecionadas e eventuais mensagens de erro. A IA pode ajudar a redigir esse registro, mas deve receber informações concretas sobre o procedimento executado.

Problemas de codificação de caracteres são frequentes em dados bibliográficos internacionais. Nomes com acentos, caracteres chineses, árabes, cirílicos, sinais matemáticos e símbolos químicos podem ser corrompidos se o formato ou a codificação forem inadequados. A perda desses caracteres afeta buscas, deduplicação, identificação de autores e análise linguística. Sempre que possível, deve-se usar UTF-8 e verificar amostras multilíngues. Em documentos com nomes e títulos em diferentes idiomas, a preservação de caracteres originais é requisito de qualidade e inclusão.

Campos multivalorados exigem atenção. Um registro pode ter vários autores, várias afiliações, várias palavras-chave, vários financiadores e dezenas de referências. Em formatos tabulares, esses campos são frequentemente concatenados. Se o objetivo é construir redes, será necessário separar os valores corretamente. O delimitador deve ser conhecido e documentado. A IA pode sugerir regras de separação, mas deve-se testar em amostras com nomes compostos, instituições com vírgulas e palavras-chave que contenham pontuação. Um erro de *parsing* pode criar autores inexistentes ou quebrar nomes institucionais.

A exportação de referências citadas é particularmente sensível. Algumas bases exportam referências como *strings* bibliográficas não estruturadas; outras fornecem DOIs, autores, ano e título; outras incluem apenas referências abertas. O formato de referência determina a qualidade de cocitação e acoplamento. Referências não estruturadas exigem algoritmos de correspondência para identificar documentos citados. Quando o estudo depende de relações de citação, o protocolo deve informar como as referências foram obtidas, qual sua estrutura e como foram tratadas. A ausência dessa informação compromete a avaliação do estudo.

A exportação de afiliações também varia. Algumas bases fornecem afiliação por autor; outras listam endereços separados sem ligação explícita; outras fornecem apenas instituição agregada. Para análises de colaboração, a ligação autor-afiliação é crucial. Se ela não está presente, atribuir países ou instituições a autores pode exigir suposições. Essas suposições devem ser evitadas ou claramente documentadas. A IA não deve ser usada para inferir automaticamente afiliações sem validação, especialmente em estudos avaliativos ou institucionais.

Em estudos de larga escala, limites de exportação podem obrigar o pesquisador a dividir buscas. Essa divisão deve evitar sobreposição ou lacunas. Por exemplo, segmentar por ano é mais seguro que segmentar por relevância, pois ranking de relevância pode mudar. Segmentos por tipo documental ou área exigem atenção se um registro puder pertencer a múltiplas categorias. Em qualquer caso, o número total de registros na interface deve ser comparado com o total exportado e importado. Diferenças precisam ser explicadas.

A documentação de exportação também deve registrar filtros aplicados na interface: anos, áreas, idiomas, tipos de documento, índices, coleções, acesso aberto, país, instituição ou fonte. Muitas falhas de reprodutibilidade decorrem de relatórios que informam a *string* de busca, mas omitem filtros. Em bases complexas, filtros podem alterar drasticamente o *corpus*. O pesquisador deve registrar capturas de tela, URLs parametrizadas quando disponíveis, ou relatórios de busca. A IA pode transformar essas informações em texto metodológico, mas não deve inventar filtros não registrados.

Quadro 8.3 - Formatos de exportação de metadados

Formato	Vantagens	Limitações	Uso recomendado
BibTeX	Compatível com gerenciadores e ferramentas bibliométricas; preserva campos bibliográficos básicos.	Pode perder campos complexos; estrutura varia por fonte.	Importação em bibliometrix, gestão de referências e análises bibliográficas tradicionais.
RIS	Amplo suporte em softwares; fácil intercâmbio.	Campos de citação e afiliação podem ser limitados.	Troca entre bases, gerenciadores e ferramentas de revisão.
CSV/ XLSX	Facilita inspeção visual, planilhas e análises tabulares.	Risco de codificação, delimitadores, alterações automáticas e perda de estrutura.	Controle de qualidade, relatórios e integração com scripts simples.
JSON	Preserva hierarquias, listas e objetos complexos; comum em APIs.	Exige transformação programática; pode ser grande e aninhado.	Coleta via API, preservação de dados brutos e *pipelines* automatizados.
XML	Estruturado e rico; comum em PubMed e repositórios.	Processamento mais técnico; campos aninhados.	Extrações sistemáticas e interoperabilidade com padrões específicos.
Plain text	Frequentemente aceito por ferramentas bibliométricas.	Pode sofrer truncamento e problemas de codificação.	Exportações de bases comerciais quando configuradas para registros completos.

Fonte: Elaborado pelo autor (2026).

Para projetos colaborativos, recomenda-se padronizar a estrutura de pastas e nomes de arquivos desde o início. Uma pasta pode conter '00_protocolo', '01_buscas', '02_dados_brutos', '03_dados_processados', '04_scripts', '05_logs', '06_figuras' e '07_relatorios'. Essa organização reduz ambiguidade e facilita auditoria. Em equipes, deve-se definir quem executa a coleta, quem valida a importação, quem aprova a deduplicação e quem documenta alterações. A coleta não deve depender da memória individual do operador.

Por fim, a exportação deve ser considerada bem-sucedida apenas após validação. Validar significa verificar contagem de registros, campos esperados, legibilidade, duplicatas evidentes, presença de campos críticos e compatibilidade com a ferramenta analítica. Não basta salvar o arquivo. Em cientometria com IA, essa validação pode ser parcialmente apoiada por *prompts* que solicitam checklists e diagnósticos, mas a verificação quantitativa deve ser feita sobre os dados reais. A qualidade do capítulo metodológico de um estudo começa na disciplina da exportação.

8.5 Uso de APIs em estudos cientométricos

APIs, ou interfaces de programação de aplicações, permitem que sistemas consultem bases bibliográficas de maneira programável. Em vez de baixar registros manualmente pela interface gráfica, o pesquisador constrói consultas por parâmetros, recebe respostas estruturadas e transforma essas respostas em arquivos analisáveis. Em cientometria, APIs são particularmente úteis para coletas de larga escala, atualizações periódicas, enriquecimento de metadados, integração de fontes e preservação de fluxos reprodutíveis. Entretanto, o

uso de APIs exige competência técnica e atenção a limites, licenças, políticas de uso e estabilidade dos *endpoints*.

Uma API bibliográfica normalmente permite consultar obras por termo, título, DOI, autor, instituição, periódico, ano, área, tipo de documento ou combinação de filtros. A resposta pode vir em *JSON*, *XML* ou outro formato estruturado. A vantagem é que a consulta pode ser documentada no próprio código. A desvantagem é que pequenos erros de parâmetro podem produzir conjuntos incompletos. Além disso, APIs diferem em sintaxe, autenticação, limite de requisições, campos retornados, paginação e políticas de atualização. Assim como uma *string* booleana, uma consulta de API é um artefato metodológico que deve ser revisado.

Crossref, por exemplo, disponibiliza metadados depositados por seus membros e pode ser consultado por DOI, título, data, tipo de obra, financiador e outros parâmetros (HENDRICKS *et al.*, 2020). OpenAlex oferece uma estrutura ampla de obras, autores, instituições, fontes e conceitos, com filtros e paginação por cursor (CULBERT *et al.*, 2025). PubMed fornece mecanismos de busca e recuperação por meio das utilidades E-utilities, relevantes para áreas biomédicas. OpenCitations disponibiliza dados abertos de citações. Dimensions e Scopus também têm APIs, em geral com condições de acesso específicas. Cada fonte exige leitura da documentação oficial antes da coleta.

A primeira etapa de uma coleta via API é definir a pergunta e os campos necessários. Não se deve começar pelo código. O pesquisador deve responder: que documentos serão recuperados? Quais filtros definem o *corpus*? Que campos são essenciais? A coleta será única ou recorrente? É necessário preservar resposta bruta? Haverá enriquecimento com outras fontes? Como serão tratados erros e limites de requisição? Somente depois dessas respostas é que se deve escrever pseudocódigo. A IA pode ajudar a transformar o protocolo em lógica computacional, mas não deve determinar silenciosamente os filtros.

A paginação é um aspecto técnico com implicações metodológicas. Muitas APIs retornam resultados em páginas ou lotes. Se o código recupera apenas a primeira página, o *corpus* fica truncado. Em bases grandes, a paginação por cursor é mais estável que a paginação por deslocamento, mas exige armazenar o cursor e repetir chamadas até o fim. O log deve registrar número de páginas, total estimado, total efetivamente coletado e eventuais interrupções. Uma falha de conexão no meio da coleta pode produzir lacunas invisíveis se o código não for robusto.

Limites de taxa, ou *rate limits*, também precisam ser respeitados. APIs abertas frequentemente impõem número máximo de requisições por segundo ou por dia para garantir estabilidade. Ignorar esses limites pode resultar em bloqueio, respostas incompletas ou comportamento antiético em relação à infraestrutura comunitária. O protocolo deve incluir pausas, identificação do usuário quando recomendada e tratamento de respostas de erro. A IA pode gerar pseudocódigo com pausas e repetição controlada, mas o pesquisador deve verificar as regras da API no momento da coleta.

Campos retornados por APIs podem ser configuráveis. Algumas APIs permitem selecionar apenas campos necessários, reduzindo volume de dados; outras retornam objetos completos. Selecionar poucos campos melhora eficiência, mas pode impedir análises futuras. Uma prática equilibrada é preservar respostas brutas completas quando o volume permitir e produzir tabelas derivadas com campos analíticos. Em coletas muito grandes, a seleção de campos deve ser justificada. O importante é que a decisão não seja acidental. Se referências ou afiliações não foram coletadas, o relatório deve informar essa limitação.

O enriquecimento de metadados é uma aplicação comum de APIs. Um conjunto exportado de uma base pode ser enriquecido com DOI via Crossref, status de acesso aberto

via Unpaywall, identificadores institucionais via ROR, conceitos via OpenAlex ou citações abertas via OpenCitations. Esse enriquecimento pode melhorar a análise, mas também cria um *corpus* híbrido. O pesquisador deve registrar a fonte original e a fonte de enriquecimento de cada campo. Caso contrário, não será possível saber se uma afiliação veio da base primária, de um serviço externo ou de inferência posterior.

A deduplicação por API pode utilizar identificadores. Quando dois registros têm o mesmo DOI, é provável que representem o mesmo documento, embora versões e erros devam ser verificados. Quando não há DOI, algoritmos podem consultar APIs por título e ano para encontrar correspondências. A IA pode sugerir lógica de correspondência, mas resultados de *matching* devem ser avaliados por precisão e *recall*. Falsos positivos podem fundir documentos distintos; falsos negativos podem manter duplicatas. Em estudos avaliativos, a deduplicação deve ser conservadora e auditável.

APIs também permitem atualização incremental. Um estudo institucional pode coletar mensalmente novas publicações de autores ou organizações. Um observatório científico pode atualizar séries temporais. Nesses casos, o protocolo deve registrar janela de coleta, data de execução, parâmetros, versão do código e diferenças entre execuções. Atualizações incrementais exigem chaves estáveis para evitar recontagem. Se registros antigos são atualizados com novos metadados, o sistema deve distinguir novas obras de obras modificadas. A coleta deixa de ser evento único e passa a ser processo contínuo de governança de dados.

O uso de APIs deve considerar a estabilidade dos serviços. *Endpoints* podem mudar, campos podem ser renomeados, políticas podem ser atualizadas e bases podem alterar cobertura. A reprodutibilidade de longo prazo depende de registrar não apenas o código, mas também versões, documentação consultada, data de acesso e amostras de resposta. Quando possível, deve-se depositar o código e os dados derivados em repositórios com identificador persistente, respeitando licenças. Para dados que não podem ser redistribuídos, pode-se disponibilizar scripts e listas de identificadores, permitindo reconstrução por usuários autorizados.

Quadro 8.4 - Exportação de metadados por API

Elemento do protocolo de API	Descrição metodológica	Exemplo de registro
Fonte e endpoint	Identifica a base e o ponto de consulta usado.	OpenAlex works *endpoint*; Crossref works *endpoint*; PubMed ESearch/EFetch.
Parâmetros e filtros	Define *corpus* recuperado.	Tema, anos, tipo documental, idioma, instituição, fonte, DOI.
Paginação	Explica como todos os resultados foram percorridos.	Cursor salvo a cada lote; total de páginas; retomada em caso de erro.
Campos coletados	Lista metadados retornados e campos derivados.	Título, autores, DOI, ano, afiliações, referências, citações, resumo.
Controle de taxa	Registra pausas e limites respeitados.	Espera entre requisições, identificação por e-mail quando recomendada.
Logs	Permite auditoria de execução.	Data, hora, URL, status, número de registros, erros e tentativas.
Armazenamento	Preserva dados brutos e processados.	JSON original em pasta bruta; CSV derivado em pasta processada.

Elemento do protocolo de API	Descrição metodológica	Exemplo de registro
Validação	Confere completude e consistência.	Total esperado versus total coletado; amostra manual; campos ausentes.

Fonte: Elaborado pelo autor (2026).

A coleta via API não elimina a necessidade de conhecimento bibliográfico. Um script pode recuperar registros de forma eficiente, mas não sabe se a query representa corretamente um campo de pesquisa. Pode extrair milhares de documentos irrelevantes ou omitir termos essenciais. A validação por especialistas, revisão de amostras e comparação com literatura conhecida continuam necessárias. A IA deve ser integrada a esse ciclo como assistente de planejamento, documentação e revisão, não como substituto da competência informacional.

Um protocolo de API deve incluir: objetivo da coleta, fonte, *endpoint*, parâmetros, filtros, campos, autenticação, limites de taxa, paginação, data e hora de execução, versão do código, ambiente computacional, armazenamento de dados brutos, transformação em tabelas, validação de contagem, tratamento de erros e critérios de exclusão. Esses elementos permitem que outra pessoa compreenda a coleta mesmo sem executar o código. Em capítulos metodológicos de artigos, nem todos os detalhes cabem no texto principal, mas podem ser incluídos em material suplementar.

A IA pode gerar pseudocódigo útil quando recebe instruções detalhadas. Um bom *prompt* deve indicar a API, os filtros, o formato da saída, os campos desejados, a necessidade de logs, tratamento de erros e preservação dos dados brutos. Um mau *prompt*, como 'faça um código para baixar artigos', pode produzir código inseguro, incompleto ou incompatível com a documentação. O pesquisador deve pedir ao modelo que marque pontos dependentes da documentação oficial e que não invente parâmetros. Também deve testar o pseudocódigo em pequena escala antes de ampliá-lo.

A segurança e a privacidade não devem ser negligenciadas. Chaves de API não devem ser inseridas em *prompts* públicos nem salvas em notebooks compartilhados. Dados institucionais podem conter informações administrativas ou pessoais. Logs podem registrar URLs, tokens ou consultas sensíveis. O protocolo deve prever armazenamento seguro, variáveis de ambiente e remoção de credenciais antes do compartilhamento. A IA pode sugerir boas práticas, mas o pesquisador é responsável por proteger dados e cumprir termos de uso.

Em síntese, APIs transformam a coleta em processo mais escalável e reprodutível, mas ampliam a necessidade de documentação técnica. Elas são particularmente adequadas para projetos de monitoramento, enriquecimento, integração de fontes e coleta em grande escala. Contudo, não devem ser usadas apenas porque são sofisticadas. Uma coleta manual bem documentada pode ser mais adequada que uma coleta automatizada opaca. O critério deve ser aderência à pergunta, qualidade dos metadados, viabilidade técnica e transparência.

8.6 Coleta manual, semiautomática e automatizada

A coleta de dados bibliográficos pode ocorrer em diferentes níveis de automação. Em um extremo está a coleta manual, realizada por interface gráfica, seleção de registros e exportação direta. No outro está a coleta automatizada por scripts, APIs e *pipelines*. Entre esses extremos existe a coleta semiautomática, na qual o pesquisador combina interface gráfica, planilhas, scripts simples, ferramentas bibliométricas e *prompts*. A escolha do nível de automação deve

considerar escala, fonte, finalidade, competência técnica, restrições de acesso, necessidade de atualização e sensibilidade dos resultados.

A coleta manual é adequada para estudos pequenos, buscas exploratórias, fontes sem API, bases com limites de exportação ou situações em que o pesquisador precisa controlar visualmente cada etapa. Seu principal benefício é a proximidade com a interface e a possibilidade de verificar filtros, categorias e resultados antes da exportação. Seu principal risco é a dependência de ações humanas não registradas: clicar em filtros, alterar ordenação, selecionar campos incompletos, baixar arquivos com nomes ambíguos ou esquecer parâmetros. A coleta manual só é cientificamente aceitável quando documentada com rigor.

A coleta semiautomática é provavelmente a mais comum em estudos cientométricos aplicados. O pesquisador executa a busca na base, exporta arquivos, usa bibliometrix, VOSviewer, CiteSpace, planilhas ou scripts para importar e combinar registros, recorre a APIs para enriquecer campos e usa IA para gerar checklist*s* ou redigir protocolo. Esse modelo é flexível e permite que equipes com diferentes competências colaborem. Entretanto, a multiplicidade de ferramentas exige controle de versões e registro das transformações. Cada passagem entre ferramentas pode alterar campos, codificação e estrutura.

A coleta automatizada é indicada para grandes volumes, atualizações recorrentes, integração de fontes e observatórios. Ela reduz trabalho manual, aumenta consistência e permite repetição. Contudo, exige testes, tratamento de erros, logs, documentação e manutenção. Um script automatizado mal validado pode reproduzir sistematicamente um erro em milhares de registros. A automação não garante qualidade; apenas executa de forma eficiente a lógica definida. Portanto, o pesquisador deve validar a lógica antes de escalar, realizar coletas piloto e comparar resultados com amostras conhecidas.

A decisão entre manual, semiautomática e automatizada também depende da finalidade. Um estudo acadêmico de mapeamento de uma área pode usar uma coleta semiautomática bem documentada. Um painel institucional atualizado mensalmente precisa de automação. Uma revisão exploratória inicial pode começar manualmente. Um estudo com dados comerciais restritos pode depender de exportação manual por causa dos termos da plataforma. O importante é que o nível de automação seja justificado e não confundido com superioridade metodológica.

A IA pode apoiar todos os níveis. Na coleta manual, pode gerar checklist*s* de cliques, campos e filtros. Na coleta semiautomática, pode ajudar a transformar arquivos de diferentes fontes em uma matriz de compatibilidade, sugerir procedimentos de deduplicação e redigir logs. Na coleta automatizada, pode gerar pseudocódigo, revisar comentários, propor tratamento de erros e explicar scripts a membros da equipe. Em todos os casos, porém, a IA deve operar sobre informações fornecidas pelo pesquisador e resultados verificáveis. Não se deve aceitar como fato uma afirmação do modelo sobre a disponibilidade de determinado campo sem consulta à fonte.

Um ponto crítico é a coleta por *scraping*. Raspar páginas Web pode violar termos de uso, sobrecarregar servidores e produzir dados instáveis. Sempre que uma API ou exportação oficial existe, ela deve ser priorizada. Quando *scraping* é a única opção legítima, o pesquisador deve verificar permissões, respeitar robots.txt quando aplicável, controlar taxa de acesso, registrar código e evitar coleta de dados pessoais desnecessários. Em muitas situações, a recomendação metodológica é não raspar, mas buscar fontes oficiais ou acordos de acesso.

A coleta manual e a semiautomática podem se beneficiar de formulários de controle. Um formulário de coleta registra fonte, operador, data, *string* de busca, filtros, número de resultados, formato exportado, campos selecionados, nome do arquivo e observações. Esse

formulário reduz dependência da memória e facilita revisão. A IA pode converter anotações soltas em formulário estruturado, mas deve manter separação entre dados observados e comentários interpretativos. Um campo como 'observação' pode incluir 'a base indicou 3.245 resultados, mas apenas 3.200 foram exportados por limite de lote'.

Para coleta automatizada, logs são equivalentes aos formulários manuais. Um log deve registrar requisições, respostas, erros, tentativas, totais e tempo. Logs excessivamente detalhados podem ser difíceis de ler; logs escassos impedem auditoria. Um equilíbrio consiste em manter logs completos para execução e produzir resumos metodológicos para relatório. A IA pode ajudar a transformar logs técnicos em narrativa científica, desde que não altere números ou simplifique lacunas. A redação metodológica deve ser fiel aos registros.

A coleta em equipe requer governança. Deve-se definir papéis: responsável pela estratégia de busca, responsável pela execução da coleta, responsável pela validação dos arquivos, responsável pela deduplicação, responsável pela documentação e responsável por aprovar alterações no protocolo. Em projetos com IA, também é útil registrar quem elaborou *prompts*, quais respostas foram incorporadas e quais foram rejeitadas. Isso evita que decisões conversacionais fiquem invisíveis. A governança da coleta é parte da governança da pesquisa.

A coleta semiautomática pode incluir enriquecimento em lote. Por exemplo, uma planilha com DOIs pode ser enviada a um script que consulta *Crossref* ou OpenAlex para completar títulos, anos, periódicos ou referências. Esse procedimento deve preservar os valores originais e acrescentar campos novos com prefixos que identifiquem a fonte do enriquecimento. Substituir campos originais sem registro reduz rastreabilidade. Quando houver conflito entre fontes, o protocolo deve definir critérios: priorizar DOI, priorizar base primária, revisar manualmente ou manter ambos os valores.

A validação por amostragem é indispensável. Após a coleta, recomenda-se selecionar amostras aleatórias e amostras críticas: registros antigos, documentos sem DOI, autores com muitos registros, instituições centrais, documentos com muitas citações, registros em idiomas não ingleses e itens provenientes de cada fonte. A amostra deve ser conferida contra a fonte original. Isso permite estimar problemas de exportação, duplicação e perda de campos. A IA pode auxiliar a gerar o plano de amostragem e a tabela de checagem, mas a verificação deve ser feita nos dados e, quando necessário, nas plataformas.

A automação também deve ser testada quanto à estabilidade. Um script deve ser executado em pequena escala, depois em escala média, antes da coleta final. Deve lidar com falhas de rede, respostas vazias, registros sem campos esperados, caracteres especiais e limites de taxa. Deve salvar resultados progressivamente para evitar perda total em caso de interrupção. Deve permitir retomada. Esses elementos parecem técnicos, mas têm impacto científico, pois determinam se o *corpus* está completo e se a coleta pode ser reproduzida.

Coletas manuais frequentemente geram maior confiança subjetiva porque o pesquisador vê a interface. Coletas automatizadas frequentemente geram maior confiança técnica porque são replicáveis. Ambas as confianças podem ser ilusórias. A interface pode ocultar filtros ou arredondar totais; o script pode conter erro silencioso. A confiança deve vir da documentação, validação, amostragem e coerência entre totais. A IA pode reforçar ilusões ao produzir textos metodológicos fluidos. Por isso, a regra é: nenhum texto elegante substitui evidência de coleta.

O resultado ideal é uma coleta proporcional. Estudos pequenos não precisam de *pipelines* complexos, mas precisam de registro. Estudos grandes não devem depender de cliques manuais não auditados. Estudos institucionais devem equilibrar eficiência e responsabilidade. A cientometria com IA não exige automatizar tudo; exige tornar explícito o que foi automatizado, o que foi decidido por humanos, o que foi sugerido por IA e o que foi validado.

8.7 Registro do processo de coleta para reprodutibilidade

Reprodutibilidade, em coleta bibliográfica, significa permitir que outro pesquisador compreenda e, quando possível, reconstrua o conjunto de dados obtido. Em bases dinâmicas, a reprodução exata pode ser impossível, pois registros são adicionados, corrigidos ou removidos. Ainda assim, a rastreabilidade é obrigatória: devem ser preservados parâmetros, datas, versões, arquivos brutos e decisões. A reprodutibilidade não é um estado absoluto, mas um compromisso com transparência suficiente para avaliar a confiabilidade dos resultados.

O registro começa antes da coleta. O protocolo deve informar pergunta de pesquisa, fontes escolhidas, justificativa das fontes, período, tipos documentais, idiomas, áreas, critérios de inclusão e exclusão, campos necessários, estratégia de busca, formatos de exportação, ferramentas e plano de validação. Esse protocolo inicial pode ser atualizado, mas alterações devem ser registradas. Em estudos assistidos por IA, os *prompts* que influenciaram decisões metodológicas também devem ser arquivados, com data, modelo utilizado quando relevante e resumo da decisão adotada.

Durante a coleta, devem ser registrados os parâmetros executados. Em interface gráfica, isso inclui *string* de busca, filtros, coleções, índice ou sub-base, número de resultados, campos selecionados para exportação, formato, data e operador. Em API, inclui *endpoint*, parâmetros, paginação, status das respostas, limites de taxa, total estimado, total coletado, erros e tentativas. Em ambos os casos, capturas de tela, logs e arquivos de manifesto ajudam a reconstruir o processo. O registro deve ser suficientemente detalhado para revelar o que foi feito, não apenas o que se pretendia fazer.

Após a coleta, o registro deve incluir controle de qualidade. Quantos registros foram exportados? Quantos foram importados? Quantos foram excluídos por duplicidade? Quantos não tinham DOI? Quantos não tinham resumo? Quantos não tinham afiliação? Houve campos truncados? Houve problemas de codificação? Houve registros removidos por critérios de exclusão? Essas informações devem compor um relatório de coleta. Em artigos, podem ser resumidas no método; em relatórios técnicos, podem aparecer em apêndices.

A reprodutibilidade depende da preservação de dados brutos. Arquivos originais exportados não devem ser sobrescritos. Se as licenças impedirem compartilhamento público, ainda assim devem ser preservados internamente e descritos. Dados processados, scripts e metadados derivados podem ser compartilhados quando permitido. Para bases abertas, recomenda-se disponibilizar listas de identificadores e código. Para bases comerciais, listas de DOI, *strings* e procedimentos podem permitir que usuários autorizados repliquem a coleta. O grau de compartilhamento deve respeitar termos de uso.

Versionamento é outro elemento central. Em projetos com múltiplas etapas, recomenda-se usar controle de versão para scripts, protocolos e relatórios. Cada alteração na estratégia de busca, nos critérios de inclusão ou nas regras de deduplicação deve ser registrada. Arquivos de dados podem ser versionados por data e hash. O hash permite verificar se o arquivo foi alterado. A IA pode ajudar a criar um modelo de registro de versões, mas a disciplina de preenchimento é humana e institucional.

A documentação deve distinguir dados brutos, dados limpos, dados enriquecidos e dados analisados. Dados brutos são a exportação original. Dados limpos passaram por correções, normalização e deduplicação. Dados enriquecidos receberam campos de outras fontes. Dados analisados são tabelas, redes e matrizes usadas para resultados. Misturar esses níveis confunde auditoria. Um erro comum é salvar apenas a planilha final e perder a origem das transformações. Em cientometria com IA, essa perda é grave porque impede verificar se classificações ou interpretações foram baseadas em dados corretos.

O diário de pesquisa é útil para registrar decisões contextuais. Nele, o pesquisador anota porque escolheu determinada fonte, porque excluiu certos tipos documentais, por que alterou uma *string*, por que combinou bases, por que adotou uma regra de deduplicação e por que rejeitou uma sugestão de IA. Essas decisões raramente aparecem integralmente no artigo final, mas são fundamentais para a integridade do projeto. O diário pode ser textual, planilhado ou integrado a um repositório de projeto.

Prompts utilizados na coleta devem ser documentados com o mesmo rigor que scripts. Um *prompt* que gerou uma checklist de metadados, uma matriz de comparação de bases ou um pseudocódigo de API faz parte do ambiente metodológico. Registrar o *prompt* não significa que a IA determinou a decisão; significa que sua participação é transparente. O registro deve incluir a entrada fornecida, a saída relevante, a decisão humana posterior e a forma de validação. Essa prática previne a naturalização de sugestões não verificadas.

A documentação também deve indicar limitações não resolvidas. Se uma base não exportou referências para parte dos registros, isso deve ser declarado. Se afiliações antigas estão incompletas, deve ser informado. Se resumos não estavam disponíveis em determinada fonte, isso deve aparecer. Se a coleta via API foi interrompida e retomada, o log deve mostrar. Limitações transparentes não invalidam necessariamente o estudo; limitações ocultas comprometem sua interpretação.

Em pesquisas com múltiplas fontes, a reprodutibilidade depende de uma matriz de proveniência. Cada registro deve manter informação sobre a fonte de origem e, quando houver fusão, sobre as fontes que contribuíram para cada campo. Um DOI pode vir de Crossref, uma afiliação de Scopus, um conceito de OpenAlex e uma citação de OpenCitations. Sem proveniência, não é possível auditar conflitos. Uma abordagem recomendada é manter colunas com prefixos ou tabelas relacionais separadas por fonte antes da harmonização final.

A criação de um dicionário de dados é recomendável. O dicionário descreve cada campo, sua fonte, tipo, formato, possíveis valores, tratamento aplicado e uso analítico. Por exemplo, 'publication_year' pode ser ano de publicação da fonte primária; 'cited_by_count' pode ser contagem da fonte X na data Y; 'institution_normalized' pode ser campo criado por regra de normalização; 'topic_ai' pode ser categoria atribuída por IA e validada por amostra. Esse dicionário é crucial para equipes, revisores e reutilização futura.

A reprodutibilidade também inclui ambiente computacional. Versões de Python, R, pacotes, bibliometrix, VOSviewer, bibliotecas de API e sistemas operacionais podem afetar resultados (ARIA; CUCCURULLO, 2017; VAN ECK; WALTMAN, 2010). Para capítulos metodológicos, pode ser suficiente informar as versões principais; para repositórios, recomenda-se arquivo de dependências. Em IA, quando possível, deve-se registrar modelo, data, parâmetros relevantes e instruções de sistema disponíveis ao usuário. Como modelos podem mudar, preservar *prompts* e saídas torna-se ainda mais importante.

Quadro 8.5 - Itens para registros metodológicos

Item de registro	Pergunta que deve responder	Evidência recomendada
Fonte escolhida	De onde vieram os registros e por quê?	Justificativa metodológica e referência à documentação da fonte.
Estratégia de busca	Como os registros foram recuperados?	String, filtros, campos, data, operador ou script.
Exportação	Que dados foram baixados e em que formato?	Arquivos originais, nomes padronizados e manifesto.

Item de registro	Pergunta que deve responder	Evidência recomendada
Processamento	Que transformações foram aplicadas?	Scripts, logs, versões e descrição das regras.
Deduplicação	Como registros repetidos foram tratados?	Critérios, chaves usadas, amostra verificada e registros removidos.
Completude	Quais campos estão ausentes?	Tabela de preenchimento por campo.
Uso de IA	Que *prompts* influenciaram decisões?	*Prompts*, saídas relevantes, decisão humana e validação.
Limitações	O que o *corpus* não representa adequadamente?	Declaração metodológica e impacto interpretativo.

Fonte: Elaborado pelo autor (2026).

A elaboração de diagramas de fluxo pode ajudar. Inspirado em práticas de revisão sistemática e de relato transparente da busca, um diagrama de fluxo cientométrico pode indicar registros identificados, registros exportados, duplicatas removidas, registros excluídos por critérios, registros incluídos na análise textual, registros incluídos em redes de citação e registros incluídos em indicadores de colaboração (PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021). Esse diagrama não deve ser copiado mecanicamente de PRISMA, mas adaptado à lógica bibliométrica. O objetivo é tornar claro como o *corpus* final foi constituído.

A IA pode ajudar a transformar registros técnicos em uma seção metodológica legível. O pesquisador pode fornecer logs, tabelas de contagem e decisões, pedindo que o modelo redija o método em tom científico. O *prompt* deve exigir que números não sejam alterados, que lacunas sejam mantidas, que inferências sejam separadas de fatos e que limitações sejam explicitadas. Após a redação, o texto deve ser conferido contra os documentos originais. A IA é boa em organizar narrativa; não deve ser autorizada a suavizar problemas.

A documentação para reprodutibilidade deve ser proporcional ao risco do uso. Um estudo exploratório para orientação interna pode ter documentação mais simples, mas ainda precisa preservar dados e decisões. Um estudo publicado em periódico deve fornecer detalhes suficientes para revisão. Um relatório usado para avaliação institucional ou distribuição de recursos exige documentação robusta, pois suas consequências são maiores. A responsabilidade metodológica cresce com o potencial de uso dos resultados.

A perda de rastreabilidade é um dos maiores riscos da cientometria com IA. Se o pesquisador usa *prompts* para decidir fontes, limpar metadados, classificar documentos e interpretar resultados, mas não registra essas etapas, o estudo torna-se uma narrativa sem trilha de auditoria. A promessa da IA na cientometria não está em automatizar a responsabilidade, mas em ampliar a capacidade de documentar, revisar e comunicar decisões. O capítulo, portanto, propõe uma ética da coleta: todo dado tem origem, toda transformação tem regra, toda sugestão de IA tem validação e toda limitação deve ser declarada.

8.8 Prompts do capítulo

Os *prompts* a seguir foram concebidos como instrumentos metodológicos. Eles não devem ser usados como comandos isolados nem aceitos sem validação. Cada *prompt* inclui contexto, tarefa, formato de saída e restrições para reduzir ambiguidade. Recomenda-se que o pesquisador registre a versão adaptada do *prompt*, a saída obtida e a decisão humana tomada a partir dela.

Prompt 8.8.1 Planejar a coleta de dados bibliográficos

Atue como especialista em Cientometria, bibliometria e gestão de dados de pesquisa.

Quero planejar a coleta de dados bibliográficos para um estudo sobre [TEMA], com período [ANO INICIAL–ANO FINAL], unidade de análise [ARTIGO/AUTOR/INSTITUIÇÃO/PAÍS/TEMA/OUTRA] e objetivo principal [OBJETIVO].

Elabore um plano de coleta contendo:

1) fontes de dados candidatas;

2) justificativa para cada fonte;

3) campos de metadados indispensáveis;

4) campos desejáveis;

5) filtros que precisam ser definidos;

6) formatos de exportação recomendados;

7) riscos de cobertura e completude;

8) necessidade ou não de combinar fontes;

9) estratégia de coleta piloto;

10) itens que devem ser registrados para reprodutibilidade.

Não invente funcionalidades de bases. Quando uma informação depender da documentação oficial da fonte, marque como "verificar na documentação". Separe recomendações firmes, hipóteses operacionais e pontos que exigem validação humana.

Prompt 8.8.2 Criar checklist de metadados

Atue como curador de metadados bibliográficos para um estudo cientométrico. A pergunta de pesquisa é: [INSERIR PERGUNTA]. As análises pretendidas são: [INDICADORES/REDES/MAPAS/*TEXT MINING*/AVALIAÇÃO INSTITUCIONAL/OUTRAS].

Crie uma checklist de metadados em tabela com as colunas: campo, obrigatoriedade, finalidade analítica, fonte provável, risco se ausente, forma de validação e observações. Considere título, resumo, palavras-chave, autores, afiliações, país, instituição, referências citadas, citações recebidas, DOI, ORCID, ROR, periódico, ISSN, tipo documental, ano, idioma, financiador e acesso aberto.

Ao final, indique quais análises não devem ser realizadas se determinados campos estiverem ausentes ou tiverem baixa completude.

Prompt 8.8.3 Transformar critérios de inclusão e exclusão em protocolo de coleta

Atue como metodologista de estudos cientométricos. A partir dos critérios abaixo, transforme-os em um protocolo operacional de coleta.

Tema: [TEMA]

Fontes previstas: [FONTES]

Período: [PERÍODO]

Critérios de inclusão: [LISTAR]

Critérios de exclusão: [LISTAR]

Tipos documentais: [LISTAR]

Idiomas: [LISTAR]

Campos obrigatórios: [LISTAR]

Produza:

1) objetivo da coleta;

2) escopo do *corpus*;

3) passos de recuperação;

4) filtros a aplicar;

5) campos a exportar;

6) regra de nomeação de arquivos;

7) controle de duplicatas;

8) critérios de exclusão pós-coleta;

9) validação por amostragem;

10) relatório de reprodutibilidade.

Não transforme critérios ambíguos em regras definitivas. Liste ambiguidades e proponha perguntas de esclarecimento, mas também ofereça uma versão provisória do protocolo com as premissas explicitadas.

Prompt 8.8.4 Gerar pseudocódigo de coleta via API

Atue como cientista de dados com experiência em APIs bibliográficas e reprodutibilidade. Preciso coletar registros na API [NOME DA API] para o tema [TEMA], período [PERÍODO], filtros [FILTROS] e campos [CAMPOS].

Gere pseudocódigo, não código final, contendo:

1) definição dos parâmetros;

2) chamada inicial;

3) paginação;

4) controle de taxa;

5) tratamento de erros;

6) salvamento de respostas brutas;

7) transformação em tabela analítica;

8) registro de logs;

9) validação de totais;

10) exportação final.

Inclua comentários sobre quais pontos dependem da documentação oficial da API. Não invente *endpoints*, parâmetros ou limites de taxa. Quando não souber, use marcadores como [VERIFICAR NA DOCUMENTAÇÃO].

Prompt 8.8.5 Auditar uma amostra de registros exportados

Atue como auditor metodológico de dados bibliográficos. Receberei uma amostra de registros exportados de [FONTE]. Analise a amostra e produza um diagnóstico preliminar.

Verifique: campos ausentes, campos truncados, problemas de codificação, variação de nomes de autores, variação de instituições, ausência de DOI, ausência de resumo, ausência de referências, problemas de tipo documental, inconsistências de ano, duplicatas aparentes e riscos para as análises planejadas.

Organize a resposta em:

1) problemas observados;

2) impacto potencial;

3) gravidade;

4) ação recomendada;

5) necessidade de validação humana.

Não extrapole a amostra para todo o *corpus* sem marcar a conclusão como hipótese.

Prompt 8.8.6 Redigir a seção metodológica da coleta

Atue como redator científico. Vou fornecer registros do processo de coleta: fontes, *strings*, filtros, datas, formatos, totais, campos, critérios de exclusão, duplicatas e limitações. Redija uma seção metodológica em tom formal científico.

Regras: não altere números; não invente justificativas; não oculte limitações; diferencie coleta, processamento e análise; informe datas e fontes; indique campos usados; descreva a validação; mencione o uso de IA apenas quando eu informar que ela foi usada; e preserve a distinção entre dado observado, decisão metodológica e limitação.

Formato de saída: texto corrido com subtítulos, seguido de uma tabela resumindo fonte, data, *string*, filtros, total exportado e formato.

8.9 Exemplo aplicado: planejamento de coleta para um estudo sobre IA generativa na educação superior

Considere um pesquisador que deseja mapear a produção científica sobre IA generativa na educação superior entre 2020 e 2026. A pergunta inicial é ampla: 'como evoluiu a pesquisa sobre IA generativa na educação superior?'. Para transformá-la em coleta bibliográfica, é necessário definir se o foco é produção, impacto, colaboração, temas emergentes ou políticas institucionais. Se a análise inclui mapas temáticos, títulos, resumos e palavras-chave são indispensáveis. Se inclui colaboração, autores e afiliações precisam estar completos. Se inclui estrutura intelectual, referências citadas tornam-se necessárias. A escolha da fonte dependerá da cobertura de educação, ciência da computação, tecnologia educacional e anais de conferências.

Uma coleta piloto poderia consultar Scopus, Web of Science, OpenAlex e ERIC, comparando número de resultados, tipos documentais, disponibilidade de resumos e presença de referências. A estratégia deveria incluir termos como *'generative artificial intelligence'*, *'generative AI'*, '*large language models*', 'ChatGPT', *'higher education'*, '*university', 'college'* e termos equivalentes. A IA poderia apoiar a expansão terminológica, mas a *string* final deveria ser revisada por especialistas em educação e informação. A coleta piloto permitiria identificar ruído, como artigos sobre IA generativa em educação médica que não se referem especificamente ao ensino superior, ou documentos sobre modelos generativos sem dimensão educacional.

Após a coleta, o pesquisador deveria gerar relatório de completude: quantos registros têm DOI, resumo, palavras-chave, afiliação, referências e país. Caso OpenAlex ofereça maior cobertura, mas menor completude de resumos em parte do *corpus*, essa diferença deve ser considerada. Caso Scopus e Web of Science tenham melhor estrutura de referências, mas menor cobertura de anais, a limitação deve ser explicitada. Caso ERIC recupere documentos relevantes não presentes em bases multidisciplinares, pode ser usado como fonte complementar, com deduplicação cuidadosa. O resultado metodológico seria um *corpus* cuja construção é transparente, e não apenas um número final de registros.

Esse exemplo mostra que o uso de IA não elimina o trabalho científico de decidir. O modelo pode sugerir fontes, termos, campos e procedimentos; pode construir uma tabela de comparação; pode redigir pseudocódigo; pode auditar amostras. Todavia, a validade do *corpus* depende da verificação na fonte, da coleta piloto, da documentação e da interpretação crítica. Em uma área emergente, como IA generativa, termos mudam rapidamente, documentos podem aparecer como *preprints*, anais podem ter papel relevante e bases podem atualizar registros com atraso. A coleta deve ser entendida como fotografia datada de uma literatura em movimento.

8.10 Considerações finais

A coleta de dados bibliográficos é uma etapa decisiva da cientometria. Ela define o universo visível da análise, condiciona indicadores e delimita interpretações. Em estudos assistidos por inteligência artificial, sua importância aumenta, pois *prompts* podem acelerar decisões, mas também podem mascarar fragilidades. O uso responsável da IA na coleta exige que cada sugestão seja tratada como hipótese operacional, que cada campo seja

verificado nos dados reais, que cada fonte seja escolhida por aderência à pergunta e que cada transformação seja documentada.

O capítulo mostrou que fontes comerciais, abertas e institucionais possuem vantagens e limitações distintas. Apresentou os metadados essenciais para análises de produção, impacto, colaboração, redes e temas. Discutiu formatos de exportação e riscos de perda de informação. Examinou APIs como instrumentos de coleta reprodutível, mas dependentes de documentação e controle técnico. Comparou coleta manual, semiautomática e automatizada. Por fim, propôs que o registro do processo seja tratado como eixo de integridade científica.

A mensagem central é simples: não existe boa cientometria sem boa coleta. Mapas sofisticados, métricas avançadas e narrativas interpretativas dependem da qualidade dos registros que lhes dão origem. A IA pode apoiar a construção desse percurso, mas não deve invisibilizar a origem dos dados. O pesquisador continua responsável por perguntar, escolher, verificar, registrar, interpretar e limitar suas conclusões. A coleta bibliográfica, quando bem documentada, torna-se não apenas uma etapa técnica, mas um argumento metodológico sobre o que se está chamando de ciência analisada.

Governança da coleta e qualidade de dados em projetos cientométricos com IA

A governança da coleta é o conjunto de regras, papéis e controles que assegura que os dados bibliográficos sejam obtidos, armazenados, transformados e utilizados de forma consistente. Em estudos pequenos, ela pode assumir a forma de uma checklist. Em estudos institucionais ou multicêntricos, deve assumir a forma de um plano de gestão de dados, com responsabilidades explícitas. A presença de IA no processo torna a governança ainda mais necessária, porque parte das decisões pode ocorrer em interações conversacionais não registradas. Sem governança, a equipe perde a capacidade de explicar por que determinada fonte foi escolhida, por que certo campo foi priorizado, ou por que um conjunto de registros foi excluído.

Uma política mínima de governança deve definir quem tem autorização para executar coletas, alterar *strings*, consultar APIs, editar scripts, normalizar campos e aprovar o *corpus* final. Essa definição não é burocrática; ela previne variações de procedimento. Em projetos nos quais diferentes membros exportam registros de bases distintas, pequenas diferenças de configuração podem produzir arquivos incompatíveis. Um pesquisador pode exportar apenas registros e citações; outro pode incluir referências completas; outro pode aplicar filtros de idioma sem perceber. A governança reduz essa variabilidade e transforma a coleta em prática coletiva controlada.

A qualidade de dados pode ser compreendida em dimensões complementares: completude, precisão, consistência, atualidade, unicidade, validade e proveniência. Completude diz respeito à presença de campos necessários. Precisão refere-se à correção do valor registrado. Consistência avalia se o mesmo campo segue padrões uniformes. Atualidade indica se o dado está atualizado em relação à fonte. Unicidade trata de duplicatas. Validade considera se valores pertencem ao domínio esperado. Proveniência identifica de onde cada dado veio. Um estudo cientométrico robusto deve tratar essas dimensões antes de produzir indicadores.

A IA pode apoiar a criação de métricas de qualidade, mas a medição deve ser feita nos dados. Por exemplo, o modelo pode sugerir que o pesquisador calcule a proporção de registros com DOI, a proporção com resumo, a proporção com afiliação vinculada a autor, a distribuição de tipos documentais e a frequência de registros sem ano. No entanto, o cálculo deve ser executado em planilhas ou scripts. A IA pode interpretar o resultado, sugerir riscos

e ajudar a redigir a seção metodológica. A distinção entre cálculo, interpretação e redação deve permanecer explícita.

Uma regra útil é produzir um relatório de qualidade antes da análise principal. Esse relatório não precisa ser longo, mas deve responder a perguntas básicas: quantos registros há no *corpus* bruto? Quantos permanecem após deduplicação? Quais campos têm maior ausência? Quais fontes contribuem com mais registros? Há discrepâncias temporais? Há áreas ou idiomas sub-representados? Há valores extremos em citações? Há instituições com múltiplas grafias? Esse relatório serve como autorização metodológica para prosseguir. Se problemas graves aparecem, a coleta deve ser revisada antes de gerar mapas ou métricas.

A gestão de riscos da coleta também deve ser planejada. Riscos frequentes incluem: escolha de fonte inadequada, busca subespecificada, limites de exportação não percebidos, perda de referências, afiliações incompletas, duplicatas persistentes, campos truncados, uso indevido de API, violação de termos de uso, dependência de uma versão dinâmica da base, documentação insuficiente e confiança excessiva em sugestões de IA. Cada risco pode receber uma ação preventiva. Por exemplo, limites de exportação podem ser mitigados por segmentação planejada; afiliações incompletas por teste de completude; *prompts* arriscados por exigência de marcação de incerteza e verificação documental.

A governança também exige decisões sobre armazenamento. Dados brutos devem ser preservados em local seguro, com controle de acesso quando necessário. Dados derivados devem ser organizados por versão. Scripts devem estar próximos aos dados que processam, mas sem credenciais embutidas. Arquivos de log devem ser mantidos. Relatórios finais devem indicar qual versão do *corpus* foi usada. Quando o estudo envolve dados comerciais, a equipe deve definir o que pode ou não ser compartilhado. Quando envolve bases abertas, deve-se preferir compartilhamento de código, identificadores e dados processados sempre que permitido.

A interoperabilidade entre fontes é um desafio de governança. Integrar dados de Web of Science, Scopus, Crossref, OpenAlex e sistemas institucionais exige decisões sobre quais campos serão considerados canônicos. O título da fonte comercial será priorizado em caso de divergência? O DOI de Crossref substituirá valores ausentes? A afiliação será normalizada por ROR? O país será inferido pela afiliação ou usado apenas quando fornecido pela base? Essas decisões devem ser pré-definidas. Se forem tomadas caso a caso sem registro, a harmonização perde transparência.

Em projetos que usam IA para normalização, a governança deve ser ainda mais cautelosa. Modelos de linguagem podem agrupar variantes institucionais, traduzir termos e sugerir categorias temáticas, mas podem também unir entidades diferentes ou impor classificações inadequadas. Recomenda-se usar IA para gerar candidatos, não para substituir a validação. Por exemplo, o modelo pode propor que 'Univ. São Paulo', 'University of Sao Paulo' e 'USP' correspondem à mesma instituição, mas a equipe deve confirmar e registrar a regra. Em instituições com nomes semelhantes, hospitais universitários e campi regionais, a decisão pode afetar indicadores.

A documentação da participação da IA pode ser feita em tabela. Cada linha registra etapa, objetivo do *prompt*, dados fornecidos, saída utilizada, decisão humana, validação e impacto. Essa tabela evita duas distorções opostas: esconder o uso de IA ou atribuir à IA uma responsabilidade que é humana. A transparência permite que revisores avaliem se o uso foi aceitável. Em muitos casos, a IA terá apoiado redação, organização e triagem, mas não terá executado a coleta. Em outros, terá gerado pseudocódigo ou sugerido regras de limpeza. Essas diferenças precisam aparecer.

A governança deve prever revisão por pares internos. Antes de fechar o *corpus*, um membro da equipe que não executou a coleta deve revisar o protocolo, uma amostra de arquivos, os logs, a tabela de completude e as regras de deduplicação. Essa revisão é semelhante ao controle de qualidade de uma busca sistemática, embora adaptada à cientometria. A revisão interna pode identificar termos ausentes, filtros indevidos ou campos não exportados. A IA pode apoiar a criação do roteiro de revisão, mas a avaliação deve ser humana e fundamentada nos dados.

A coleta deve ser documentada em linguagem compreensível para diferentes públicos. Um cientista de dados pode entender logs e JSON; um gestor de pesquisa pode precisar de tabela executiva; um revisor de periódico pode exigir detalhes metodológicos; um especialista temático pode avaliar pertinência da busca. Produzir camadas de documentação é uma boa prática. A documentação técnica preserva reprodutibilidade; a documentação metodológica explica escolhas; a documentação executiva comunica limites e usos. A IA pode adaptar a linguagem sem alterar o conteúdo factual.

O controle de qualidade deve incluir testes de plausibilidade. Uma série temporal com queda abrupta pode indicar mudança real, mas também erro de filtro. Uma instituição com produção zero em determinado ano pode refletir ausência de afiliação ou alteração no nome. Um autor com número extraordinário de documentos pode representar homônimo não desambiguado. Uma explosão de palavras-chave pode decorrer de mudança terminológica ou de introdução de termos automáticos. Antes de interpretar padrões como fenômenos científicos, deve-se descartar artefatos de coleta.

A governança de coleta também dialoga com princípios de ciência aberta. Quando possível, o pesquisador deve tornar disponíveis estratégias de busca, listas de identificadores, scripts de coleta, dicionários de dados e relatórios de qualidade. A abertura fortalece a crítica e o reuso. No entanto, ciência aberta não significa violar licenças de bases comerciais. A solução é compartilhar o que pode ser compartilhado: metodologia, parâmetros, código, agregados, identificadores ou dados derivados. A documentação deve explicar quaisquer restrições de compartilhamento.

Uma coleta governada deve ser preparada para auditoria futura. Meses depois da publicação, outro pesquisador pode perguntar por que determinado documento foi excluído, por que a contagem de citações difere de consulta atual, ou por que uma instituição aparece com certo nome. Sem registros, a resposta dependerá de memória. Com registros, a equipe pode explicar que a coleta ocorreu em determinada data, com determinada fonte, e que os dados mudaram posteriormente. Essa capacidade de resposta é sinal de maturidade metodológica.

Em síntese, governança e qualidade de dados são componentes estruturais da cientometria com IA. Elas conectam decisão humana, infraestrutura bibliográfica, automação e interpretação. A coleta governada não impede erros, mas permite identificá-los, corrigi-los e descrevê-los. A ausência de governança transforma erros em ruído invisível. Para um livro dedicado à análise da ciência com *prompts*, essa lição é central: a IA pode acelerar a coleta, mas somente a governança assegura que aceleração não se converta em opacidade.

Erros comuns na coleta de dados bibliográficos e como preveni-los

O primeiro erro comum é confundir resultado de busca com *corpus* validado. Ao executar uma busca em uma base, o pesquisador recebe uma lista inicial de registros, não necessariamente um conjunto pronto para análise. Essa lista pode conter ruído, duplicatas, tipos documentais indesejados, registros sem campos necessários e documentos fora do escopo. O *corpus* só se constitui após validação, aplicação de critérios, deduplicação e

controle de qualidade. A IA pode ajudar a resumir resultados, mas não deve antecipar a validade do *corpus* sem inspeção.

O segundo erro é escolher a fonte por prestígio, e não por adequação. Bases reconhecidas são valiosas, mas não cobrem igualmente todos os campos, idiomas e tipos documentais. Um estudo sobre produção latino-americana, por exemplo, pode exigir bases regionais. Um estudo sobre ciência da computação pode exigir anais. Um estudo sobre políticas públicas pode exigir documentos governamentais. A decisão metodológica correta é aquela que alinha pergunta, fonte e unidade de análise. Prestígio da base não compensa desajuste ao objeto.

O terceiro erro é não exportar campos essenciais. Muitos pesquisadores descobrem tarde que não exportaram referências citadas, resumos ou afiliações. Quando isso ocorre, a análise planejada precisa ser reduzida ou a coleta refeita. A prevenção é simples: criar checklist de metadados antes da coleta final e executar piloto. O checklist deve derivar das análises pretendidas. Se haverá rede de coautoria, autores precisam estar completos; se haverá mapa conceitual, termos textuais precisam estar presentes; se haverá análise institucional, afiliações precisam ser utilizáveis.

O quarto erro é alterar dados brutos diretamente. Abrir uma planilha exportada, corrigir nomes manualmente e salvar sobre o arquivo original elimina a trilha de auditoria. A regra é preservar o bruto e trabalhar em cópia. Qualquer limpeza deve ser rastreável. Se correções manuais forem necessárias, devem ser registradas em tabela auxiliar: valor original, valor corrigido, justificativa e responsável. A IA pode sugerir correções, mas os valores originais devem permanecer acessíveis.

O quinto erro é tratar identificadores como infalíveis. DOI, ORCID e ROR são extremamente úteis, mas podem estar ausentes, incorretos ou duplicados. Um DOI pode apontar para versão diferente; um ORCID pode não estar associado a todos os trabalhos do autor; uma instituição pode não ter identificador nos registros antigos. Identificadores devem ser usados como evidências fortes, não como substitutos absolutos de verificação. Em caso de conflito, o protocolo deve indicar como decidir.

O sexto erro é ignorar a data da coleta. Contagens de citação, status de acesso aberto, metadados e cobertura de bases mudam. Sem data, não é possível interpretar diferenças entre resultados. A data deve aparecer no arquivo, no log e na seção metodológica. Em coletas via API, data e hora são recomendáveis. Em análises de impacto, a data da contagem de citações é fundamental, pois citações acumulam continuamente.

O sétimo erro é misturar fontes sem proveniência. Ao combinar bases, o pesquisador pode criar um *dataset* maior, mas perde interpretabilidade se não sabe de onde veio cada campo. A solução é manter colunas de origem e preservar identificadores internos. Se o registro aparece em três bases, essa informação pode ser útil para avaliar cobertura. Se um campo foi enriquecido por API, isso deve ser indicado. Proveniência é a memória do dado.

O oitavo erro é usar IA para preencher lacunas sem marcar inferência. Um modelo pode sugerir país a partir de instituição, área a partir de título, ou categoria temática a partir de resumo. Essas operações podem ser úteis, mas são inferências, não metadados originais. Devem ser armazenadas em campos separados e avaliadas por amostragem. Misturar metadado coletado com metadado inferido cria falsa objetividade. A prevenção é usar nomenclatura explícita, como 'categoria_ia_validada' ou 'pais_inferido'.

O nono erro é produzir seção metodológica insuficiente. Frases como 'os dados foram coletados na Scopus não bastam. É preciso informar data, *string*, filtros, tipos documentais, campos exportados, formato, total recuperado, total incluído e tratamentos aplicados. A falta

desses elementos impede reprodutibilidade e reduz confiança. A IA pode redigir uma seção mais clara, mas somente se receber informações completas. Um *prompt* não substitui registro.

O décimo erro é superinterpretar o *corpus*. Um conjunto de registros coletado em uma base específica, em uma data específica, com uma busca específica, representa apenas o recorte definido. Se o estudo não inclui livros, não deve concluir sobre toda a produção de uma área baseada em livros. Se não inclui literatura não inglesa, não deve generalizar para comunidades locais. A coleta delimita o alcance das conclusões. A honestidade interpretativa começa no reconhecimento do recorte.

A prevenção desses erros pode ser organizada em um ciclo: planejar, pilotar, coletar, validar, documentar, analisar e revisar. Planejar define fontes e campos. Pilotar testa a estratégia. Coletar executa com registro. Validar mede completude e consistência. Documentar preserva rastreabilidade. Analisar aplica indicadores e mapas. Revisar verifica coerência entre dados e conclusões. A IA pode apoiar todas as etapas, mas o ciclo deve continuar sob controle humano.

Em ambientes de ensino, esses erros são oportunidades pedagógicas. Pedir a estudantes que comparem duas exportações, uma completa e outra incompleta, revela rapidamente o impacto da coleta. Pedir que auditem uma amostra com registros sem DOI ou afiliação mostra limitações de metadados. Pedir que uma IA gere uma metodologia sem dados suficientes e depois comparar com uma metodologia documentada ajuda a compreender os riscos da falsa autoridade textual. A coleta é um bom campo para ensinar rigor científico aplicado.

A prevenção de erros deve ser incorporada aos *prompts*. Em vez de pedir 'faça a coleta', o pesquisador deve pedir 'liste riscos da coleta', 'verifique quais campos são indispensáveis', 'diferencie metadado original de inferência', 'não invente parâmetros', 'marque pontos que exigem documentação oficial'. *Prompts* orientados ao risco produzem respostas mais úteis que *prompts* orientados apenas à produtividade. Na cientometria com IA, um bom *prompt* não é o que promete resposta final rápida, mas o que melhora a qualidade da decisão.

O tratamento dos erros também deve ser transparente. Se um erro foi descoberto depois de análises preliminares, o pesquisador deve registrar a correção e, quando necessário, recalcular resultados. Não há vergonha metodológica em corrigir; há problema em ocultar. Estudos cientométricos frequentemente lidam com dados imperfeitos. A integridade reside em reconhecer imperfeições, mitigá-las e comunicar seus efeitos. A IA pode auxiliar a revisar o relatório de limitações, mas não deve ser usada para obscurecê-las.

O objetivo final não é eliminar toda incerteza, mas produzir uma coleta suficientemente confiável para responder à pergunta proposta. Essa suficiência depende da pergunta. Um estudo exploratório tolera mais incompletude que uma avaliação institucional. Uma análise temática pode aceitar ausência de citações, mas não de títulos ou resumos. Uma rede de citação exige referências. Um mapa de colaboração exige afiliações. A qualidade da coleta deve ser julgada em relação ao uso pretendido, e não por critério único universal.

Apêndice A.8 - Perguntas de auditoria da coleta

Uma auditoria de coleta deve iniciar pela relação entre pergunta e fonte. A equipe deve perguntar se a base escolhida representa adequadamente o tipo de produção que se pretende analisar, se o período coberto pela base coincide com o período do estudo, se a área disciplinar possui tradição de publicação nos canais indexados e se há documentos relevantes fora do universo capturado. Essa reflexão impede que a coleta seja confundida com realidade empírica total. A fonte não contém a ciência; contém uma representação da ciência sob critérios específicos de indexação e disponibilidade.

A segunda pergunta de auditoria incide sobre a estratégia de recuperação. Deve-se verificar se os termos usados cobrem sinônimos, variantes, siglas, traduções, termos antigos e termos emergentes. Deve-se também revisar se a *string* é excessivamente ampla ou excessivamente estreita. Uma busca muito sensível pode trazer grande ruído, enquanto uma busca muito específica pode excluir subáreas relevantes. A IA pode ajudar a sugerir vocabulário, mas a equipe deve testar a recuperação contra um conjunto de artigos conhecidos. Se artigos centrais não aparecem, a estratégia precisa ser revista.

A terceira pergunta trata dos filtros. O que foi excluído por ano, idioma, tipo documental, área, índice, país ou fonte? Esses filtros são coerentes com o objetivo? Foram aplicados antes ou depois da exportação? Foram registrados? Muitas vezes, a *string* é relatada, mas os filtros permanecem invisíveis. Em bases multidisciplinares, filtros de área podem mudar substancialmente o *corpus*. Em estudos sobre áreas interdisciplinares, filtros disciplinares podem excluir justamente os documentos mais relevantes. A auditoria deve identificar esse risco.

A quarta pergunta diz respeito aos campos. O conjunto exportado contém tudo que a análise promete? Se o relatório apresenta colaboração internacional, os dados têm afiliação e país? Se apresenta cocitação, há referências citadas? Se apresenta temas por resumo, os resumos foram exportados para a maioria dos registros? Se apresenta impacto, a fonte das citações e a data da contagem estão registradas? A coerência entre promessa analítica e metadados disponíveis é um dos critérios mais importantes de qualidade cientométrica.

A quinta pergunta refere-se à integridade dos arquivos. O total na interface corresponde ao total exportado? Os lotes se sobrepõem? Há registros ausentes? A codificação preserva caracteres? Os nomes dos arquivos indicam fonte, data e lote? Os arquivos brutos estão separados dos processados? Essas questões parecem administrativas, mas erros nesse nível podem alterar estatísticas básicas. A auditoria deve conferir contagens antes e depois da importação, porque perdas podem ocorrer na conversão.

A sexta pergunta avalia a deduplicação. Que chaves foram usadas? DOI, título, ano, autores, identificadores internos? Como foram tratados documentos sem DOI? A fusão preservou a proveniência dos campos? Uma regra agressiva pode fundir documentos distintos; uma regra permissiva pode manter duplicatas. Em estudos de produção, duplicatas inflacionam contagens; em redes, podem criar nós e arestas artificiais. A deduplicação deve ser testada em amostras e descrita de modo transparente.

A sétima pergunta aborda o uso de IA. Que *prompts* foram utilizados? Que decisões eles influenciaram? A saída da IA foi verificada? O modelo sugeriu campos, fontes, filtros ou pseudocódigo? Houve risco de invenção de funcionalidades? A auditoria não deve proibir o uso de IA, mas torná-lo visível. Quando a IA participa de uma etapa metodológica, sua participação precisa ser documentada como parte do protocolo, assim como se documenta um software estatístico ou uma função de importação.

A oitava pergunta é sobre limitações. O estudo declara o que não foi capturado? Reconhece lacunas de cobertura, idioma, período, tipo documental e metadados? Explica o impacto dessas lacunas sobre os resultados? Uma limitação bem descrita orienta a interpretação; uma limitação oculta produz falsa precisão. Em cientometria, a honestidade metodológica é particularmente importante porque resultados podem ser usados para comparar pessoas, grupos e instituições.

A nona pergunta relaciona coleta e comunicação. O relatório final permite que o leitor compreenda como o *corpus* foi constituído? Apresenta tabela de fontes, datas, *strings*, filtros e totais? Inclui apêndice técnico quando necessário? Disponibiliza dados ou identificadores

quando permitido? Um bom relatório não exige que o leitor confie cegamente no pesquisador; oferece elementos para julgar a robustez da coleta. Essa é a função científica da documentação.

A décima pergunta trata da proporcionalidade. O nível de documentação é adequado ao uso dos resultados? Estudos exploratórios podem ter registros mais simples, mas avaliações institucionais exigem documentação robusta. Quanto maior o impacto potencial da análise sobre decisões de financiamento, reputação ou política científica, maior deve ser o padrão de rastreabilidade. A coleta de dados bibliográficos é, portanto, uma prática técnica e ética.

Capítulo 9

Pré-processamento e limpeza de dados com prompts

9.1 Introdução

O pré-processamento e a limpeza de dados constituem uma das etapas menos visíveis e, ao mesmo tempo, mais determinantes da cientometria. Em muitos relatórios, artigos e painéis institucionais, os resultados aparecem sob a forma de indicadores, rankings, mapas de coautoria, redes de coocorrência, mapas temáticos ou séries temporais. Entretanto, antes que qualquer visualização seja produzida, existe um conjunto de decisões sobre grafias de nomes, fusão de registros, tratamento de duplicatas, padronização de palavras-chave, agrupamento terminológico, resolução de afiliações e preservação de identificadores persistentes. Essas decisões definem o universo empírico do estudo e, por isso, precisam ser descritas como parte do método, e não como operação de bastidor.

A literatura metodológica recente reforça essa preocupação. Nowakowska (2025) demonstra que o pré-processamento bibliométrico influencia diretamente a confiabilidade de análises baseadas em dados bibliográficos. Lim, Kumar e Donthu (2024) mostram que a combinação de bases e ferramentas exige limpeza cuidadosa para evitar resultados incompatíveis. Donthu *et al.* (2021) e Aria e Cuccurullo (2017) também destacam que o valor de indicadores, mapas e redes depende de uma preparação de dados coerente com a pergunta de pesquisa. Desse modo, a qualidade da cientometria não reside apenas no algoritmo de visualização, mas na inte*grid*ade dos metadados que alimentam a análise.

A inserção de modelos de linguagem e *prompts* nessa etapa desloca a limpeza de dados de uma atividade puramente operacional para uma prática metodológica documentável. Quando se pede a uma IA que agrupe palavras-chave, compare nomes de autores, sugira equivalências institucionais ou identifique duplicatas prováveis, não se está apenas acelerando uma tarefa mecânica. Está-se produzindo uma mediação interpretativa que pode alterar a estrutura de uma rede científica, modificar a centralidade de uma instituição, aumentar ou reduzir a produtividade atribuída a um pesquisador e redefinir a composição temática de uma área.

Este capítulo adota a premissa de que *prompts* podem apoiar o pré-processamento, mas não devem substituir procedimentos formais de controle de qualidade. A IA pode sugerir agrupamentos, detectar padrões de variação, explicar possíveis ambiguidades, gerar listas de verificação, propor regras de normalização, escrever pseudocódigo e produzir relatórios de auditoria. Entretanto, a mesma IA pode criar falsas equivalências, apagar diferenças disciplinares relevantes, atribuir erroneamente publicações a autores homônimos, converter instituições distintas em uma mesma entidade ou inventar justificativas plausíveis para decisões equivocadas. O ganho de produtividade precisa ser acompanhado por mecanismos de contenção metodológica.

A noção de *prompt*, neste capítulo, é compreendida como instrução metodológica formalizada. Um bom *prompt* de limpeza bibliográfica explicita a tarefa, o campo analisado, os critérios de equivalência, os limites da decisão automatizada, o formato de saída e o modo de sinalizar incerteza. Em vez de solicitar apenas "limpe esta lista", o pesquisador deve informar o tipo de dado, o objetivo da análise, os campos disponíveis, os riscos conhecidos e as regras de não decisão. A IA deve separar casos resolvidos de casos duvidosos, declarar justificativas, preservar dados originais e produzir trilhas de decisão que possam ser revisadas por especialistas.

A etapa de pré-processamento também é um espaço privilegiado para aplicar a diferença entre automação e curadoria. Automação significa executar operações repetitivas, como remover espaços extras, unificar maiúsculas e minúsculas, normalizar acentos, padronizar separadores e identificar padrões óbvios de duplicidade. Curadoria significa decidir se "Univ Sao Paulo", "Universidade de São Paulo", "USP" e "University of São Paulo" devem ser tratados como a mesma instituição; se "*machine learning*" e "aprendizado de máquina" são equivalentes para fins de coocorrência; ou se dois registros com títulos semelhantes, mas anos diferentes, representam o mesmo documento ou versões distintas.

Em estudos cientométricos assistidos por IA, a limpeza deve ser registrada como parte do método. Isso implica documentar a versão da base utilizada, a data de coleta, os filtros aplicados, os campos exportados, as regras de deduplicação, os vocabulários normalizados, os *prompts* utilizados, as respostas da IA, as alterações feitas manualmente e a amostragem de qualidade. A reprodutibilidade não exige que todas as decisões sejam automatizadas; exige que as decisões sejam compreensíveis, justificáveis e, quando possível, replicáveis. Assim, a IA deixa de ser uma caixa-preta informal e passa a compor um protocolo auditável de preparação de dados.

Quadro 9.1 - Problemas típicos de pré-processamento em bases bibliográficas

Campo	Problema frequente	Risco cientométrico	Apoio possível da IA
Autores	Variações de grafia, abreviações, homônimos e identificadores incompletos.	Fragmentação ou fusão indevida de produção, coautoria e trajetória científica.	Sugerir agrupamentos candidatos com evidência, grau de confiança e marcação de ambiguidade.
Instituições	Siglas, traduções, campi, unidades subordinadas e nomes históricos.	Erro em redes institucionais, mapas geográficos e rankings.	Extrair entidades, separar níveis institucionais e sinalizar casos que exigem validação externa.
Documentos	Duplicatas exatas e aproximadas entre bases, versões e registros com metadados divergentes.	Contagem inflada de produção, citações e vínculos de rede.	Identificar pares candidatos e justificar a decisão sem exclusão automática.
Palavras-chave	Sinônimos, plural, siglas, traduções e vocabulários de origens distintas.	*Clusters* temáticos artificiais ou excessivamente fragmentados.	Agrupar termos equivalentes, preservar casos duvidosos e registrar regra de normalização.
Referências	Citações incompletas, abreviações e grafias diversas de obras citadas.	Distorção em cocitação, acoplamento bibliográfico e historiografia científica.	Sugerir chaves de unificação com base em campos fornecidos, sem inventar dados.

Fonte: Elaborado pelo autor (2026).

9.2 Padronização de nomes de autores

A padronização de nomes de autores é uma das tarefas mais sensíveis da preparação de dados bibliográficos. Nomes abreviados, iniciais omitidas, acentos, partículas familiares, transliterações e variações de sobrenome podem fragmentar a produção de um mesmo

pesquisador. Por outro lado, autores distintos podem compartilhar a mesma forma abreviada, especialmente em bases que registram sobrenome e iniciais. A limpeza inadequada pode produzir dois erros simétricos: separar indevidamente o que deveria ser unido ou unir indevidamente autores homônimos.

A literatura de desambiguação autoral mostra que esse problema é estrutural em análises bibliométricas. Müller, Reitz e Roy (2017) discutem a importância de conjuntos de dados confiáveis para avaliar procedimentos de desambiguação, enquanto Tekles e Bornmann (2020) comparam abordagens não supervisionadas e demonstram que a desambiguação é condição para análises confiáveis em nível de autor. Em estudos avaliativos, o problema deixa de ser apenas técnico: uma fusão indevida pode atribuir produção a quem não a realizou; uma separação indevida pode invisibilizar parte da trajetória de um pesquisador.

Em uma abordagem orientada por *prompts*, a primeira regra é definir a finalidade da padronização. Em um estudo de produtividade individual, a atribuição de publicações exige rigor maior do que em uma análise agregada de temas. Em uma rede de coautoria, a distinção entre autores homônimos altera nós e arestas. Em uma análise institucional, nomes de autores podem ser usados apenas como campos auxiliares para validação. O *prompt* deve solicitar que a IA apresente candidatos a agrupamento, e não decisões finais, indicando evidências usadas, grau de confiança e casos que exigem revisão humana.

Identificadores persistentes podem reduzir incertezas, mas não eliminam a necessidade de curadoria. ORCID, ResearcherID, Scopus Author ID e identificadores internos de bases podem estar ausentes, incompletos ou associados a apenas parte da produção. A regra editorial recomendada é preservar sempre o nome original, criar campo de nome normalizado, registrar a regra aplicada e manter uma coluna de status. Assim, a base permite retornar ao dado bruto quando uma decisão for contestada.

A IA pode apoiar a triagem ao comparar variantes nominais com coautores, afiliações, áreas temáticas, títulos e períodos de publicação. Contudo, deve ser instruída a não inferir identidade quando os campos são insuficientes. Comandos como “não una autores apenas por sobrenome e iniciais”, “marque casos de baixa confiança” e “explique a evidência disponível” reduzem o risco de limpeza agressiva. A melhor saída nem sempre é a fusão; muitas vezes, a decisão metodologicamente correta é manter registros separados até validação externa.

9.3 Desambiguação de instituições

A desambiguação institucional envolve a normalização de afiliações, siglas, nomes traduzidos, unidades subordinadas, hospitais, laboratórios, campi e departamentos. Em estudos de colaboração, internacionalização e avaliação institucional, esse campo é decisivo. Uma mesma organização pode aparecer com formas diferentes; instituições distintas podem compartilhar siglas; unidades internas podem ser relevantes em uma pergunta e irrelevantes em outra. A limpeza deve, portanto, explicitar o nível de análise: universidade, sistema, faculdade, departamento, laboratório, hospital ou campus.

Trabalhos recentes sobre desambiguação institucional, como o AffilGood de Duran-Silva *et al.* (2024), mostram que afiliações em literatura científica são complexas, multilíngues e frequentemente ruidosas. O uso de identificadores como ROR pode apoiar a normalização, mas nem todas as bases exportam esses identificadores e nem todas as organizações estão representadas de modo inequívoco. Em registros antigos ou regionais, os metadados podem ser incompletos, o que exige decisão humana fundamentada.

O *prompt* para desambiguação institucional deve solicitar uma tabela com afiliação original, instituição normalizada, unidade subordinada, país, identificador quando fornecido, justificativa, grau de confiança e alerta de ambiguidade. A IA não deve inventar identificadores nem resolver siglas sem evidência. Quando "USP" aparece em um conjunto brasileiro, pode indicar Universidade de São Paulo; em outro contexto, pode significar outra organização. A evidência contextual - país, cidade, coautores, periódico área - é indispensável.

A decisão entre agrupar no nível de universidade ou preservar unidades subordinadas depende da pergunta. Para análise macro de colaboração internacional, pode ser suficiente normalizar no nível da instituição principal. Para diagnóstico interno de uma universidade, departamentos, programas, institutos e laboratórios podem ser unidades centrais. Essa decisão deve ser registrada porque altera mapas, rankings e indicadores. O mesmo conjunto de dados pode produzir resultados distintos conforme a granularidade adotada.

A governança da limpeza institucional recomenda manter colunas separadas para afiliação original, instituição normalizada, unidade subordinada, país, identificador, regra aplicada, responsável e status. Essa estrutura preserva a reversibilidade e permite auditoria. A IA deve ser usada para sugerir candidatos, detectar padrões e organizar inconsistências, não para substituir a validação por especialistas institucionais quando o resultado tiver consequências avaliativas.

9.4 Tratamento de duplicatas

Duplicatas são registros que representam o mesmo documento ou versões muito próximas de um documento. Podem surgir dentro de uma mesma base, mas são mais frequentes quando o pesquisador combina Scopus, Web of Science, Dimensions, OpenAlex, Lens, PubMed, Crossref e repositórios institucionais. A presença de duplicatas inflaciona contagens de produção, altera indicadores de citação, distorce redes de coautoria e pode duplicar relações em mapas de acoplamento ou cocitação.

A detecção de duplicatas é problema clássico de limpeza de dados. Elmagarmid, Ipeirotis e Verykios (2007) mostram que registros duplicados podem apresentar erros, ausência de chaves comuns e variações de formato, exigindo medidas de similaridade e critérios de decisão. Em cientometria, a situação é agravada por versões documentais: *preprint*, artigo aceito, versão publicada, capítulo, registro de conferência e repositório podem parecer duplicatas, embora representem objetos distintos conforme a pergunta de pesquisa.

A primeira etapa deve separar duplicata exata, duplicata aproximada e versão relacionada. Registros com mesmo DOI, título, ano e fonte podem ser fundidos com baixa incerteza. Registros com títulos semelhantes, autores iguais e anos próximos exigem revisão. Um *preprint* e um artigo publicado podem ser tratados como versões do mesmo trabalho ou como objetos distintos, dependendo do objetivo. Se o estudo analisa circulação do conhecimento, o *preprint* pode ser relevante; se analisa artigos revisados por pares, deve ser tratado separadamente.

O *prompt* de identificação de duplicatas deve usar DOI, título normalizado, primeiro autor, ano, periódico ou evento, páginas, volume, resumo e identificadores de base. A saída recomendada deve classificar cada par como duplicata provável, possível duplicata ou não duplicata, indicando evidência e ação recomendada: fundir, manter separado ou revisar manualmente. O modelo não deve excluir registros automaticamente. A decisão conservadora é preferível quando o impacto sobre indicadores é alto.

O relatório de deduplicação deve informar número inicial de registros, número de duplicatas removidas, critérios usados, casos ambíguos e impacto esperado sobre a análise.

Quando múltiplas bases são combinadas, a proveniência deve ser preservada. Um registro final pode indicar que o documento foi recuperado em mais de uma fonte. Essa informação não é ruído, pode ser útil para compreender cobertura e sobreposição entre bases.

9.5 Normalização de palavras-chave

A normalização de palavras-chave é decisiva para análises de coocorrência, mapas temáticos, identificação de temas emergentes e interpretação de *clusters*. Palavras-chave autorais, palavras-chave indexadas, descritores controlados, tópicos gerados por bases e termos extraídos de resumos têm naturezas diferentes. Combinar esses campos sem distinção pode produzir redes artificiais ou superestimar termos que derivam de vocabulários controlados.

Variações morfológicas simples, como singular e plural, maiúsculas, hifens, acentos e abreviações, podem fragmentar termos semanticamente equivalentes. Já a polissemia exige cautela: termos como "rede", "modelo", "impacto", "plataforma" e "avaliação" podem ter sentidos diferentes em áreas distintas. A normalização deve distinguir ajustes formais de agrupamentos conceituais. A primeira operação pode ser mais automatizada, a segunda depende de curadoria disciplinar.

Ferramentas como VOSviewer e bibliometrix oferecem recursos para mapear termos, mas a qualidade do mapa depende da preparação do vocabulário (VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017). A IA pode apoiar a criação de tabelas de equivalência, sugerindo termo preferencial, variantes, justificativa e grau de confiança. Contudo, deve preservar termos ambíguos para revisão. Agrupar "responsible AI", "ethical AI", "AI governance" e "algorithmic accountability" pode simplificar a visualização, mas empobrecer a interpretação se a pergunta busca diferenciar agendas.

A normalização deve ser documentada por meio de uma tabela com valor original, termo normalizado, tipo de alteração, regra aplicada, justificativa, status e responsável. Essa tabela pode ser disponibilizada como material suplementar. Em estudos longitudinais, é importante registrar se termos antigos foram agrupados com termos recentes, pois essa decisão pode alterar narrativas de emergência temática. Um mapa limpo, porém excessivamente harmonizado, pode ocultar transformações históricas relevantes.

O *prompt* ideal para palavras-chave deve proibir a eliminação de termos sem registro, pedir separação entre variação formal e equivalência semântica, solicitar marcação de polissemia e exigir exemplos de evidência textual. A saída deve ser mais verificável do que elegante. Tabelas com colunas de confiança e ação recomendada são preferíveis a listas finais supostamente prontas.

9.6 Agrupamento de sinônimos e variantes terminológicas

O agrupamento de sinônimos e variantes terminológicas é uma extensão da normalização de palavras-chave, mas envolve decisões conceituais mais profundas. Termos como *"bibliometrics"* e "*bibliometric* analysis" podem ser agrupados quando representam o mesmo conceito operacional. Já "cientometria", "bibliometria", "avaliação da pesquisa" e "mapeamento científico" podem ser próximos, mas não necessariamente equivalentes. A decisão de agrupá-los depende da pergunta, da escala da análise e do nível de granularidade desejado.

Em *corpus* multilíngues, a tradução terminológica exige atenção especial. Português, espanhol, inglês e francês podem coexistir em títulos, resumos e palavras-chave. A IA pode sugerir equivalências, mas deve distinguir tradução literal, termo técnico usual e aproximação contextual. Termos equivalentes em dicionário nem sempre são equivalentes em comunidade

científica. Essa precaução é particularmente importante em estudos latino-americanos, nos quais expressões locais podem carregar tradições teóricas próprias.

Acronímia é outro ponto crítico. Siglas como AI, IA, ML, LLM, NLP e RAG podem ter usos diferentes conforme área, idioma e período. O *prompt* deve solicitar que siglas sejam expandidas apenas quando houver evidência textual. Em caso de dúvida, a recomendação é preservar o termo separado ou marcar como pendente. Essa prática evita que a IA produza normalização excessiva com aparência técnica convincente.

O nível de granularidade deve ser definido antes da limpeza. Em uma análise macro, agrupar termos amplos pode aumentar legibilidade. Em uma análise de frentes emergentes, preservar subtermos específicos pode revelar diferenciações importantes. Por isso, *prompts* de agrupamento devem pedir ao modelo que apresente alternativas de granularidade e explique perdas informacionais associadas a cada uma. A decisão final deve ser humana e alinhada aos objetivos do estudo.

O agrupamento por IA deve seguir o princípio da reversibilidade. Todo termo agrupado deve continuar rastreável até sua forma original. A tabela de normalização deve preservar vínculos entre termos, frequência original e decisão tomada. Sem essa rastreabilidade, o pesquisador não consegue explicar por que um *cluster* cresceu, desapareceu ou mudou de posição após a limpeza.

9.7 Verificação humana e amostragem de qualidade

A verificação humana é o mecanismo que transforma limpeza assistida por IA em procedimento cientificamente controlado. Nenhuma regra de normalização deve ser aceita apenas porque produziu resultado visualmente limpo. O pesquisador precisa verificar se as regras são plausíveis em diferentes partes do *corpus*: autores produtivos e periféricos, instituições centrais e regionais, termos frequentes e raros, duplicatas evidentes e limítrofes, registros antigos e recentes.

A amostragem de qualidade pode ser aleatória, estratificada ou orientada ao risco. Amostras aleatórias estimam problemas gerais; amostras estratificadas garantem cobertura de subgrupos; amostras orientadas ao risco focam casos ambíguos. Em estudos sensíveis, dois revisores podem avaliar uma amostra de modo independente, discutir divergências e ajustar regras. Essa prática aumenta a confiabilidade e reduz o risco de que a IA imponha decisões sem controle.

Métricas de erro ajudam a quantificar a qualidade da limpeza. Entre elas estão falso agrupamento, falso desagrupamento, falso positivo de duplicata e falso negativo de duplicata. A taxa aceitável depende do uso pretendido. Um mapa exploratório pode tolerar maior incerteza; uma avaliação institucional exige padrão mais rigoroso. O relatório deve informar não apenas que houve revisão, mas como ela foi feita, quantos casos foram verificados e quais limitações persistem.

A IA pode apoiar a criação de planos de amostragem, checklists e tabelas de auditoria. Pode também resumir divergências entre revisores e apontar padrões de erro. No entanto, não deve ser árbitro final. Modelos generativos podem justificar decisões erradas com linguagem convincente, por isso, a etapa de verificação precisa manter evidência documental, dado original e decisão humana.

A trilha de decisão deve registrar valor original, valor normalizado, regra aplicada, agente responsável, justificativa, evidência e status. Essa estrutura é especialmente relevante quando a limpeza altera indicadores ou rankings. Em termos editoriais, recomenda-se que o capítulo

metodológico de um artigo ou relatório descreva essa trilha de forma sintética e disponibilize a tabela completa quando possível.

9.8 A limpeza como parte do desenho de pesquisa

A preparação dos dados não começa depois da coleta, ela começa no momento em que se formula a pergunta cientométrica. Se o objetivo é analisar a evolução de temas emergentes, a normalização de palavras-chave será central. Se o objetivo é comparar instituições, a desambiguação de afiliações ocupará lugar decisivo. Se a pergunta envolve impacto, será necessário examinar referências citadas, documentos duplicados e variações de DOI. Se o foco está em trajetórias de pesquisadores, a desambiguação de nomes e a preservação da ordem de autoria serão elementos críticos.

A IA pode apoiar esse planejamento ao transformar a pergunta de pesquisa em um plano de preparação de dados. Um *prompt* adequado pode listar campos necessários, riscos de inconsistência, operações de limpeza, critérios de exclusão, regras de agrupamento e amostras de verificação. A contribuição da IA está menos em executar a limpeza e mais em tornar visível o encadeamento de decisões que o pesquisador talvez não tivesse explicitado.

O desenho de pesquisa também deve especificar o nível de tolerância ao erro. Alguns estudos permitem aproximação maior, como mapas exploratórios de uma literatura ampla. Outros exigem alta precisão, como diagnósticos de desempenho institucional, avaliações de programas ou análises que subsidiarão decisões de financiamento. Em análises exploratórias, pode-se solicitar à IA agrupamentos amplos e hipóteses de equivalência. Em análises avaliativas, deve-se exigir evidência forte, identificadores persistentes, confirmação externa e revisão humana.

A limpeza também é um lugar de escolhas éticas. Ao atribuir produção a autores, instituições ou países, o pesquisador distribui reconhecimento. Um erro de desambiguação pode invisibilizar pesquisadores de regiões periféricas, apagar variações linguísticas, privilegiar nomes anglófonos ou reforçar assimetrias entre bases internacionais e produção local. Os princípios do Leiden Manifesto recomendam que métricas apoiem avaliação qualitativa, respeitem diferenças entre áreas e sejam usadas com transparência (HICKS *et al.*, 2015). Esses princípios também se aplicam à preparação dos dados.

A documentação da limpeza deve ser proporcional ao impacto da análise. Em um exercício didático, basta registrar regras principais. Em um artigo científico, convém incluir seção metodológica detalhada e, quando possível, planilhas de normalização. Em relatório de avaliação institucional, recomenda-se preservar logs, versões de base, parâmetros de software, *prompts* e decisões de revisão. Quanto maior a consequência prática da análise, maior deve ser a transparência sobre a preparação dos dados.

9.9 Modelos de prompts para pré-processamento e limpeza

Os *prompts* a seguir foram revisados como instrumentos metodológicos. Eles não devem ser copiados de modo acrítico, mas adaptados à base, à área do conhecimento e ao objetivo do estudo. A literatura sobre modelos de linguagem mostra que técnicas como *prompt engineering*, *chain-of-thought* e *self-consistency* podem melhorar a organização de tarefas complexas, mas não eliminam a necessidade de validação externa e revisão humana (VASWANI *et al.*, 2017; WEI *et al.*, 2022; WANG *et al.*, 2023; CHEN *et al.*, 2025).

Prompt 9.9.1 Prompt central para agrupamento de palavras-chave

Receba a lista de palavras-chave abaixo e agrupe apenas termos equivalentes ou muito próximos.

Para cada grupo, indique: termo preferencial, termos agrupados, justificativa semântica, evidência textual disponível, grau de confiança e ação recomendada.

Não elimine termos ambíguos.

Marque como "revisão humana" todos os casos em que a equivalência depender de contexto disciplinar, tradução imprecisa, sigla ambígua ou diferença conceitual relevante.

Preserve a lista original e não invente descritores externos.

Prompt 9.9.2 Prompt para padronização de nomes de autores

Atue como curador de dados bibliográficos. Analise as variantes de nomes de autores abaixo.

Agrupe apenas variantes que provavelmente se referem ao mesmo pesquisador, considerando coautores, afiliações, ORCID, área temática, título dos trabalhos e período de publicação.

Para cada agrupamento, apresente nome preferencial, variantes, evidências utilizadas, grau de confiança e casos que exigem revisão humana.

Não una autores homônimos apenas por sobrenome e iniciais.

Prompt 9.9.3 Prompt para desambiguação de instituições

Atue como especialista em normalização de afiliações institucionais.

Receba as afiliações abaixo e proponha uma tabela com: afiliação original, instituição normalizada, unidade subordinada quando existir, país, possível identificador ROR quando fornecido no conjunto de dados, justificativa, grau de confiança e alerta de ambiguidade. Não invente identificadores.

Quando houver siglas iguais para instituições diferentes, marque como caso pendente.

Prompt 9.9.4 Prompt para identificação de duplicatas

Analise os registros bibliográficos abaixo e identifique pares ou grupos de possíveis duplicatas.

Use DOI, título normalizado, primeiro autor, ano, periódico ou evento, páginas, volume, resumo e identificadores de base.

Classifique cada par como duplicata provável, possível duplicata ou não duplicata.

Explique a evidência que sustenta a classificação e recomende ação: fundir, manter separado ou revisar manualmente.

Não exclua registros automaticamente.

Prompt 9.9.5 Prompt 9.5 - Prompt para checklist de qualidade

> Com base no protocolo de limpeza descrito abaixo, produza um checklist de verificação humana.
>
> Inclua itens para autores, instituições, duplicatas, palavras-chave, sinônimos, identificadores, campos ausentes, amostragem, registro de decisões e limitações remanescentes.
>
> Para cada item, indique objetivo da verificação, método sugerido, evidência necessária e critério de aceite.

Prompt 9.9.6 Prompt para relatório de auditoria

> A partir da tabela de alterações de limpeza, redija um relatório metodológico sintético.
>
> O relatório deve informar número de registros originais, número de registros após deduplicação, regras aplicadas, campos normalizados, número de casos revisados manualmente, exemplos de alterações, limitações e impacto esperado sobre indicadores e mapas.
>
> Separe fatos observados, decisões humanas e inferências metodológicas.

9.10 Estudo de caso ilustrativo: limpeza de uma base sobre inteligência artificial na educação

Considere uma equipe interessada em mapear a literatura sobre inteligência artificial na educação entre 2015 e 2025. A coleta inicial reúne registros de Scopus, Web of Science e OpenAlex. A base contém títulos, resumos, palavras-chave, autores, afiliações, DOI, ano, periódico, referências citadas e número de citações por fonte. Antes da análise, a equipe percebe que termos como *artificial intelligence*, AI, inteligência artificial, *machine learning*, aprendizagem de máquina, *learning analytics*, *educational data mining* e *adaptive learning* aparecem com sobreposições. Também observa que universidades aparecem em inglês, português e siglas locais. Sem limpeza, o mapa de coocorrência tende a fragmentar o campo e a rede institucional tende a subestimar colaborações.

O primeiro passo é preservar os dados brutos. A equipe cria uma cópia imutável da exportação original e trabalha em uma versão de processamento. Em seguida, define um dicionário de campos: identificador interno, base de origem, DOI, título original, título normalizado, autores originais, autores normalizados, afiliações originais, instituições normalizadas, palavras-chave originais, palavras-chave normalizadas e status de revisão. Essa estrutura evita que alterações sucessivas destruam a rastreabilidade.

Na etapa de autores, o *prompt* solicita agrupamento com base em coautores, afiliação e tema. O modelo sugere grupos de alta confiança e casos de baixa confiança envolvendo sobrenomes comuns e instituições diferentes. A equipe decide manter os casos de baixa confiança separados até confirmação por identificadores ou consulta a fontes externas. Esse exemplo mostra que a melhor resposta da IA pode ser a não decisão.

Na etapa de instituições, aparecem registros como “Univ Sao Paulo”, “Universidade de São Paulo”, “University of São Paulo”, “USP”, “Faculdade de Educação da USP” e “São

Paulo State University". Um *prompt* mal formulado poderia unir indevidamente USP e Universidade Estadual Paulista. Um *prompt* robusto exige país, cidade, evidência textual e marcação de ambiguidade. A equipe agrupa variantes inequívocas da Universidade de São Paulo, mantém unidade subordinada em campo separado e preserva Universidade Estadual Paulista como entidade distinta.

Na etapa de duplicatas, um mesmo artigo aparece em Scopus e Web of Science com DOI idêntico, mas títulos com pequena diferença de pontuação. Esse caso é classificado como duplicata provável e fundido. Outro registro aparece como *preprint* em OpenAlex e como artigo publicado em periódico no ano seguinte. Apesar da semelhança de título e autores, a equipe preserva ambos temporariamente, marcando-os como versões relacionadas. A decisão depende do objetivo analítico.

Na normalização de palavras-chave, o modelo sugere agrupar *"artificial intelligence"*, "AI" e "inteligência artificial", mas marca "AI" como ambíguo em registros de áreas técnicas. Sugere agrupar "*machine learning*" e "aprendizado de máquina", mas manter "*learning analytics*" separado de "*educational data mining*". Essa distinção preserva a estrutura temática do campo e evita transformar subáreas em bloco genérico.

A verificação humana utiliza amostragem estratificada. A equipe seleciona autores produtivos, autores com nomes comuns, instituições centrais, afiliações ambíguas, pares de duplicatas prováveis e palavras-chave normalizadas. Dois revisores analisam a amostra de modo independente. Divergências são discutidas e o protocolo é ajustado. Esse procedimento não elimina todos os erros, mas permite estimar a qualidade da limpeza e identificar regras problemáticas antes da análise final.

O caso mostra que a IA deve ser componente do fluxo, não substituto do método. O modelo ajudou a reduzir o espaço de revisão, produzir justificativas e organizar tabelas. Entretanto, decisões críticas dependeram da pergunta, da interpretação de campo e da validação humana. A IA acelerou a triagem, mas não autorizou eliminação cega de registros nem fusão automática de entidades.

9.11 Matriz de riscos do pré-processamento assistido por IA

A limpeza com apoio de IA envolve riscos específicos. O primeiro é a falsa equivalência, quando o modelo agrupa entidades que parecem semelhantes linguisticamente, mas são distintas no contexto científico. O segundo é a falsa separação, quando variantes de uma mesma entidade permanecem dispersas porque o modelo não reconhece traduções, transliterações ou abreviações. Ambos os erros são importantes: a falsa equivalência produz agregação indevida, a falsa separação produz fragmentação indevida.

O terceiro risco é a alucinação metodológica. O modelo pode afirmar que determinado identificador corresponde a uma instituição, que uma sigla tem certo significado ou que um termo pertence a um tesauro sem que isso esteja nos dados fornecidos. O quarto risco é a normalização excessiva, quando o desejo de produzir mapas mais limpos apaga diferenças conceituais relevantes. O quinto é a dependência de uma única execução, pois modelos generativos podem variar respostas em tarefas semânticas sensíveis. O sexto é a perda de rastreabilidade, quando a saída substitui valores originais sem preservar vínculos.

Esses riscos não invalidam o uso de IA; indicam que ela deve ser incorporada em um sistema de qualidade. A mitigação exige *prompts* restritivos, evidência explícita, grau de confiança, revisão humana, comparação entre execuções, preservação dos dados originais e documentação das decisões. Em conformidade com princípios de avaliação

responsável, a IA deve ampliar a transparência da análise, não reduzir a responsabilidade do pesquisador (HICKS *et al.*, 2015).

Quadro 9.2 - Matriz de riscos e controles no uso de IA para limpeza bibliográfica

Risco	Descrição	Controle recomendado
Falsa equivalência	Agrupar entidades distintas por semelhança textual.	Exigir evidência, grau de confiança e revisão humana.
Falsa separação	Manter variantes de uma mesma entidade em grupos distintos.	Usar identificadores persistentes e metadados auxiliares.
Alucinação	Inventar identificadores, autores, instituições ou relações.	Proibir inferência externa e verificar em fontes confiáveis.
Normalização excessiva	Apagar diferenças conceituais relevantes.	Definir granularidade conforme a pergunta de pesquisa.
Perda de rastreabilidade	Substituir valores sem preservar originais.	Manter colunas de valor original, valor normalizado, regra e justificativa.
Dependência de uma execução	Confiar em uma única resposta do modelo.	Comparar execuções, amostras e revisores humanos.

Fonte: Elaborado pelo autor (2026).

9.12 Protocolo reprodutível de limpeza com prompts

Um protocolo reprodutível de limpeza pode ser organizado em dez etapas.

Figura 9.1 - As 10 etapas de limpeza de dados

Fonte: Elaborado pelo autor (2026).

O uso de IA aparece em várias dessas etapas, mas com funções diferentes. No diagnóstico, ajuda a identificar padrões de inconsistência. Na formulação de regras, sugere critérios e exceções. Na normalização, organiza candidatos. Na auditoria, aponta possíveis problemas remanescentes. Na redação metodológica, ajuda a transformar logs em texto explicativo. Em nenhuma dessas funções a IA dispensa a responsabilidade do pesquisador.

O protocolo deve indicar explicitamente onde houve assistência da IA, onde houve revisão humana e onde houve decisão automatizada por regra.

A reprodutibilidade também exige versionamento. A base v0 corresponde ao dado bruto. A base v1 pode conter normalização mecânica de espaços, acentos e capitalização. A base v2 pode incluir deduplicação. A base v3 pode incluir autores e instituições normalizados. A base v4 pode incluir palavras-chave agrupadas. Cada versão deve ter data, responsável, descrição de alterações e arquivo de regras. Esse procedimento permite retornar a uma versão anterior se uma decisão for contestada ou se a pergunta de pesquisa mudar.

O protocolo deve prever critérios de parada. A limpeza nunca é perfeita; sempre haverá casos ambíguos. Critérios possíveis incluem taxa de erro abaixo de limite definido em amostra, ausência de duplicatas óbvias entre registros mais citados, revisão completa de instituições centrais ou estabilidade dos principais *clusters* após nova rodada de normalização. Sem critérios de parada, a limpeza pode tornar-se interminável ou ser encerrada cedo demais.

Em termos de publicação, o relatório deve explicar quais problemas foram encontrados, quais regras foram aplicadas, que papel a IA desempenhou, quantos casos foram revisados, quais limitações persistem e como essas limitações podem afetar resultados. Essa prática fortalece a credibilidade do estudo e permite que outros pesquisadores avaliem a robustez dos indicadores e mapas apresentados.

Figura 9.2 - Onde e como a IA pode apoiar o pesquisador

Fonte: Imagem gerada por IA, ChatGPT (2026).

9.13 Checklist comentado para equipes de pesquisa

O checklist a seguir substitui o conjunto repetitivo de parágrafos da versão original por uma matriz sintética de controle editorial e metodológico. Cada item deve ser preenchido pela equipe com responsável, data, evidência, decisão e status. A função do checklist é impedir que a limpeza se torne uma sequência de intervenções desconectadas.

Quadro 9.3 - Checklist da verificação dos metadados

Item de controle	Verificação recomendada	Critério de aceite
Integridade dos identificadores	Verificar DOI, PMID, OpenAlex ID, Scopus EID, WoS Accession Number, ORCID e ROR.	Ausências e conflitos registrados; identificadores inválidos marcados para revisão.
Completude dos campos	Mensurar ausência em título, resumo, palavras-chave, autores, afiliações, ano, fonte, referências e citações.	Percentuais de preenchimento documentados antes da análise principal.
Consistência temporal	Conferir ano de publicação, ano de versão, ano de indexação e data de coleta.	Séries temporais baseadas em campo definido e justificado.
Consistência autoral	Comparar variantes nominais, coautorias, afiliações e identificadores.	Casos de baixa confiança preservados para revisão humana.
Consistência institucional	Normalizar instituições no nível adequado e preservar unidades subordinadas quando relevantes.	Regra de granularidade explicitada e reversível.
Consistência terminológica	Revisar termos frequentes, emergentes, siglas, traduções e polissemia.	Tabela de normalização com termo original e termo preferencial.
Duplicatas documentais	Identificar duplicatas por DOI, título normalizado, primeiro autor, ano e fonte.	Pares duvidosos mantidos como pendentes ou revisados manualmente.
Impacto da limpeza	Comparar indicadores, *clusters* e redes antes e depois da limpeza.	Mudanças inesperadas investigadas antes da interpretação.
Amostragem de qualidade	Selecionar amostras aleatórias, estratificadas e orientadas ao risco.	Taxas de erro e divergências de revisão documentadas.
Relato metodológico	Registrar regras, *prompts*, versões, decisões humanas, limitações e arquivos suplementares.	Seção metodológica capaz de sustentar auditoria externa.

Fonte: Elaborado pelo autor (2026).

9.14 Considerações finais

O pré-processamento e a limpeza de dados são condições de possibilidade da cientometria. Sem padronização de autores, desambiguação institucional, deduplicação documental, normalização terminológica e verificação de qualidade, indicadores e mapas podem adquirir aparência científica sem representar adequadamente o fenômeno estudado. A IA amplia a capacidade de lidar com grandes volumes de metadados, mas também aumenta a necessidade de explicitar critérios. O principal erro seria interpretar a fluência do modelo como garantia de correção.

Ao longo do capítulo, defendeu-se que *prompts* devem funcionar como instrumentos de curadoria assistida. Eles ajudam a diagnosticar problemas, propor agrupamentos, criar checklists, transformar critérios em protocolos e redigir relatórios de auditoria. Contudo, *prompts* também podem produzir falsas equivalências, alucinações, justificativas frágeis e apagamento de ambiguidades. A qualidade da saída depende do desenho do *prompt*, da qualidade dos dados fornecidos, da validação humana e da clareza da pergunta cientométrica.

A melhor prática é conservar sempre três camadas: o dado original, o dado transformado e a justificativa da transformação. Essa estrutura simples protege a rastreabilidade e permite

auditoria. Quando enriquecida com *prompts*, logs, amostras e revisão humana, transforma a limpeza em parte explícita do método. Em vez de esconder o trabalho de preparação, o pesquisador passa a apresentá-lo como componente essencial da validade do estudo.

Finalmente, a limpeza não deve buscar uma pureza impossível dos dados. Deve buscar adequação transparente ao problema de pesquisa. Dados bibliográficos são produtos de sistemas de indexação, práticas editoriais, políticas de base, idiomas, campos disciplinares e infraestruturas desiguais. A IA pode ajudar a navegar essa complexidade, mas não eliminá-la. O pesquisador responsável reconhece limitações, registra escolhas e interpreta resultados à luz das operações realizadas. Essa é a base de uma cientometria com IA eficiente, auditável e cientificamente prudente.

Capítulo 10

Indicadores cientométricos

10.1 Introdução

A cientometria contemporânea opera em um cenário paradoxal. Nunca houve tanta disponibilidade de registros bibliográficos, citações, identificadores persistentes, dados de financiamento, vínculos institucionais, perfis autorais, repositórios de código, registros de patentes e rastros de atenção social. Ao mesmo tempo, nunca foi tão evidente que a abundância de dados não elimina a necessidade de interpretação qualificada. Indicadores cientométricos podem revelar padrões de produção, circulação e influência científica, mas também podem induzir leituras simplificadoras quando são tratados como equivalentes diretos de qualidade, mérito ou relevância social. Neste capítulo, os indicadores são apresentados como instrumentos de observação e inferência, não como veredictos automáticos sobre pessoas, instituições, áreas do conhecimento ou políticas de pesquisa (WALTMAN, 2016; HICKS *et al.,* 2015).

No contexto da trilha Cientometria com Inteligência Artificial: *Prompts* para analisar, mapear e interpretar a Ciência, os indicadores ganham uma função adicional. Eles deixam de ser apenas resultados numéricos obtidos por softwares bibliométricos e passam a constituir objetos de diálogo com sistemas de inteligência artificial. Um modelo generativo pode ajudar a selecionar métricas, comparar alternativas, explicar limitações, transformar tabelas em narrativas e advertir contra extrapolações indevidas. Contudo, esse apoio só é metodologicamente aceitável quando as perguntas, as bases, os recortes, os procedimentos de normalização e as limitações são explicitados. A IA pode ampliar a inteligibilidade dos indicadores, mas não pode tornar invisíveis as escolhas que os produzem (LIU *et al.,* 2023; JI *et al.*, 2023).

A história dos indicadores bibliométricos mostra que a mensuração da ciência nasceu associada tanto à necessidade de organizar a literatura quanto ao desejo de compreender a dinâmica da produção científica. A lei de Lotka sobre produtividade autoral, a proposta de Garfield de indexação por citações, os estudos de Price sobre redes de documentos e a noção de cocitação de Small constituem marcos que tornaram possível observar a ciência por meio de registros documentais. Esses marcos não devem ser lidos como origem de uma tecnocracia métrica, mas como expressão de um esforço para tornar analisável a comunicação científica, suas regularidades, assimetrias e estruturas de influência (LOTKA, 1926; GARFIELD, 1955; PRICE, 1965; SMALL, 1973).

Quando se introduz IA generativa nesse campo, o desafio não é substituir a competência bibliométrica por comandos de linguagem natural. O desafio é traduzir a pergunta científica em um conjunto coerente de indicadores, explicitar o que cada métrica mede, reconhecer o que ela não mede e construir uma narrativa interpretativa prudente. Um *prompt* bem formulado pode solicitar que o modelo diferencie produtividade de impacto, impacto acadêmico de atenção social, colaboração formal de cooperação substantiva e visibilidade de qualidade. Essa capacidade de explicitação é particularmente relevante em ambientes de avaliação, nos quais uma métrica mal escolhida pode orientar decisões de financiamento, promoção, contratação ou reorganização institucional.

O capítulo está organizado em sete partes. A primeira discute indicadores de produção, suas possibilidades e armadilhas. A segunda examina indicadores de citação e impacto, incluindo o índice *prompt*, métricas normalizadas e distribuições assimétricas de citações. A terceira trata de colaboração científica, coautoria e redes. A quarta aborda interdisciplinaridade,

diversidade cognitiva e integração entre campos. A quinta discute internacionalização e suas ambiguidades. A sexta apresenta indicadores alternativos e sinais de atenção social. A sétima sintetiza limitações éticas e técnicas do uso de métricas. Ao final, são apresentados *prompts* aplicados para seleção de indicadores, explicação de limitações, interpretação de crescimento de publicações e redação de resultados sem extrapolação.

A perspectiva adotada é a das métricas responsáveis. Isso significa que indicadores devem apoiar, e não substituir, avaliação qualitativa; devem ser adequados aos contextos locais, devem considerar diferenças disciplinares, devem ser transparentes quanto à fonte e ao cálculo, devem reconhecer incertezas e devem evitar consequências adversas, como incentivos a produtivismo, manipulação de citações, fragmentação em salame de resultados e exclusão de formas de contribuição não capturadas por bases bibliográficas. A inteligência artificial, quando usada por *prompts*, deve reforçar esses princípios, funcionando como mecanismo de explicitação, revisão e auditoria, e não como autoridade avaliativa autônoma (HICKS *et al.,* 2015; RUSHFORTH; HAMMARFELT, 2023).

10.2 Indicadores de produção

Indicadores de produção são os instrumentos mais básicos da análise cientométrica. Eles respondem a perguntas como: quantos documentos foram publicados, por quem, em que período, em quais veículos, em quais áreas e sob quais formas documentais. Em sua formulação mais simples, contam artigos, revisões, capítulos, anais de congresso, livros, *preprints*, patentes ou outros produtos. Essa simplicidade é simultaneamente sua vantagem e sua limitação. A contagem de publicações oferece uma primeira aproximação da atividade científica, mas não expressa por si mesma a qualidade epistemológica, a contribuição original, o valor social ou a robustez metodológica dos trabalhos (LOTKA, 1926; DONTHU *et al.*, 2021).

Na perspectiva da produção científica, é necessário distinguir volume de intensidade, intensidade de continuidade e continuidade de maturidade. Um campo emergente pode apresentar crescimento acentuado de publicações sem ter consolidado métodos, conceitos ou consensos. Uma instituição pode aumentar sua produção em razão de novas políticas de indexação, mudança de base, expansão de programas de pós-graduação ou adesão a colaborações multicêntricas. Um autor pode publicar muito porque coordena um grupo amplo, porque participa de redes internacionais, porque atua em área altamente colaborativa ou porque adota práticas de fragmentação de resultados. Portanto, a pergunta cientométrica deve preceder a métrica: não basta perguntar quantos documentos existem, mas que tipo de atividade se deseja caracterizar.

Os principais indicadores de produção incluem número total de publicações, produção anual, taxa de crescimento, distribuição por tipo documental, distribuição por periódico ou conferência, produção por autor, por instituição, por país, por área e por palavra-chave. Em estudos longitudinais, a produção anual permite identificar fases de emergência, crescimento, estabilização ou declínio. Em diagnósticos institucionais, a distribuição por unidade acadêmica ou programa permite observar concentração e dispersão da atividade. Em análises de áreas, a produção por periódico revela canais centrais de comunicação, enquanto a produção por tipo documental indica a cultura de publicação do campo.

A inteligência artificial pode apoiar a interpretação desses indicadores ao transformar séries de produção em narrativas analíticas, mas essa transformação deve ser regulada por *prompts* que impeçam conclusões causalistas. Um aumento de publicações não prova, por si só, maior qualidade, maior financiamento, maior impacto ou maior relevância. O *prompt* deve solicitar que o modelo apresente explicações possíveis, evidências necessárias e

riscos de interpretação. A IA pode sugerir hipóteses, como expansão de uma agenda de pesquisa, maior indexação de periódicos, mudança terminológica ou efeito de políticas de incentivo, mas deve marcar tais hipóteses como inferências dependentes de validação (LIU *et al.*, 2023; JI *et al.*, 2023).

A contagem fracionária é uma questão central nos indicadores de produção. Na contagem inteira, cada unidade participante recebe crédito completo por uma publicação. Na contagem fracionária, o crédito é distribuído entre autores, instituições ou países. A escolha entre esses métodos altera substancialmente rankings e comparações, sobretudo em áreas com grandes colaborações. A contagem inteira tende a favorecer unidades integradas em redes amplas, enquanto a contagem fracionária busca representar contribuição proporcional. Nenhuma escolha é neutra: ela depende da finalidade do estudo. Se o objetivo é mapear presença em redes, a contagem inteira pode ser apropriada, se o objetivo é comparar contribuição relativa, a contagem fracionária tende a ser mais prudente.

Outro problema dos indicadores de produção é a heterogeneidade disciplinar. Áreas biomédicas, físicas, computacionais, humanísticas e artísticas apresentam ritmos e formatos de publicação distintos. Conferências podem ser centrais na ciência da computação, enquanto livros e capítulos têm maior relevância em diversas áreas das humanidades. Comparar unidades sem considerar essas diferenças viola um princípio elementar da cientometria responsável. A IA pode ser instruída a perguntar, antes de sugerir indicadores, qual é a área analisada, quais tipos documentais são relevantes e qual base de dados captura adequadamente esses produtos. Essa etapa evita a imposição de um modelo único de produtividade.

A produção também deve ser situada em relação ao tempo. Um pesquisador em início de carreira não pode ser comparado diretamente a um pesquisador sênior sem ajuste por idade acadêmica. Uma instituição recém-criada não pode ser avaliada pela mesma janela temporal de uma universidade centenária. Um tema emergente pode parecer pouco produtivo em contagem absoluta e, ainda assim, apresentar taxa de crescimento elevada. Por isso, os indicadores de produção devem ser acompanhados de medidas temporais, como produção por ano, crescimento médio anual, crescimento relativo e distribuição por coortes.

Em estudos assistidos por IA, é recomendável que o *prompt* solicite explicitamente uma leitura em três camadas: descrição do padrão, interpretação possível e cautelas. A descrição deve dizer o que os dados mostram; a interpretação deve propor explicações plausíveis; as cautelas devem indicar o que não pode ser concluído. Essa separação reduz a probabilidade de o modelo converter uma tendência numérica em afirmação avaliativa indevida. Também permite que o pesquisador identifique onde faltam dados complementares, como financiamento, número de docentes, tamanho de grupos, políticas editoriais ou mudanças na cobertura da base.

Indicadores de produção são particularmente úteis em diagnósticos de campos emergentes. A identificação de crescimento acelerado, entrada de novos países, diversificação de periódicos e aumento de palavras-chave associadas pode sinalizar a consolidação de uma agenda. No entanto, crescimento quantitativo pode refletir modismos, pressões editoriais, indexação oportunista ou deduplicação terminológica. A IA pode ajudar a comparar termos e detectar mudanças semânticas, mas não deve decidir sozinha que um campo está consolidado. Consolidação exige evidências adicionais, como presença de periódicos especializados, comunidades de conferência, financiamento recorrente, redes de autores, estabilidade conceitual e impacto acumulado.

Em análises institucionais, indicadores de produção precisam ser normalizados por tamanho e missão. Uma universidade com muitos docentes naturalmente publicará mais que uma instituição pequena; um hospital universitário terá perfil diferente de uma

faculdade de artes; um instituto de pesquisa aplicada produzirá documentos distintos de um departamento teórico. A simples contagem total favorece unidades maiores e mais indexadas. Indicadores complementares, como produção por docente, por pesquisador ativo, por programa ou por recurso investido, podem oferecer leitura mais informativa, desde que os denominadores sejam confiáveis.

A IA pode contribuir para a seleção de denominadores, perguntando quais dados institucionais estão disponíveis e alertando para riscos de comparabilidade. Por exemplo, produção por docente exige definição de quem conta como docente ativo; produção por pesquisador exige cadastro atualizado; produção por orçamento exige dados financeiros comparáveis; produção por programa exige classificação consistente das publicações. Um *prompt* útil solicita que o modelo liste variáveis necessárias, fontes possíveis, riscos de viés e procedimentos de validação antes de sugerir qualquer indicador derivado.

Portanto, indicadores de produção são indispensáveis, mas devem ser interpretados como medidas de atividade documentada. Eles dizem algo sobre o volume e a distribuição da comunicação científica registrada, mas não esgotam a compreensão da ciência produzida. Seu uso com IA exige que *prompts* explicitem recorte temporal, base de dados, tipo documental, unidade de análise, método de contagem e finalidade interpretativa. Sem essas condições, a IA tende a produzir narrativas genéricas, frequentemente persuasivas, mas metodologicamente frágeis.

Quadro 10.1 - Indicadores de produção e usos interpretativos

Indicador	O que descreve	Uso adequado	Cautela metodológica
Número total de publicações	Volume de documentos indexados em uma fonte e período.	Caracterizar atividade documentada de uma área, instituição ou autor.	Não mede qualidade, originalidade ou relevância social.
Produção anual	Distribuição temporal da produção.	Identificar crescimento, estabilização, declínio ou efeitos de eventos.	Pode refletir mudanças de cobertura, indexação ou estratégia de busca.
Taxa de crescimento	Variação relativa no tempo.	Comparar fases de expansão de temas ou programas.	Sensível a valores iniciais pequenos e a janelas curtas.
Distribuição por tipo documental	Peso relativo de artigos, revisões, anais, livros, capítulos ou *preprints*.	Adequar a análise à cultura de publicação da área.	Bases podem subrepresentar livros, capítulos ou literatura regional.
Contagem fracionária	Participação proporcional de autores, instituições ou países.	Comparar contribuição relativa em colaborações.	Depende da regra de fracionamento e das convenções de autoria.

Fonte: Elaborado pelo autor (2026).

10.3 Indicadores de citação e impacto

Indicadores de citação e impacto ocupam posição central na cientometria porque prometem captar alguma forma de recepção, uso ou influência dos trabalhos científicos. A citação é um vínculo documental entre um texto que cita e um texto citado. Esse vínculo pode expressar reconhecimento intelectual, dependência conceitual, uso metodológico, crítica,

contextualização, obrigação retórica, convenção disciplinar ou até referência periférica. A literatura sobre comportamento de citação mostra que as citações não têm significado único. Assim, contar citações é contar relações documentais, não medir diretamente qualidade intrínseca (BORNMANN; DANIEL, 2008; TAHAMTAN; BORNMANN, 2019).

A contagem de citações de uma publicação indica quantas vezes ela foi referida por outros documentos indexados em determinada base. Essa medida é sensível à idade do documento, à área do conhecimento, ao idioma, à cobertura da base, ao tipo de documento, ao tamanho da comunidade de pesquisa, à existência de autocitações, à disponibilidade em acesso aberto e à centralidade do tema. Artigos mais antigos tendem a acumular mais citações; áreas com maior densidade de publicação citam mais; revisões frequentemente recebem mais citações que estudos primários; documentos metodológicos podem ter impacto transversal. Portanto, citações brutas raramente devem ser usadas isoladamente para comparação avaliativa.

O fator de impacto de periódicos, o CiteScore, o índice *prompt*, o número médio de citações por documento, o total de citações, a proporção de publicações altamente citadas e os indicadores normalizados por campo e ano são exemplos de métricas de impacto. Cada uma responde a pergunta distinta. O fator de impacto descreve média de citações recebidas por artigos de um periódico em janela específica; o índice *prompt* combina produtividade e citações de um autor, grupo ou instituição; percentis indicam posição relativa de uma publicação em sua área e ano; indicadores normalizados procuram corrigir diferenças disciplinares. A escolha inadequada da métrica distorce a interpretação.

O índice *prompt* é uma das métricas mais difundidas e também uma das mais criticadas. Sua atratividade reside na simplicidade: um pesquisador tem índice *prompt* igual a *prompt* quando *prompt* de seus trabalhos receberam pelo menos *prompt* citações. Essa definição combina produção e impacto em um único número. Entretanto, o índice *prompt* favorece carreiras longas, áreas mais citantes, bases mais amplas e pesquisadores com muitos trabalhos medianamente citados. Ele é pouco sensível a trabalhos extraordinariamente citados, não distingue autoria principal de participação marginal, pode variar conforme a base consultada e é inadequado para comparação entre áreas. Explicar essas limitações é tarefa central da alfabetização cientométrica (HIRSCH, 2005; WALTMAN, 2016).

Modelos generativos podem ajudar a traduzir o índice *prompt* para públicos não especialistas, mas precisam evitar metáforas absolutas. Um *prompt* adequado deve pedir que a IA explique o índice como medida parcial de trajetória citacional, indique exemplos simples, apresente limitações e advirta que ele não deve ser usado sozinho para decisões de carreira. Em contextos institucionais, a IA pode gerar quadros comparativos entre índice *prompt*, total de citações, média de citações, mediana, percentis e indicadores normalizados. Entretanto, os cálculos devem vir de dados verificados; a IA não deve inventar números, completar séries ausentes ou estimar valores sem base.

A normalização por campo e ano é uma das respostas metodológicas à heterogeneidade das citações. Como áreas diferem em densidade citacional, comparar citações brutas de artigos de matemática, biomedicina, engenharia, sociologia e história é metodologicamente problemático. Indicadores normalizados comparam publicações com outras do mesmo campo, ano e tipo documental, buscando avaliar desempenho relativo. Exemplos incluem citações normalizadas por campo, proporção de publicações entre as 10% mais citadas e indicadores baseados em percentis. Ainda assim, a normalização depende de classificações de área, que podem ser imperfeitas, sobretudo para trabalhos interdisciplinares.

As distribuições de citações são fortemente assimétricas. Poucos trabalhos concentram muitas citações, enquanto a maioria recebe poucas ou nenhuma citação em determinado

período. Por isso, médias podem ser enganosas. Em muitos casos, medianas, percentis e distribuição completa são mais informativos que médias simples. Um relatório cientométrico responsável deve indicar se o impacto está concentrado em poucos documentos ou distribuído de maneira ampla. A IA pode auxiliar na redação de uma interpretação que diferencie impacto agregado, impacto típico e presença de *outliers*, mas essa interpretação deve partir de estatísticas reais (BORNMANN; LEYDESDORFF, 2017).

A autocitação merece atenção específica. Autocitações de autores, grupos, periódicos ou instituições podem ser legítimas quando refletem continuidade de pesquisa, mas podem inflar artificialmente indicadores quando usadas estrategicamente. O tratamento das autocitações depende da pergunta. Se o objetivo é medir circulação externa, pode ser apropriado excluí-las ou reportá-las separadamente. Se o objetivo é compreender a trajetória intelectual de um grupo, autocitações podem ser relevantes. A IA deve ser instruída a não tratar autocitação automaticamente como fraude, nem como irrelevante, mas como componente que precisa ser contextualizado.

Citações negativas ou críticas também desafiam interpretações simplistas. Um artigo pode ser muito citado porque é fundamental, controverso, metodologicamente criticado ou historicamente influente. A contagem não distingue automaticamente esses significados. Métodos de análise de contexto de citação e processamento de linguagem natural podem classificar funções citacionais, como uso metodológico, fundamentação teórica, comparação, extensão ou crítica. A IA generativa pode apoiar a leitura de trechos de citação, mas deve operar com evidências textuais e amostras auditáveis. Em estudos cientométricos, a combinação entre indicadores quantitativos e análise qualitativa de citações fortalece a interpretação.

Indicadores de impacto também são afetados pela fonte de dados. Web of Science, Scopus, Google Scholar, Dimensions, OpenAlex, Lens, Crossref e OpenCitations têm coberturas diferentes. Algumas bases incluem mais livros, teses, *preprints* ou documentos não tradicionais; outras priorizam periódicos selecionados. Como resultado, o mesmo autor ou artigo pode ter contagens diferentes em bases distintas. Um relatório deve informar a fonte, a data de coleta e as regras de deduplicação. A IA pode comparar vantagens e limites das bases, mas não deve harmonizar automaticamente números incompatíveis (MARTÍN-MARTÍN *et al.*, 2018; MARTÍN-MARTÍN *et al.*, 2021).

Em avaliação institucional, indicadores de citação devem ser usados com especial prudência. Métricas de impacto podem apoiar diagnóstico, identificar áreas de influência e orientar perguntas estratégicas, mas não devem substituir julgamento por pares, análise de missão institucional e consideração de impactos não bibliométricos. O Manifesto de Leiden recomenda que indicadores quantitativos apoiem avaliação qualitativa, que dados e métodos sejam transparentes e que diferenças entre áreas sejam consideradas. A IA pode ser instruída a revisar um texto de resultados à luz desses princípios, apontando afirmações excessivas e lacunas de contextualização.

A escrita de resultados de impacto é uma área em que *prompts* são especialmente úteis. Um bom *prompt* pode pedir que o modelo transforme uma tabela de citações em texto com quatro restrições: descrever apenas padrões observados, não inferir qualidade intrínseca, mencionar limitações da fonte e indicar a necessidade de normalização. Essa técnica reduz a retórica inflacionada comum em relatórios métricos. O objetivo não é empobrecer a narrativa, mas torná-la epistemicamente responsável.

Quadro 10.2 - Indicadores de citação e perguntas correspondentes

Indicador	Pergunta que responde	Limitação principal
Total de citações	Quanto a coleção foi citada na base escolhida?	Favorece documentos antigos, áreas mais citantes e bases com maior cobertura.
Média de citações	Qual é o impacto médio por documento?	Pode ser distorcida por poucos documentos muito citados.
Mediana de citações	Qual é o impacto típico da coleção?	Pode ocultar *outliers* relevantes.
Índice prompt	Há combinação entre volume de produção e citações acumuladas?	Favorece carreiras longas e não é comparável entre áreas.
Percentil de citações	Qual é a posição relativa do documento em sua área e ano?	Depende da classificação de campo e do conjunto de referência.
Top 10% mais citados	Qual proporção da produção atinge impacto elevado relativo?	Exige base consistente e normalização apropriada.

Fonte: Elaborado pelo autor (2026).

10.4 Indicadores de colaboração

Indicadores de colaboração procuram medir a dimensão coletiva da produção científica. A ciência moderna é amplamente colaborativa, embora o grau e a forma da colaboração variem entre áreas. A coautoria é o indicador mais comum porque está disponível em metadados bibliográficos e permite construir redes de autores, instituições e países. No entanto, coautoria é uma aproximação documental da colaboração, não sua totalidade. Pessoas podem colaborar sem assinar publicações; podem assinar por critérios de grupo; podem contribuir com dados, infraestrutura, financiamento, supervisão, análise ou redação em graus muito distintos. Portanto, indicadores de coautoria precisam ser lidos como sinais formais de colaboração registrada (KATZ; MARTIN, 1997; NEWMAN, 2001).

Entre os indicadores básicos de colaboração estão o número médio de autores por documento, a proporção de documentos em coautoria, a proporção de colaboração internacional, o número de instituições por artigo, o número de países por publicação, o índice de colaboração e medidas de centralidade em redes. Esses indicadores permitem identificar se um campo é predominantemente individual, local, institucional, nacional ou internacional. Também permitem observar a formação de comunidades, a presença de atores ponte e a distribuição da conectividade entre unidades.

A análise de redes amplia a leitura dos indicadores de colaboração. Em uma rede de coautoria, nós podem representar autores, instituições ou países; arestas representam publicações compartilhadas. A força da aresta pode indicar número de documentos em conjunto. Medidas como grau, centralidade de intermediação, centralidade de proximidade, centralidade de autovetor, densidade, modularidade e componentes conectados permitem caracterizar a estrutura relacional. Contudo, tais medidas não devem ser usadas mecanicamente. Um autor com alta centralidade pode ser líder intelectual, coordenador administrativo, participante de grandes consórcios ou simplesmente alguém inserido em uma área altamente colaborativa (NEWMAN, 2001; VAN ECK; WALTMAN, 2010).

A inteligência artificial pode apoiar a interpretação de redes, especialmente quando há necessidade de transformar medidas abstratas em explicações compreensíveis. Um *prompt* pode pedir que o modelo diferencie descrição estrutural, hipóteses explicativas e evidências necessárias. Por exemplo, uma rede com poucos *clusters* e alta densidade pode sugerir comunidade integrada, mas também pode refletir base de dados restrita ou período curto. Uma rede fragmentada pode indicar campo emergente, subáreas independentes ou falha na padronização de nomes. A IA deve ser instruída a apresentar alternativas interpretativas, não conclusões únicas.

O método de contagem afeta profundamente indicadores de colaboração. Na contagem inteira, cada país ou instituição participante recebe crédito completo por uma publicação colaborativa. Na contagem fracionária, o crédito é dividido entre participantes. Em análises de internacionalização, a contagem inteira pode inflar a produção de países altamente conectados a consórcios; a contagem fracionária pode reduzir a visibilidade de sua participação em redes estratégicas. Ambas as perspectivas são úteis, desde que a finalidade seja explícita. A IA pode auxiliar na escolha entre contagem inteira e fracionária, perguntando se o objetivo é medir presença, contribuição proporcional ou capacidade de articulação.

A colaboração não é sempre um bem em si. Embora trabalhos colaborativos e internacionais frequentemente apresentem maior visibilidade citacional, essa associação não implica causalidade simples. Colaborações podem ampliar acesso a recursos, diversidade de competências, redes de citação e canais de publicação; também podem reproduzir assimetrias de liderança, dependência científica e desigualdade na distribuição de autoria. Indicadores de colaboração devem, portanto, ser combinados com análise de posição autoral, autoria correspondente, financiamento, afiliações e liderança intelectual.

A posição de autoria é uma informação importante em algumas áreas, mas seu significado varia. Em biomedicina, primeiro autor e último autor costumam ter interpretações específicas; em matemática, a ordem alfabética pode ser comum; em grandes consórcios, regras coletivas podem prevalecer. Indicadores que atribuem liderança por posição autoral devem ser usados apenas quando a convenção disciplinar é conhecida. A IA pode ajudar a identificar possíveis convenções, mas deve pedir confirmação da área e não assumir regras universais.

Redes de colaboração também são sensíveis à desambiguação de autores e instituições. Homônimos, variações de nomes, mudanças de afiliação, siglas institucionais e fusões organizacionais podem criar nós falsos ou fragmentar atores reais. Antes de calcular centralidade, é necessário limpar dados. Caso contrário, um indicador sofisticado apenas quantifica erro de metadados. *Prompts* de auditoria podem pedir que a IA identifique nomes potencialmente duplicados, instituições equivalentes e padrões de afiliação inconsistentes, sempre marcando casos duvidosos para revisão humana.

Indicadores de colaboração são úteis para políticas de pesquisa. Eles permitem mapear parceiros estratégicos, lacunas de cooperação, dependência de poucos atores, potencial de internacionalização e capacidade de articulação interdisciplinar. Em uma universidade, podem revelar que a produção de alto impacto está concentrada em redes internacionais específicas; em um programa de pós-graduação, podem mostrar isolamento institucional; em um país, podem evidenciar assimetrias regionais. Contudo, a política derivada desses indicadores deve considerar objetivos institucionais, áreas prioritárias e condições de infraestrutura.

A IA pode ser usada para elaborar diagnósticos de colaboração a partir de tabelas e medidas de rede. Um *prompt* responsável deve pedir que o modelo não confunda coautoria com colaboração substantiva, não atribua liderança sem evidência, não interprete centralidade

como qualidade e não recomende estratégias sem considerar missão institucional. Dessa forma, a IA atua como tradutora e crítica, não como decisora.

Quadro 10.3 - Indicadores de colaboração e riscos de leitura

Indicador	Interpretação possível	Risco de extrapolação
Média de autores por documento	Grau de colaboração formal por publicação.	Não indica intensidade real nem qualidade da cooperação.
Proporção de coautoria internacional	Inserção transnacional da produção.	Pode refletir convenções da área ou consórcios amplos.
Centralidade em rede de coautoria	Posição relacional de autores ou instituições.	Não deve ser confundida com liderança intelectual.
Densidade da rede	Grau de conectividade entre atores.	Redes densas podem resultar de recorte restrito ou erro de desambiguação.
Modularidade	Existência de comunidades ou *clusters*.	*Clusters* dependem do algoritmo, do limiar e da qualidade dos dados.

Fonte: Elaborado pelo autor (2026).

10.5 Indicadores de interdisciplinaridade

Indicadores de interdisciplinaridade buscam captar a circulação de conhecimentos entre áreas, disciplinas, especialidades ou comunidades epistêmicas. O tema é metodologicamente complexo porque interdisciplinaridade pode significar variedade de referências, integração conceitual, colaboração entre departamentos, publicação em periódicos de diferentes áreas, uso de métodos heterogêneos ou impacto em campos diversos. Cada definição exige indicadores diferentes. Antes de medir interdisciplinaridade, é necessário dizer se a pergunta trata de diversidade, integração, difusão, combinação metodológica ou posição de ponte entre campos (CASSI *et al.*, 2017; LEYDESDORFF; WAGNER; BORNMANN, 2019).

Uma das abordagens mais comuns mede interdisciplinaridade pela diversidade das referências citadas. Um artigo que cita literatura de várias áreas pode ser considerado mais interdisciplinar que outro que se apoia em um campo homogêneo. Medidas de diversidade combinam três dimensões: variedade, equilíbrio e disparidade. Variedade refere-se ao número de categorias distintas; equilíbrio indica como as citações se distribuem entre categorias; disparidade expressa a distância cognitiva entre as categorias. O índice Rao-Stirling é uma medida conhecida porque procura integrar essas dimensões (CASSI *et al.*, 2017).

Outra abordagem observa a diversidade dos periódicos citantes ou citados. Um trabalho pode ser interdisciplinar porque recebe citações de múltiplas áreas, mesmo que tenha sido produzido dentro de uma disciplina. Essa diferença entre interdisciplinaridade de entrada e de saída é relevante. A primeira indica diversidade das fontes mobilizadas; a segunda indica difusão em diferentes comunidades. Em termos de IA, um *prompt* deve pedir que o modelo diferencie integração de conhecimentos, difusão de resultados e mera variedade classificatória.

Indicadores de interdisciplinaridade dependem fortemente das classificações disciplinares. Bases como Web of Science e Scopus atribuem categorias a periódicos, não necessariamente a artigos. Um periódico multidisciplinar pode reunir trabalhos muito distintos; um artigo interdisciplinar pode ser publicado em periódico disciplinar; uma categoria pode ser ampla demais para captar diferenças internas. O uso de classificações

algorítmicas por *clusters* de citação ou por tópicos pode oferecer alternativas, mas também introduz dependências metodológicas. Portanto, resultados de interdisciplinaridade devem informar a classificação utilizada.

A inteligência artificial pode apoiar a análise interdisciplinar pela leitura de títulos, resumos e palavras-chave, sugerindo áreas envolvidas e relações conceituais. Contudo, essa classificação textual deve ser validada. Modelos generativos podem reconhecer termos de áreas diferentes, mas também podem superestimar interdisciplinaridade quando um termo é polissêmico. A palavra "rede", por exemplo, pode pertencer à sociologia, computação, epidemiologia, bibliometria ou teoria organizacional. Um *prompt* adequado solicita que o modelo indique evidências textuais, grau de confiança e termos ambíguos.

A interdisciplinaridade não deve ser confundida automaticamente com excelência. Trabalhos interdisciplinares podem ser inovadores, mas também podem ser superficiais se apenas combinam vocabulários sem integração teórica. Da mesma forma, trabalhos disciplinares podem ser profundamente relevantes e rigorosos. Indicadores de interdisciplinaridade informam sobre composição cognitiva e circulação entre campos, não sobre superioridade científica. A IA deve ser instruída a evitar frases como "quanto mais interdisciplinar, melhor", substituindo-as por leituras contextualizadas: interdisciplinaridade pode ser desejável quando a pergunta exige múltiplos conhecimentos.

Em diagnósticos institucionais, indicadores de interdisciplinaridade podem revelar capacidades de conexão entre departamentos, centros e programas. Uma universidade pode desejar mapear áreas que funcionam como pontes cognitivas, programas que combinam métodos ou grupos que conectam saúde, tecnologia e ciências sociais. Contudo, a leitura de interdisciplinaridade institucional exige cuidado com escala. Uma instituição grande parecerá mais diversa por tamanho; uma pequena pode ser especializada por missão. Indicadores devem ser normalizados e interpretados à luz da estratégia institucional.

A IA pode auxiliar na construção de perguntas sobre interdisciplinaridade. Em vez de perguntar genericamente "esta instituição é interdisciplinar?", um *prompt* pode ajudar a formular questões como: "quais áreas aparecem combinadas nas referências?", "quais campos citam os trabalhos da instituição?", "quais programas atuam como pontes em redes de coautoria?", "quais temas emergentes combinam vocabulários de áreas distintas?". Essas perguntas são operacionalizáveis e permitem escolher indicadores adequados.

Há também interdisciplinaridade socialmente orientada, quando a pesquisa conecta conhecimento acadêmico a problemas complexos, políticas públicas, inovação ou demandas sociais. Essa dimensão nem sempre aparece em citações disciplinares. Indicadores alternativos, análise de políticas, patentes, documentos técnicos e menções em relatórios podem complementar a análise. A IA pode apoiar a triagem desses materiais, mas deve distinguir circulação social de validação científica.

Assim, indicadores de interdisciplinaridade são poderosos quando tratados como medidas de diversidade cognitiva e relações entre campos. Seu uso exige definição conceitual, escolha de unidade de análise, classificação explícita, cautela com escala e interpretação qualitativa. *Prompts* bem elaborados podem ajudar a evitar o uso ornamental do termo interdisciplinaridade, convertendo-o em objeto mensurável e discutível.

10.6 Indicadores de internacionalização

Indicadores de internacionalização descrevem a presença, a circulação e a articulação transnacional da produção científica. Eles podem incluir proporção de publicações com

coautoria internacional, número de países participantes, diversidade geográfica dos parceiros, citações recebidas de outros países, publicações em periódicos internacionais, participação em redes globais, autoria correspondente estrangeira ou colaboração com instituições de alto prestígio. Entretanto, internacionalização não é sinônimo automático de qualidade, excelência ou relevância. Trata-se de uma dimensão relacional e geopolítica da ciência.

A colaboração internacional é frequentemente associada a maior impacto citacional, mas a relação é mediada por múltiplos fatores. Projetos internacionais podem envolver maior escala, financiamento mais robusto, acesso a infraestrutura, diversidade de competências, visibilidade editorial e redes de difusão. Ao mesmo tempo, podem refletir desigualdades na divisão do trabalho científico, em que determinados países contribuem com dados, campo empírico ou amostras, enquanto outros concentram liderança analítica e autoria correspondente. Indicadores de internacionalização devem, portanto, ser acompanhados de análise de papéis.

A proporção de publicações em coautoria internacional é o indicador mais usado. Ela é calculada como percentual de documentos com autores afiliados a mais de um país. É útil para comparar abertura de redes, mas tem limitações. Países pequenos tendem a colaborar internacionalmente com maior frequência por necessidade de escala; países grandes podem ter redes domésticas suficientes para produzir sem coautoria internacional; áreas como física de partículas apresentam colaboração global estrutural; humanidades podem ter padrões mais locais e linguísticos. Comparações sem contexto podem produzir injustiças interpretativas.

A diversidade de parceiros internacionais pode ser medida pelo número de países distintos, pela distribuição das parcerias ou por índices de concentração. Uma instituição pode ter muitas publicações internacionais concentradas em um único país parceiro, ou menos publicações distribuídas em rede mais diversa. A escolha depende da pergunta. Se o objetivo é fortalecer parceria estratégica, concentração pode ser positiva; se o objetivo é diversificar exposição, maior variedade pode ser desejável. A IA pode ajudar a explicitar essas alternativas, mas não deve impor uma hierarquia única.

A internacionalização também se relaciona à circulação linguística. Publicações em inglês tendem a ter maior visibilidade em bases internacionais, mas a produção em português, espanhol ou outras línguas pode ser essencial para problemas locais, políticas públicas e formação acadêmica. Uma métrica que privilegia apenas indexação internacional pode desvalorizar conhecimento orientado a contextos nacionais. Indicadores responsáveis devem reconhecer que a missão de uma instituição pode incluir tanto circulação global quanto relevância local.

Outro indicador relevante é a liderança internacional. Publicar com autores estrangeiros não significa liderar a colaboração. A posição de autoria, a autoria correspondente, a instituição de afiliação principal, a coordenação do projeto e a origem do financiamento podem indicar papéis diferentes. Em algumas áreas, a autoria correspondente é aproximação de liderança; em outras, não. A IA pode sugerir indicadores de liderança, mas deve pedir confirmação das convenções disciplinares e dos metadados disponíveis.

As redes de países e instituições permitem visualizar a internacionalização. Mapas de coautoria internacional mostram centros, periferias, corredores de cooperação e comunidades regionais. Entretanto, mapas podem reforçar ilusões visuais: linhas espessas podem representar muitos documentos, mas não necessariamente colaboração profunda; nós grandes podem resultar de tamanho populacional ou investimento; ausência de conexão pode refletir lacunas da base. A IA pode redigir legendas analíticas que alertem para esses limites, evitando que visualizações sejam interpretadas como fotografia completa da ciência global.

Internacionalização também pode ser analisada por impacto. Citações internacionais, presença em periódicos de circulação global e inserção em frentes de pesquisa internacionais indicam recepção além do contexto doméstico. Contudo, impacto internacional não substitui impacto social nacional. Pesquisas sobre saúde pública, educação, desenvolvimento regional e políticas locais podem ser altamente relevantes mesmo com menor circulação internacional. O *prompt* de interpretação deve pedir que o modelo separe visibilidade internacional, colaboração transnacional e relevância pública.

Em políticas científicas, indicadores de internacionalização podem orientar programas de mobilidade, acordos de cooperação, redes temáticas e estratégias de inserção global. Mas podem também gerar incentivos perversos, como busca de parcerias meramente formais ou desvalorização de agendas nacionais. A avaliação responsável deve reconhecer pluralidade de missões. Uma universidade regional pode ter como objetivo resolver problemas locais; um instituto nacional pode buscar soberania tecnológica; um programa de excelência pode priorizar redes globais. Os indicadores devem servir a essas missões, não substitui-las.

A IA pode apoiar diagnósticos de internacionalização se for alimentada com dados consistentes: lista de países, frequência de coautoria, áreas envolvidas, posição autoral, citações, períodos e bases. Um *prompt* robusto deve exigir que a resposta distinga fatos observados, inferências plausíveis e recomendações condicionais. Deve também perguntar quais dados faltam para avaliar liderança, reciprocidade e benefício mútuo. Assim, indicadores de internacionalização deixam de ser números de prestígio e tornam-se instrumentos de análise estratégica.

10.7 Indicadores alternativos e atenção social

Indicadores alternativos e sinais de atenção social surgiram como tentativa de captar dimensões da circulação científica não representadas por citações acadêmicas tradicionais. Eles incluem menções em redes sociais, notícias, *b*logs, documentos de política pública, repositórios de referência, plataformas de leitura, downloads, visualizações, comentários, Wikipedia, patentes e outros ambientes digitais. O termo altmetria indica um conjunto heterogêneo de rastros de atenção, discussão e uso, não uma métrica única. Por isso, é inadequado falar em "impacto alternativo" de forma genérica sem especificar fonte, evento e público (THELWALL *et al.*, 2013; ERDT *et al.*, 2016).

A principal vantagem desses indicadores é a velocidade e a amplitude. Enquanto citações acadêmicas podem levar anos para se acumular, menções em notícias, redes sociais ou documentos de política podem ocorrer rapidamente. Isso é útil para observar difusão inicial, repercussão pública e circulação fora da literatura científica. Em áreas aplicadas, indicadores de atenção social podem revelar que um estudo foi usado em debate público, diretriz clínica, relatório governamental ou mídia especializada. Contudo, atenção não é sinônimo de qualidade, aprovação ou influência positiva.

Menções em redes sociais são particularmente ambíguas. Um artigo pode ser muito compartilhado porque é relevante, controverso, mal interpretado, politicamente mobilizado ou divulgado por estratégias de comunicação. A contagem de *tweets*, posts ou compartilhamentos não indica leitura cuidadosa, concordância ou uso científico. Após mudanças em plataformas digitais e restrições de acesso a dados, a estabilidade desses indicadores tornou-se ainda mais problemática. Portanto, a altmetria deve ser interpretada como sinal contextual de atenção, não como medida direta de impacto científico (HAUSTEIN *et al.*, 2014; ERDT *et al.,* 2016).

Mendeley readership, ou leitores salvos em gerenciadores de referência, é frequentemente tratado como sinal intermediário entre atenção e uso acadêmico. Leitores podem ser

estudantes, pesquisadores, profissionais ou interessados. Estudos mostram correlações entre leitores e citações em certas áreas, mas a interpretação depende da composição do público. Um número alto de leitores pode indicar uso educacional, interesse inicial ou potencial de citação futura. A IA pode ajudar a interpretar esse sinal, desde que receba informações sobre tipo de fonte e público provável (THELWALL *et al.*, 2013).

Documentos de política pública são especialmente relevantes para medir circulação social da ciência. Quando uma publicação é citada em relatórios governamentais, diretrizes técnicas ou documentos de organismos internacionais, há indício de uso em processos decisórios. Mesmo assim, é necessário analisar contexto: a citação pode fundamentar recomendação, fornecer dado, contextualizar problema ou aparecer apenas em bibliografia. Modelos de linguagem podem auxiliar a ler trechos de documentos de política e classificar funções de uso, mas a validação humana é indispensável.

Indicadores de notícias e mídia captam visibilidade pública. Eles podem ser relevantes para estudos sobre comunicação científica, controvérsias públicas e impacto social percebido. Contudo, a cobertura midiática é influenciada por assessorias, interesse jornalístico, temas sensíveis, idioma, localização e eventos conjunturais. Uma métrica de notícias não deve ser usada para comparar qualidade de pesquisas de áreas diferentes. Também pode haver viés em favor de temas com apelo público imediato, em detrimento de pesquisas fundamentais de longo prazo.

Patentes e citações em patentes podem indicar relação com inovação tecnológica, mas sua interpretação também é complexa. Nem toda pesquisa relevante gera patente; nem toda patente representa inovação socialmente útil; nem toda citação patentária indica aplicação direta. Em áreas tecnológicas, tais indicadores podem complementar análise de transferência de conhecimento, mas devem ser combinados com dados de contratos, licenças, *startups*, normas técnicas e adoção industrial. A IA pode ajudar a mapear possíveis rotas de aplicação, mas não deve transformar menção em patente em prova de impacto econômico.

Downloads e visualizações são sinais de acesso, não necessariamente de leitura ou uso. Eles podem ser afetados por acesso aberto, mecanismos de busca, *bots*, campanhas de divulgação, interface do repositório e contagem duplicada. Ainda assim, podem ser úteis em repositórios institucionais para observar alcance de teses, relatórios, *datasets* e materiais educacionais que não recebem muitas citações formais. O pesquisador deve distinguir acesso, engajamento, uso e impacto.

A IA tem papel importante na integração interpretativa de indicadores alternativos. Como esses sinais são heterogêneos, um *prompt* pode pedir que o modelo classifique cada indicador segundo dimensão predominante: atenção pública, uso acadêmico inicial, uso em política, potencial tecnológico, engajamento educacional ou circulação midiática. O *prompt* deve exigir que a IA apresente limitações específicas para cada fonte. Essa abordagem impede a agregação acrítica em escores compostos, que frequentemente misturam eventos incomparáveis.

Indicadores alternativos são promissores para complementar a cientometria, mas não substituem citações, revisão por pares, avaliação qualitativa ou evidências de impacto social. Seu valor está em abrir novas perguntas: quem presta atenção à ciência? Onde ela circula? Que públicos a mobilizam? Que documentos a reutilizam? Que temas ultrapassam fronteiras acadêmicas? Quando usados com IA, esses indicadores exigem *prompts* que diferenciem atenção de influência, visibilidade de qualidade e repercussão de benefício social.

10.8 Limitações éticas e técnicas do uso de métricas

As limitações éticas e técnicas dos indicadores cientométricos são inseparáveis de seu uso prático. A primeira limitação é a redução da complexidade. A ciência envolve criatividade, rigor, formação de pessoas, abertura de dados, desenvolvimento metodológico, impacto social, contribuição teórica, manutenção de infraestrutura, revisão por pares e muitas outras atividades. Nenhuma métrica isolada representa essa pluralidade. Quando um indicador é transformado em objetivo, os agentes podem adaptar comportamentos para maximizá-lo, produzindo distorções conhecidas como efeitos de incentivo (HICKS *et al.*, 2015; RUSHFORTH; HAMMARFELT, 2023).

A segunda limitação é a dependência de bases de dados. Indicadores são tão confiáveis quanto os registros que os sustentam. Cobertura desigual por idioma, área, tipo documental, região e período afeta resultados. Bases comerciais tendem a selecionar periódicos segundo critérios próprios; bases abertas podem ter metadados incompletos; perfis autorais podem conter duplicatas ou atribuições erradas; dados de afiliação podem ser inconsistentes. A IA não corrige magicamente esses problemas. Ao contrário, pode ocultá-los se produzir narrativa fluente sem explicitar incertezas.

A terceira limitação é a comparabilidade. Comparar autores, grupos, instituições ou países exige que se considere área, idade acadêmica, tamanho, missão, recursos, cultura de publicação e janela temporal. A métrica que parece neutra pode favorecer sistematicamente determinados campos, idiomas ou formatos. O uso responsável de indicadores requer normalização, contextualização e, muitas vezes, recusa de comparações simplistas. A IA pode ser instruída a identificar comparações injustas e sugerir ajustes, mas a decisão final depende de juízo metodológico e político.

A quarta limitação refere-se à performatividade das métricas. Indicadores não apenas descrevem a ciência; eles podem alterar a ciência. Quando programas de avaliação privilegiam quantidade de artigos, pesquisadores podem fragmentar resultados. Quando rankings valorizam citações, instituições podem incentivar publicação em áreas mais citantes. Quando periódicos buscam fator de impacto, podem priorizar revisões ou temas de alta visibilidade. Quando *altmetrics* são valorizadas, estratégias de divulgação podem ser confundidas com impacto. Assim, indicadores têm efeitos sociais e devem ser tratados como tecnologias de governança.

A quinta limitação diz respeito à ética da avaliação individual. Métricas agregadas podem ser úteis para analisar campos, instituições e programas, mas tornam-se perigosas quando aplicadas rigidamente a indivíduos. Carreiras são heterogêneas; períodos de licença, cuidado, gestão, ensino, extensão, mudança de país, desigualdades de gênero e raça, acesso desigual a redes e recursos influenciam trajetórias. O uso de indicadores para promoção, contratação ou financiamento deve ser acompanhado por avaliação qualitativa e sensibilidade a contextos. A IA não deve automatizar julgamento de pessoas com base em métricas.

A sexta limitação é a incerteza dos próprios cálculos. Variações de data de coleta, atualização de bases, deduplicação, atribuição de autoria, remoção de autocitações, escolha de janela temporal e método de contagem podem alterar resultados. Relatórios cientométricos devem documentar essas escolhas. A IA pode ajudar a produzir um registro metodológico, mas deve ser alimentada com informações precisas. Sem isso, pode inventar procedimentos ou apresentar cálculos como se fossem universais.

A sétima limitação é a interpretação causal. Indicadores observam associações e padrões, mas raramente demonstram causalidade. O fato de publicações internacionais receberem mais citações não prova que a internacionalização causou o impacto; o fato de

uma área crescer em publicações não prova que ela é mais relevante; o fato de um artigo receber muitas menções não prova que influenciou políticas. A escrita cientométrica deve distinguir correlação, associação temporal, hipótese e causalidade. *Prompts* podem exigir essa distinção de forma explícita.

A oitava limitação envolve viés algorítmico e opacidade em sistemas de IA. Quando modelos generativos participam da análise, eles podem reproduzir vieses de treinamento, favorecer literatura anglófona, sugerir referências inexistentes, atribuir significado indevido a métricas ou construir narrativas excessivamente confiantes. O uso de IA em indicadores exige auditoria: separar dados fornecidos pelo pesquisador de inferências geradas pelo modelo; exigir nível de incerteza; verificar referências; manter logs de *prompts*; preservar planilhas e scripts; e submeter interpretações a revisão especializada (JI *et al.*, 2023; LIU *et al.*, 2023).

A nona limitação é a tendência à agregação indevida. Escores compostos que combinam produção, citações, *altmetrics*, internacionalização e colaboração podem parecer sofisticados, mas frequentemente misturam dimensões incomensuráveis. A agregação só é justificável quando pesos, escalas, objetivos e fundamentos são explicitados. Caso contrário, o indicador composto torna-se uma caixa-preta avaliativa. A IA pode ajudar a criticar escores compostos, perguntando que dimensões foram combinadas, com que pesos, para qual finalidade e com quais consequências.

A décima limitação é a ausência de voz dos avaliados. Indicadores são frequentemente construídos por avaliadores, gestores ou plataformas, sem participação dos pesquisadores e comunidades afetadas. Métricas responsáveis exigem transparência, possibilidade de contestação e adequação à missão. Em uma abordagem assistida por IA, a documentação dos *prompts* e das decisões analíticas pode facilitar auditoria e debate. O objetivo não é eliminar indicadores, mas civilizar seu uso.

Em síntese, as limitações éticas e técnicas não tornam os indicadores inúteis; tornam sua interpretação exigente. A cientometria responsável não abandona métricas, mas as contextualiza. A IA pode fortalecer esse processo se for usada para explicitar pressupostos, listar riscos, revisar afirmações e gerar versões prudentes de resultados. Porém, quando usada para acelerar relatórios sem validação, a IA pode amplificar exatamente os problemas que a bibliometria crítica procura corrigir.

10.9 Prompts aplicados ao uso de indicadores cientométricos

O uso de *prompts* em indicadores cientométricos deve obedecer a uma lógica de pergunta, dado, método, interpretação e limite. Um *prompt* não deve pedir apenas "analise estes indicadores". Deve especificar o papel do modelo, a finalidade do estudo, a unidade de análise, a fonte de dados, o período, os indicadores disponíveis, o público-alvo, o formato de saída e as restrições interpretativas. Quanto mais sensível for o uso avaliativo, maior deve ser a exigência de prudência, rastreabilidade e separação entre fato e inferência.

Prompts para seleção de indicadores são úteis quando o pesquisador ainda está formulando o desenho do estudo. A IA pode sugerir métricas de produção, impacto, colaboração, internacionalização ou atenção social de acordo com a pergunta. Contudo, a resposta deve incluir justificativa, dados necessários, limitações e alternativas. O modelo não deve recomendar uma métrica porque ela é popular, mas porque responde a uma dimensão analítica da pergunta. A popularidade do índice *prompt*, por exemplo, não justifica seu uso em análise de pesquisadores em início de carreira ou em comparação interdisciplinar.

Prompts para explicar indicadores a públicos não especialistas têm função pedagógica. Gestores, docentes, estudantes e formuladores de políticas frequentemente usam métricas sem conhecer seus pressupostos. A IA pode transformar uma explicação técnica em linguagem acessível, desde que preserve cautelas. Um *prompt* adequado pode pedir exemplos, analogias limitadas, contraindicações e uma seção "o que este indicador não mede". Essa última parte é fundamental para combater a falsa autoridade dos números.

Prompts para interpretar crescimento de publicações devem evitar causalidade. O aumento de produção em uma área pode indicar expansão real, mudança terminológica, maior cobertura da base, efeito de políticas editoriais, crescimento de financiamento, aumento de *preprints* ou modismo temporário. Um bom *prompt* pede que a IA apresente essas hipóteses e indique dados necessários para testá-las. A interpretação deve distinguir tendência observada de explicação causal.

Prompts para escrever resultados sem extrapolação são particularmente importantes. Relatórios cientométricos muitas vezes transformam achados modestos em afirmações fortes. Um texto responsável deve dizer "a produção indexada na base X aumentou no período *PROMPT*", não "a área tornou-se mais importante" sem evidência adicional. Deve dizer "os documentos receberam maior média de citações", não "os documentos têm maior qualidade". A IA pode revisar o texto procurando termos avaliativos indevidos e sugerindo formulações mais precisas.

A documentação dos *prompts* também é parte do método. Quando a IA é usada para selecionar indicadores, interpretar resultados ou redigir seções, o relatório deve indicar como ela foi utilizada. Não é necessário publicar toda interação quando isso comprometer legibilidade, mas é recomendável manter apêndice metodológico com *prompts* principais, versão do modelo, data de uso, dados fornecidos e procedimentos de validação. Essa prática permite reprodutibilidade parcial e auditoria.

O pesquisador deve conservar separadamente dados brutos, dados processados, scripts, tabelas, saídas de IA e versão final interpretada. A IA não deve ser o único local onde a análise existe. Ela é uma camada de mediação, não repositório metodológico. Em estudos de avaliação, essa separação é também uma exigência ética, pois permite contestação e correção.

A seguir, são apresentados *prompts* aplicados que operacionalizam esses princípios. Eles podem ser adaptados a diferentes temas, bases e unidades de análise. Sua formulação inclui instruções para explicitar incertezas, evitar extrapolações e distinguir indicadores apropriados de indicadores inadequados.

Prompt 10.9.1 Selecionar indicadores adequados à pergunta de pesquisa

Atue como especialista em Cientometria e métricas responsáveis.

A partir da pergunta de pesquisa [PERGUNTA], proponha um conjunto de indicadores adequados.

Para cada indicador, apresente: dimensão analítica medida, unidade de análise, dados necessários, base de dados possível, método de cálculo, visualização recomendada, limitações, riscos de interpretação e necessidade de validação humana.

Não sugira indicadores que não respondam diretamente à pergunta. Separe indicadores essenciais, complementares e inadequados.

Prompt 10.9.2 Explicar limitações do índice prompt a públicos não especialistas

Explique o índice *prompt* para um público não especialista, em português claro e tom institucional.

Inclua: definição simples, exemplo numérico, o que o indicador tenta combinar, cinco limitações importantes, situações em que seu uso é inadequado, alternativas complementares e uma advertência final sobre porque ele não deve ser usado sozinho para avaliar qualidade científica ou decidir promoção, financiamento ou contratação.

Prompt 10.9.3 Interpretar crescimento de publicações sem inferência causal indevida

Receba a série temporal de publicações abaixo e escreva uma interpretação cientométrica prudente.

Separe a resposta em: descrição objetiva dos dados; hipóteses possíveis para o crescimento ou declínio; dados adicionais necessários para testar cada hipótese; limitações da base e da estratégia de busca; e formulações que devem ser evitadas por extrapolarem os dados.

Não afirme que crescimento quantitativo prova qualidade, impacto ou relevância social.

Prompt 10.9.4 Escrever uma seção de resultados sem extrapolar os dados

A partir da tabela de indicadores abaixo, redija uma seção de resultados em tom formal científico.

Regras: descreva apenas o que os dados sustentam; diferencie produção, impacto, colaboração, internacionalização e atenção social; não use métricas como sinônimo de qualidade; informe limitações de fonte, período e normalização; destaque incertezas; e conclua com uma frase que indique quais interpretações exigem evidência adicional.

Prompt 10.9.5 Auditoria de afirmações métricas

Revise o texto abaixo e identifique afirmações métricas potencialmente problemáticas.

Para cada frase, classifique o problema: causalidade indevida, generalização excessiva, confusão entre impacto e qualidade, ausência de normalização, comparação injusta, omissão da fonte ou linguagem avaliativa forte demais.

Em seguida, proponha uma versão revisada, mais precisa e compatível com métricas responsáveis.

10.10 Matriz de decisão para escolha de indicadores

Uma regra prática para qualquer indicador é perguntar: qual fenômeno se pretende observar, qual rastro documental representa esse fenômeno, qual base registra esse rastro, qual cálculo o transforma em número e qual interpretação é permitida pelo resultado? Essa sequência impede que métricas sejam escolhidas apenas por disponibilidade. Ela também ajuda a IA a operar dentro de limites. Se o pesquisador informa que possui apenas título, ano, autores e afiliações, o modelo não deve sugerir análise de citações ou *altmetrics* sem indicar que os dados são insuficientes.

Outra regra consiste em distinguir indicadores descritivos, comparativos, avaliativos e prospectivos. Indicadores descritivos apresentam padrões de uma coleção. Indicadores comparativos situam unidades em relação a outras. Indicadores avaliativos entram em decisões de mérito ou desempenho. Indicadores prospectivos buscam antecipar tendências. Quanto mais avaliativa ou prospectiva for a finalidade, maior deve ser a exigência de normalização, validação e revisão por especialistas. A IA pode classificar indicadores segundo essas finalidades e alertar para riscos de uso inadequado.

Em relatórios institucionais, recomenda-se sempre combinar pelo menos três camadas: métricas de volume, métricas de influência e métricas de contexto. Volume sem influência pode indicar atividade pouco visível; influência sem contexto pode refletir poucos *outliers*; contexto sem métrica pode gerar narrativa impressionista. A triangulação não elimina incerteza, mas reduz conclusões unidimensionais. *Prompts* de síntese devem pedir que o modelo preserve tensões entre indicadores, em vez de forçar uma conclusão única.

A visualização também influencia a interpretação. Gráficos de linha sugerem evolução, mapas de rede sugerem relação, rankings sugerem hierarquia e mapas temáticos sugerem estrutura cognitiva. Quando a IA recomenda visualizações, deve justificar o tipo gráfico em função da pergunta. Um ranking pode ser inadequado quando diferenças são pequenas ou dados incertos; uma rede pode ser inadequada quando a desambiguação é frágil; uma média pode ser inadequada quando a distribuição é assimétrica. O *prompt* deve pedir alternativas e contraindicações.

Quadro 10.4 - Matriz de decisão por objetivo de análise

Objetivo	Indicadores recomendados	Indicadores complementares	Evitar
Descrever crescimento de um tema	Produção anual, taxa de crescimento, tipos documentais, palavras-chave.	Redes temáticas, periódicos centrais, autores recorrentes.	Concluir relevância social apenas por volume.
Avaliar recepção acadêmica	Citações normalizadas, percentis, top 10%, mediana.	Análise de contexto de citação, autocitações.	Comparar citações brutas entre áreas.
Mapear colaboração	Coautoria, países, instituições, redes e centralidade.	Posição autoral, financiamento, autoria correspondente.	Confundir coautoria com colaboração substantiva.
Analisar interdisciplinaridade	Diversidade de categorias, Rao-Stirling, áreas citadas e citantes.	Análise textual de resumos e palavras-chave.	Assumir que diversidade sempre significa integração profunda.

Observar circulação social	Menções em políticas, notícias, Mendeley, repositórios e *altmetrics*.	Análise qualitativa dos documentos que mencionam a pesquisa.	Usar atenção social como prova de qualidade científica.

Fonte: Elaborado pelo autor (2026).

A redação de resultados deve adotar linguagem proporcional à evidência. Expressões como "demonstra", "prova", "determina" e "garante" raramente são adequadas para indicadores cientométricos descritivos. Termos como "sugere", "indica", "é compatível com", "pode refletir" e "deve ser interpretado com cautela" são mais apropriados quando há inferência. A IA pode revisar a linguagem para calibrar força argumentativa. Essa revisão linguística é uma forma de controle epistemológico.

10.11 Aprofundamento metodológico sobre indicadores e prompts

A escolha de indicadores cientométricos pode ser descrita como um problema de alinhamento entre dimensão teórica, rastro empírico e finalidade interpretativa. A dimensão teórica corresponde ao fenômeno que se deseja compreender, como produtividade, influência acadêmica, cooperação, diversidade cognitiva, circulação internacional ou atenção pública. O rastro empírico corresponde ao vestígio documentado que permite observar parcialmente esse fenômeno, como registros bibliográficos, citações, afiliações, palavras-chave, referências, menções em notícias ou documentos de política. A finalidade interpretativa corresponde ao uso que será dado ao resultado: descrição, comparação, monitoramento, avaliação, planejamento estratégico ou prestação de contas. Quando esses três níveis não estão alinhados, a métrica passa a responder a uma pergunta diferente daquela que motivou o estudo.

A IA generativa pode contribuir justamente para explicitar esse alinhamento. Um *prompt* bem construído pode solicitar que o modelo preencha uma matriz com quatro colunas: pergunta de pesquisa, dimensão analítica, indicador recomendado e justificativa. Em seguida, pode pedir que o modelo identifique incoerências. Por exemplo, se a pergunta trata de impacto social de uma política pública, mas o indicador sugerido é apenas número de citações acadêmicas, o modelo deve apontar que há desalinhamento entre pergunta e métrica. Da mesma forma, se a pergunta trata de colaboração substantiva, mas os dados disponíveis são apenas coautorias, o modelo deve advertir que coautoria é *proxy*, e não equivalência perfeita.

Essa função de verificação é especialmente útil porque muitos estudos cientométricos começam pela disponibilidade de dados, não pela pergunta. O pesquisador possui uma planilha exportada de uma base, observa colunas de autores, citações e afiliações e tende a calcular tudo que o software permite. O resultado pode ser tecnicamente correto, mas analiticamente disperso. A engenharia de *prompts* ajuda a reorganizar o processo: primeiro a pergunta, depois a dimensão, depois o indicador, depois o cálculo e, somente então, a visualização. Essa ordem metodológica evita o fascínio por painéis ricos em gráficos, mas pobres em inferência.

Outro ponto central é que indicadores podem ser usados em diferentes escalas. No nível micro, analisam autores, artigos ou grupos; no nível meso, programas, departamentos, periódicos, instituições e áreas; no nível macro, países, regiões, sistemas científicos e campos globais. A escala altera o significado da métrica. O índice *prompt* de um autor não tem a mesma interpretação que o índice *prompt* de uma instituição; a colaboração internacional de um país pequeno não tem o mesmo significado que a de um país grande; a produção de um

periódico não pode ser comparada diretamente à de uma área. *Prompts* devem incluir a escala para impedir analogias indevidas.

A janela temporal é outro elemento decisivo. Indicadores de produção podem ser calculados em janelas curtas, mas indicadores de citação exigem tempo de maturação. Trabalhos recém-publicados tendem a receber poucas citações, mesmo quando são relevantes. Áreas com ciclos rápidos citam em poucos meses; áreas com ciclos longos podem levar anos para incorporar literatura. Portanto, o período de análise deve ser coerente com o fenômeno. A IA pode ajudar a perguntar se a janela temporal é adequada, se é necessário separar coortes e se indicadores de atenção rápida, como downloads ou menções, são mais apropriados para captar circulação inicial.

A normalização é uma das operações mais importantes e menos compreendidas. Normalizar não significa manipular dados, mas tornar comparações mais justas diante de diferenças conhecidas. Citações precisam ser normalizadas por área e ano; produção institucional pode exigir normalização por tamanho; colaboração pode exigir ajuste por cultura disciplinar; internacionalização pode exigir consideração do porte do país. Um *prompt* de auditoria pode perguntar: que diferença estrutural entre unidades pode distorcer a comparação? Há denominador adequado? A base cobre igualmente as unidades? A métrica foi ajustada ao tempo? Essa sequência transforma a IA em instrumento de crítica metodológica.

Contudo, normalização também tem limites. Classificações disciplinares podem ser inadequadas, especialmente em temas interdisciplinares. Denominadores institucionais podem ser inconsistentes. Dados de docentes, pesquisadores ativos ou orçamento podem não estar disponíveis. Percentis dependem do conjunto de referência. Assim, indicadores normalizados não eliminam julgamento; apenas deslocam a discussão para outro nível. A IA deve explicar a normalização como tentativa de reduzir viés, não como garantia de justiça absoluta.

Em relação aos indicadores de produção, há uma tensão entre produtividade e contribuição. Muitos documentos não significam necessariamente muitas contribuições relevantes; poucos documentos não significam baixa relevância. A qualidade do problema, a originalidade do método, a robustez dos dados, a abertura dos materiais, a capacidade de formar pessoas e a utilidade pública podem não aparecer na contagem. Por isso, indicadores de produção devem ser apresentados como medidas de atividade registrada. Um *prompt* de redação deve substituir expressões como "os autores mais produtivos são os mais importantes" por "os autores com maior número de documentos indexados no *corpus* foram". Essa diferença linguística é metodologicamente importante.

Nos indicadores de citação, a cautela deve ser ainda maior. A citação é um ato social e retórico situado em práticas disciplinares. Um artigo pode ser citado porque oferece método, dados, teoria, revisão, crítica, controvérsia ou conveniência. A contagem de citações agrega todos esses usos em um número. Isso não a torna inútil, mas limita sua semântica. A IA pode ajudar a redigir explicações que distingam impacto citacional de qualidade científica. Também pode sugerir análise de contexto de citação quando a pergunta exigir compreensão mais fina da influência intelectual.

O uso de indicadores de periódicos para avaliar artigos ou pessoas é uma prática criticada por movimentos de avaliação responsável. O fator de impacto descreve desempenho médio de um periódico em uma janela específica, não a qualidade de cada artigo nele publicado. Como distribuições de citações são assimétricas, poucos artigos podem responder por grande parte das citações de um periódico. Portanto, inferir qualidade individual a partir de métrica agregada de veículo é erro ecológico. Um *prompt* de auditoria deve identificar frases que atribuem valor a artigos exclusivamente pelo periódico em que foram publicados.

Nos indicadores de colaboração, também é necessário distinguir quantidade de relação e qualidade de relação. O número de coautores pode refletir cooperação genuína,

divisão de trabalho, consórcio multicêntrico, prática de laboratório ou convenção de autoria. Uma rede densa pode indicar campo integrado, mas também pode resultar de um grande projeto comum. A centralidade de uma instituição pode significar liderança ou simplesmente presença em muitos consórcios. Assim, indicadores de colaboração devem ser combinados com informações qualitativas sobre coordenação, financiamento, papéis e produtos.

A interdisciplinaridade exige especial atenção porque é frequentemente usada como valor normativo. Instituições desejam parecer interdisciplinares, editais incentivam interdisciplinaridade e pesquisadores mobilizam o termo para justificar agendas. Porém, medir interdisciplinaridade requer definição. Um estudo pode medir diversidade de referências, variedade de categorias, combinação de palavras-chave, coautoria entre departamentos ou difusão em múltiplas áreas. Esses indicadores não são equivalentes. A IA pode ajudar a formular a pergunta correta: deseja-se medir variedade cognitiva, integração teórica, circulação entre comunidades ou colaboração organizacional?

A internacionalização, por sua vez, é atravessada por assimetrias geopolíticas. Colaborações Norte-Sul, redes de excelência, consórcios globais e mobilidade acadêmica produzem benefícios, mas também podem reproduzir dependências. Métricas que apenas contam países parceiros podem ocultar quem define perguntas, quem controla dados, quem recebe financiamento e quem assina como autor correspondente. Um relatório responsável deve ir além da contagem e perguntar sobre reciprocidade, liderança e distribuição de benefícios. A IA pode ser usada para gerar perguntas críticas, não apenas para celebrar conexões internacionais.

Indicadores alternativos ampliam o horizonte da cientometria, mas trazem problemas próprios. Rastros digitais são instáveis, dependem de plataformas privadas, mudam com políticas de acesso e podem ser manipulados por campanhas. Além disso, diferentes eventos têm significados diferentes. Uma menção em diretriz clínica não equivale a um post em rede social; um salvamento em gerenciador de referências não equivale a citação em artigo; uma notícia jornalística não equivale a adoção em política pública. A IA pode classificar esses sinais por tipo de atenção e sugerir que não sejam somados sem justificativa.

A escrita de relatórios é onde muitos erros se materializam. Mesmo quando indicadores são calculados corretamente, a linguagem pode extrapolar. “O tema cresceu” é diferente de “o tema tornou-se prioritário”; “o artigo foi muito citado” é diferente de “o artigo é de alta qualidade”; “a instituição colaborou internacionalmente” é diferente de “a instituição liderou redes globais”. *Prompts* de revisão crítica devem procurar exatamente esses deslocamentos semânticos. A IA pode funcionar como revisor de prudência, identificando verbos fortes, adjetivos avaliativos e inferências não sustentadas.

A construção de narrativas científicas com indicadores deve seguir uma arquitetura simples. Primeiro, apresentar o *corpus* e suas limitações. Segundo, descrever produção. Terceiro, discutir impacto com normalização. Quarto, interpretar colaboração e circulação. Quinto, examinar sinais complementares, como interdisciplinaridade e atenção social. Sexto, sintetizar achados sem transformar métricas em julgamento absoluto. Essa ordem ajuda o leitor a compreender que indicadores são peças de um argumento, não conclusões automáticas.

Em termos de governança, relatórios assistidos por IA devem indicar quais tarefas foram apoiadas por modelos generativos. A seleção de indicadores, a criação de *prompts*, a revisão de linguagem e a elaboração de hipóteses são usos aceitáveis quando documentados. A geração de números sem fonte, a invenção de referências, a atribuição de causalidade e a classificação de pessoas por mérito são usos inadequados. A transparência sobre o papel da IA protege o pesquisador, o leitor e os sujeitos avaliados.

A educação cientométrica deve incluir alfabetização em IA. Saber calcular indicadores não basta; é necessário saber perguntar à IA, desconfiar de suas respostas, verificar referências, revisar interpretações e manter rastreabilidade. O pesquisador deve tratar a resposta do modelo como minuta, não como resultado. Essa postura é especialmente importante em métricas, porque a combinação entre números e linguagem fluente produz forte efeito de autoridade.

A elaboração de *prompts* para indicadores deve incorporar comandos negativos explícitos. Por exemplo: "não use citações como sinônimo de qualidade", "não compare áreas sem normalização", "não atribua causalidade", "não crie dados ausentes", "não recomende ranking se as diferenças forem pequenas", "não oculte limitações". Esses comandos reduzem, ainda que não eliminem, respostas problemáticas. Também sinalizam ao pesquisador quais riscos devem ser monitorados.

Em uma perspectiva prática, cada tabela de indicadores deveria ser acompanhada de uma legenda interpretativa e de uma legenda de limitação. A primeira explica o que o leitor pode observar; a segunda explica o que não pode concluir. A IA pode ajudar a produzir ambas. Por exemplo, em uma tabela de citações, a legenda interpretativa pode destacar documentos com maior recepção; a legenda de limitação pode informar que as citações não foram normalizadas ou que a janela é curta. Essa prática simples aumenta a qualidade comunicacional do relatório.

A relação entre indicadores e visualização também merece ser integrada aos *prompts*. Quando se solicita um gráfico, deve-se indicar a finalidade: comparar distribuição, mostrar evolução, mapear rede, evidenciar concentração ou demonstrar composição. A IA pode sugerir que rankings sejam substituídos por intervalos quando há incerteza, que médias sejam acompanhadas de medianas, que redes tenham limiares explícitos e que mapas incluam notas sobre algoritmos. Assim, visualização torna-se parte do raciocínio metodológico.

Por fim, a escolha de indicadores deve ser vista como decisão ética. Medir é selecionar o que conta; selecionar o que conta influencia comportamentos; influenciar comportamentos afeta a ciência. Uma instituição que valoriza apenas volume estimulará volume; uma agência que valoriza apenas citações estimulará estratégias de citação; um programa que valoriza apenas internacionalização pode negligenciar problemas locais. A IA não resolve esse dilema, mas pode torná-lo mais visível ao perguntar: que comportamentos este indicador pode incentivar? Quem pode ser favorecido? Quem pode ser prejudicado? Que dimensões ficam invisíveis?

10.12 Complementos operacionais para indicadores em relatórios assistidos por IA

Em estudos cientométricos aplicados, a seleção de indicadores deve ser acompanhada por um plano de análise. Esse plano precisa declarar quais indicadores serão calculados, por que foram escolhidos, que dados são necessários, que transformações serão realizadas, que softwares serão utilizados e quais critérios orientarão a interpretação. A ausência desse plano favorece a análise oportunista: calcula-se uma grande quantidade de métricas e selecionam-se apenas aquelas que produzem uma narrativa conveniente. Em uma abordagem assistida por IA, o plano de análise deve ser fornecido ao modelo antes da interpretação, para que a resposta seja orientada por decisões previamente estabelecidas, e não por improvisação retórica.

A IA pode ajudar a transformar um plano amplo em tarefas executáveis. Por exemplo, diante da instrução "avaliar impacto científico de um programa", o modelo pode decompor o objetivo em perguntas: impacto em relação a que campo? em qual período? com quais tipos documentais? a partir de que base? com citações brutas ou normalizadas? com ou sem autocitações? em comparação com quais unidades? Essa decomposição evita que a palavra

"impacto" seja tratada como categoria autoevidente. Na cientometria, impacto sempre precisa de objeto, público, janela temporal e fonte de observação.

A relação entre indicadores e dados ausentes também deve ser enfrentada. Em muitas bases, afiliações antigas são incompletas, identificadores de autores não são uniformes, referências podem estar ausentes, citações podem estar embargadas ou metadados de financiamento podem ser irregulares. Uma IA que recebe uma tabela incompleta pode tentar compensar a ausência com generalizações. Por isso, *prompts* devem incluir instruções como: "identifique campos ausentes", "não presuma valores", "marque indicadores inviáveis" e "proponha alternativas compatíveis com os dados disponíveis". Essa estratégia transforma lacuna de dados em informação metodológica, em vez de ocultá-la.

Indicadores de produção podem ser combinados com medidas de concentração. Não basta saber quantas publicações existem; muitas vezes é necessário saber se a produção está distribuída entre muitos atores ou concentrada em poucos. Curvas de Lorenz, coeficientes de Gini, percentuais de contribuição dos autores mais produtivos e distribuição por quartis podem revelar concentração. Uma área emergente dominada por poucos grupos pode ter dinâmica distinta de uma área distribuída em muitas instituições. A IA pode auxiliar a interpretar concentração, mas deve evitar inferir automaticamente hegemonia intelectual, pois concentração também pode decorrer de especialização, infraestrutura rara ou estratégia de busca.

Nos indicadores de citação, a análise de concentração é igualmente relevante. Um conjunto pode ter média elevada porque poucos artigos receberam muitas citações. Relatórios que apresentam apenas médias podem criar impressão de impacto generalizado. Por isso, é recomendável apresentar mediana, distribuição, proporção de documentos sem citação, top 10% e identificação de *outliers*. Um *prompt* de interpretação deve pedir que o modelo informe se o impacto parece distribuído ou concentrado. Essa distinção é importante para políticas científicas: impacto concentrado em poucos documentos demanda leitura diferente de impacto amplamente distribuído.

A análise de autocitação deve ser contextualizada em múltiplos níveis. Autocitação de autor pode indicar continuidade de linha de pesquisa; autocitação de periódico pode afetar métricas de periódico; autocitação institucional pode refletir grupos consolidados; autocitação nacional pode refletir comunidades locais. Remover todas as autocitações pode empobrecer a análise histórica de um campo; ignorá-las pode inflar impacto externo. A decisão deve ser relatada. Um *prompt* responsável pode solicitar duas versões interpretativas: uma com autocitações incluídas e outra com autocitações excluídas, destacando diferenças e implicações.

Indicadores de colaboração também podem se beneficiar de medidas de equilíbrio. Uma instituição pode colaborar internacionalmente, mas sempre com os mesmos parceiros; outra pode colaborar com muitos países, mas de forma esporádica. Indicadores como número de parceiros, distribuição de parcerias, repetição de vínculos e persistência temporal ajudam a diferenciar cooperação ocasional de rede consolidada. A IA pode organizar essas dimensões em categorias interpretativas: colaboração pontual, colaboração recorrente, dependência de parceiro dominante, diversificação de rede ou papel de ponte.

Para redes, a escolha de limiares é uma decisão crítica. Incluir todas as conexões pode gerar mapas visualmente densos e pouco interpretáveis; excluir conexões fracas pode apagar atores periféricos importantes. Não há limiar universal. O pesquisador deve justificar o critério, como número mínimo de documentos, força de ligação ou período. A IA pode ajudar a comparar efeitos de limiares diferentes, mas os resultados devem ser recalculados em software apropriado. O modelo pode interpretar cenários, não substituir análise de rede.

Indicadores de internacionalização também podem ser relacionados à mobilidade e à diáspora científica. Afiliações internacionais podem refletir colaboração entre instituições, mas também autores com múltiplas afiliações, pesquisadores expatriados ou programas de dupla vinculação. Em alguns contextos, uma publicação com afiliações de dois países não necessariamente resulta de colaboração entre equipes distintas. Essa nuance é importante em análises de políticas. *Prompts* de auditoria podem solicitar que a IA liste ambiguidades possíveis nas afiliações internacionais e sugira verificações complementares.

No caso da interdisciplinaridade, uma análise responsável deve verificar se a diversidade observada é substantiva ou apenas classificatória. Um artigo pode ser classificado em múltiplas áreas porque o periódico é amplo, não porque o conteúdo integra campos. Por outro lado, uma pesquisa realmente interdisciplinar pode aparecer em uma única categoria de base. A análise textual de resumos, palavras-chave e referências pode complementar classificações bibliográficas. A IA é especialmente útil para triagem semântica, mas seus rótulos devem ser testados por especialistas ou por amostras codificadas manualmente.

Os indicadores alternativos exigem atenção a eventos efêmeros. Um pico de atenção em rede social pode estar ligado a uma notícia, controvérsia ou campanha específica. Sem análise temporal, a contagem agregada de menções perde significado. Por isso, é recomendável observar curvas de atenção, fontes de menção, perfis de usuários e contexto. A IA pode resumir eventos associados a picos, mas apenas quando os textos ou metadados das menções são fornecidos. Caso contrário, qualquer explicação será especulativa.

Um aspecto pouco discutido é a relação entre indicadores e comunicação institucional. Universidades e centros de pesquisa frequentemente usam métricas em relatórios públicos, notícias e materiais de reputação. Nesse contexto, a tentação de converter indicadores em slogans é alta. A IA pode produzir textos persuasivos com facilidade, mas essa capacidade precisa ser contida por comandos de prudência. Um *prompt* para comunicação institucional deve exigir transparência, evitar rankings sem contexto, explicar bases e não transformar métricas em declarações absolutas de excelência.

Em relatórios para gestores, a visualização sintética é necessária, mas não deve apagar limitações. Painéis com semáforos, notas ou índices compostos podem facilitar leitura executiva, mas também podem ocultar incertezas. Uma abordagem equilibrada apresenta síntese executiva acompanhada de notas metodológicas curtas. A IA pode gerar as duas camadas: uma versão sintética para tomada de decisão e uma versão técnica para auditoria. Essa duplicidade melhora a comunicação sem sacrificar rigor.

Na formação de estudantes, indicadores devem ser ensinados por contraste. É útil apresentar exemplos de afirmações problemáticas e versões revisadas. Por exemplo, "o pesquisador A é melhor que o pesquisador B porque tem índice *prompt* maior" deve ser reescrito como "na base consultada, o pesquisador A apresenta índice *prompt* maior, mas a comparação requer considerar área, idade acadêmica, tipo de produção e contexto institucional". Esse exercício mostra que a cientometria não é apenas cálculo, mas linguagem responsável. A IA pode gerar exercícios, casos e rubricas de avaliação.

A aplicação de IA também permite criar rotinas de dupla checagem. Após gerar uma interpretação, o pesquisador pode submeter o próprio texto a um segundo *prompt*: "critique esta interpretação como se fosse um avaliador de métricas responsáveis". Em seguida, pode pedir: "aponte frases que dependem de dados não apresentados". Essa prática de revisão adversarial reduz o risco de complacência. O modelo não é infalível, mas a alternância entre geração e crítica melhora a qualidade do processo (LIU *et al.*, 2023; JI *et al.*, 2023).

Em contextos de avaliação institucional, recomenda-se que a análise de indicadores seja acompanhada por uma declaração de uso responsável. Essa declaração deve informar que métricas são parciais, que diferenças disciplinares foram consideradas, que os indicadores não substituem avaliação qualitativa, que dados e métodos são auditáveis e que conclusões avaliativas devem ser contextualizadas. A IA pode ajudar a redigir essa declaração, mas sua adoção deve refletir compromisso real da instituição.

O uso de métricas para indivíduos exige protocolo ainda mais rigoroso. Deve-se evitar rankings públicos simplificados, considerar trajetória, reconhecer diversidade de contribuições e permitir contestação de dados. Em muitos casos, indicadores individuais devem ser usados apenas como subsídio para reflexão, não como critério automático. A IA não deve gerar listas de "melhores pesquisadores" com base em métricas sem contexto. Um *prompt* ético deve recusar esse tipo de solicitação ou convertê-la em análise contextualizada.

Indicadores também podem ser usados para identificar lacunas e oportunidades, não apenas desempenho. Uma área com baixa colaboração internacional pode representar oportunidade estratégica; uma produção local com poucas citações internacionais pode indicar tema de relevância regional; uma rede fragmentada pode sugerir necessidade de articulação; baixa atenção social pode revelar falha de comunicação, não ausência de valor. A IA pode ajudar a transformar indicadores em perguntas de gestão, desde que evite recomendações normativas automáticas.

Por fim, o uso responsável de indicadores com IA requer cultura de revisão. Nenhum *prompt* substitui leitura especializada; nenhum indicador elimina contexto; nenhuma visualização encerra debate. A contribuição da IA está em ampliar a capacidade de formular, comparar, explicar e revisar, mas a legitimidade da análise depende de transparência, validação e responsabilidade humana. Em cientometria, a pergunta mais importante não é apenas "qual é o número?", mas "o que este número permite afirmar, o que ele oculta e que consequências pode produzir?".

10.13 Biblioteca comentada de usos de IA para indicadores

Um primeiro uso recorrente da IA consiste em preparar a conversa metodológica antes do cálculo. Nessa etapa, o pesquisador pode pedir que o modelo transforme um objetivo amplo em um conjunto de perguntas mensuráveis. Se o objetivo é "avaliar a produção de um programa de pós-graduação", a IA deve ajudar a decompor o problema em dimensões, como volume, perfil de publicação, colaboração, impacto relativo, internacionalização e aderência a linhas de pesquisa. Essa decomposição é útil porque impede que a avaliação seja reduzida a uma única medida. O resultado esperado não é uma resposta final, mas uma matriz preliminar que será revisada por especialistas.

Um segundo uso é a explicação pedagógica de indicadores. Em ambientes acadêmicos, muitas controvérsias decorrem do uso de termos que parecem transparentes, mas não são. Impacto, produtividade, excelência, colaboração e internacionalização são palavras carregadas de pressupostos. A IA pode produzir explicações adaptadas a diferentes públicos, como estudantes, gestores, avaliadores externos ou membros de colegiados. Entretanto, o *prompt* deve exigir que cada explicação inclua uma seção de limites. A educação métrica responsável depende tanto de ensinar o que o indicador mostra quanto de ensinar o que ele não mostra.

Um terceiro uso é a revisão de coerência entre indicador e conclusão. O pesquisador pode fornecer uma seção de resultados e pedir que a IA identifique conclusões não sustentadas. Essa prática é especialmente útil antes da submissão de relatórios institucionais ou artigos. A IA pode sinalizar frases que confundem produção com qualidade, citação com relevância social,

coautoria com colaboração efetiva, internacionalização com liderança global ou atenção social com impacto positivo. O texto final deve ser revisado por humanos, mas a etapa automatizada funciona como triagem crítica.

Um quarto uso é a construção de metadados metodológicos. Para cada indicador, a IA pode auxiliar a redigir campos padronizados: nome do indicador, definição, fórmula, fonte de dados, período, unidade de análise, método de contagem, tratamento de ausências, limitações e interpretação permitida. Esse tipo de documentação é valioso para reprodutibilidade. Em vez de apresentar apenas resultados, o relatório passa a mostrar como os resultados foram produzidos. A transparência reduz disputas e facilita atualização futura.

Um quinto uso é a comparação entre alternativas métricas. Diante de uma mesma pergunta, pode haver indicadores concorrentes. Para impacto, pode-se usar citações brutas, citações normalizadas, percentis, top 10%, média, mediana ou índice *prompt*. Para colaboração, pode-se usar média de autores, densidade de rede, centralidade, coautoria internacional ou diversidade de parceiros. A IA pode produzir uma comparação estruturada, indicando vantagens, limitações, dados necessários e risco de mau uso. Essa comparação apoia decisões metodológicas e evita escolha baseada apenas na familiaridade do indicador.

Um sexto uso é a produção de linguagem prudente. A IA generativa é especialmente eficaz na reformulação textual. Um *prompt* pode pedir que uma conclusão seja reescrita em três níveis de força: descrição estrita, interpretação moderada e hipótese especulativa. Essa técnica permite ao pesquisador escolher formulações compatíveis com a evidência. Também pode ser usada em treinamento, mostrando como pequenas mudanças linguísticas alteram o grau de afirmação. Em cientometria, a precisão verbal é parte do método.

Um sétimo uso é a elaboração de recomendações condicionais. Gestores frequentemente desejam recomendações a partir de indicadores. A IA pode ajudar, desde que as recomendações sejam condicionadas a evidências e não apresentadas como consequência automática. Por exemplo, baixa colaboração internacional pode sugerir criação de editais de mobilidade, mas também pode indicar missão local, área de pesquisa regional ou cobertura inadequada da base. Uma recomendação responsável dirá: "se a instituição deseja ampliar circulação internacional e se os dados forem confirmados, então...". Essa estrutura condicional reduz prescrições precipitadas.

Um oitavo uso é a criação de rubricas de avaliação de relatórios. A IA pode propor critérios para julgar se um relatório cientométrico é robusto: clareza da pergunta, adequação dos indicadores, transparência de dados, normalização, contextualização, limitação, ausência de extrapolação, qualidade das visualizações, documentação dos *prompts* e validação humana. Essa rubrica pode orientar estudantes e equipes técnicas. Ao mesmo tempo, deixa claro que o valor de um estudo cientométrico não está apenas no número de gráficos produzidos, mas na integridade do processo analítico.

Um nono uso é a identificação de indicadores ausentes. Ao receber um relatório que apresenta apenas produção e citações, a IA pode sugerir dimensões complementares, como colaboração, interdisciplinaridade, internacionalização, atenção social ou distribuição institucional. Isso não significa que todas as dimensões devam ser sempre incluídas. Significa que a ausência deve ser consciente. A IA pode perguntar: a dimensão omitida é irrelevante para a pergunta, indisponível nos dados ou apenas não foi considerada? Essa pergunta melhora a maturidade metodológica.

Um décimo uso é a preparação de apêndices reprodutíveis. *Prompts*, versões de bases, datas de coleta, filtros, *strings* de busca, campos exportados, critérios de exclusão, scripts e fórmulas podem ser organizados em apêndice metodológico. A IA pode formatar

esse material de maneira legível. Isso é particularmente importante em estudos de avaliação, nos quais decisões podem ser contestadas. A reprodutibilidade não exige que todo relatório seja excessivamente longo, mas exige que as condições de produção dos indicadores possam ser verificadas.

Esses usos mostram que a IA não precisa ocupar o lugar do analista. Seu papel mais produtivo é atuar como assistente de formulação, explicação, revisão, documentação e comunicação. Os cálculos continuam dependentes de dados, softwares e decisões metodológicas. A interpretação continua dependente de conhecimento de domínio. A responsabilidade continua humana. Quando esses limites são respeitados, a IA não enfraquece a cientometria; ela pode fortalecer sua transparência e sua capacidade pedagógica.

10.14 Estudo de caso ilustrativo: indicadores para IA em saúde pública

Considere um estudo hipotético sobre a produção científica em inteligência artificial aplicada à saúde pública entre 2014 e 2025. A pergunta inicial é ampla: como evoluiu o campo e quais dimensões indicam consolidação científica? Uma abordagem ingênua poderia contar publicações e concluir que o campo cresceu e se tornou relevante. Uma abordagem cientométrica responsável formularia dimensões distintas: produção anual, distribuição por áreas, impacto normalizado, colaboração internacional, interdisciplinaridade das referências, atenção em documentos de política e presença de temas emergentes.

Para produção, seriam usados número de documentos por ano, taxa de crescimento e distribuição por tipo documental. Para impacto, seriam usadas citações normalizadas por campo e ano, proporção de documentos altamente citados e mediana de citações. Para colaboração, seriam analisados coautoria internacional, rede de instituições e posição de países. Para interdisciplinaridade, seria examinada diversidade de áreas citadas e combinação de termos de saúde pública, epidemiologia, ciência de dados e ética. Para atenção social, seriam verificados documentos de política, diretrizes e menções em notícias especializadas.

A IA poderia apoiar o estudo em diferentes etapas. No planejamento, poderia sugerir indicadores adequados à pergunta e apontar limitações. Na análise, poderia transformar tabelas em descrições prudentes. Na interpretação, poderia gerar hipóteses concorrentes para o crescimento observado. Na redação, poderia revisar seções para remover extrapolações. Em todos os casos, os cálculos viriam de planilhas, scripts e bases verificadas; a IA não inventaria valores.

Um relatório responsável poderia afirmar que a produção indexada aumentou, que a colaboração internacional se tornou mais frequente, que alguns documentos alcançaram alto impacto relativo e que há sinais de circulação em políticas públicas. Não deveria afirmar, sem evidência adicional, que a área resolveu problemas de saúde, que a IA melhorou sistemas públicos ou que os artigos mais citados são necessariamente os mais úteis. Essa diferença entre achado bibliométrico e conclusão substantiva é uma das competências centrais da cientometria assistida por *prompts*.

10.15 Nota metodológica final sobre prudência interpretativa

A prudência interpretativa é uma competência central em qualquer capítulo dedicado a indicadores. Ela não significa enfraquecer resultados, mas ajustar a força da conclusão ao tipo de evidência disponível. Um indicador pode sustentar afirmações descritivas fortes quando os dados são completos e o cálculo é transparente; pode sustentar interpretações moderadas quando há coerência entre diferentes métricas; e pode sustentar apenas hipóteses quando há lacunas,

recorte estreito ou ausência de normalização. A IA deve ser orientada a reconhecer esses níveis, evitando que a fluência textual transforme evidência parcial em conclusão categórica.

Essa prudência também vale para recomendações. Um relatório pode sugerir que determinada instituição amplie colaboração, revise política de publicação ou fortaleça comunicação pública, mas tais recomendações devem ser condicionadas ao contexto, aos objetivos institucionais e a evidências complementares. Indicadores não substituem deliberação. Eles organizam sinais, reduzem opacidade e estimulam perguntas melhores. Em uma abordagem assistida por *prompts*, o valor da IA está em ajudar a formular essas perguntas, explicitar limites e tornar visível o caminho que liga dados, métricas e interpretação.

Síntese operacional

Em síntese, bons indicadores não encerram o debate científico; eles qualificam a conversa, tornam pressupostos visíveis e permitem que decisões sejam discutidas com base em evidências documentadas, limitações explícitas e responsabilidade interpretativa compartilhada.

Fecho metodológico

Essa é a base de uma prática cientométrica tecnicamente competente, socialmente responsável e metodologicamente reprodutível.

10.16 Checklist de qualidade para relatórios de indicadores

Antes de finalizar um relatório de indicadores cientométricos assistido por IA, recomenda-se verificar se a pergunta de pesquisa está explicitada, se a unidade de análise foi definida, se a base de dados foi identificada, se a data de coleta foi registrada, se o período de análise é coerente, se os tipos documentais foram delimitados, se nomes de autores e instituições foram limpos, se o método de contagem foi declarado, se autocitações foram tratadas ou justificadas, se citações foram normalizadas quando necessário, se comparações disciplinares foram contextualizadas e se a linguagem dos resultados evita extrapolações.

Também é necessário verificar se os *prompts* usados para selecionar indicadores, interpretar dados ou redigir resultados foram arquivados; se as saídas da IA foram revisadas por pessoa competente; se referências bibliográficas citadas pela IA foram verificadas; se os dados brutos permanecem preservados; se scripts ou planilhas estão disponíveis para auditoria; se decisões de exclusão ou inclusão foram registradas; se limitações estão presentes no texto; e se recomendações de política ou gestão são proporcionais às evidências.

Um relatório que não cumpre esses requisitos pode ser útil como exploração preliminar, mas não deve ser usado como base de decisão avaliativa. Em ambientes institucionais, a adoção de IA exige governança: definição de responsabilidades, controle de versões, transparência sobre modelos utilizados, proteção de dados sensíveis e possibilidade de contestação. Indicadores cientométricos não são apenas números; são dispositivos que podem afetar reputações, recursos e prioridades.

10.17 Considerações finais

Indicadores cientométricos são ferramentas indispensáveis para observar a ciência em escala, mas sua força depende da qualidade das perguntas e da prudência das interpretações. Indicadores de produção revelam atividade documentada; indicadores de citação indicam recepção acadêmica parcial; indicadores de colaboração descrevem

relações formais; indicadores de interdisciplinaridade captam diversidade e circulação entre campos; indicadores de internacionalização mostram articulação transnacional; indicadores alternativos registram atenção e usos fora da literatura acadêmica. Nenhum deles, isoladamente, mede a qualidade plena da ciência.

A inteligência artificial amplia a capacidade de explicar, comparar e narrar indicadores, mas também aumenta o risco de fluência sem fundamento. Por isso, *prompts* devem ser desenhados como instrumentos de controle metodológico. Eles precisam exigir contexto, fonte, unidade de análise, distinção entre dado e inferência, nível de incerteza, limitações e validação humana. O objetivo não é automatizar a avaliação científica, mas tornar a análise mais transparente, crítica e comunicável.

O princípio orientador deste capítulo é que métricas devem servir à compreensão da ciência, não governá-la de maneira cega. A IA pode contribuir para esse princípio quando ajuda a formular perguntas melhores, selecionar indicadores adequados, evitar extrapolações e comunicar resultados de forma responsável. Porém, a responsabilidade final permanece humana. Pesquisadores, avaliadores e gestores devem decidir quais dimensões importam, quais métricas são aceitáveis, quais limitações precisam ser reconhecidas e quais consequências podem decorrer do uso dos indicadores.

Em um livro dedicado à Cientometria com Inteligência Artificial, o capítulo sobre indicadores ocupa posição estratégica. Ele oferece a linguagem quantitativa que alimenta mapas, diagnósticos, relatórios e políticas. Ao mesmo tempo, lembra que toda medida é parcial, situada e carregada de pressupostos. A competência cientométrica do futuro não será apenas calcular indicadores com ferramentas automatizadas, mas saber perguntar, selecionar, auditar, interpretar e comunicar métricas em diálogo crítico com sistemas de IA.

Capítulo 11

Redes científicas e análise relacional

11.1 Introdução

A cientometria nasceu e se consolidou em estreita relação com a possibilidade de contar, comparar e interpretar registros da comunicação científica. Em suas formas mais conhecidas, ela mede produção, impacto, colaboração, circulação e visibilidade por meio de indicadores construídos a partir de artigos, livros, patentes, autores, periódicos, instituições, países, citações e outros metadados. Todavia, uma parte essencial da vida científica não aparece adequadamente quando se examinam apenas grandezas isoladas. Um autor não é apenas uma unidade produtiva; ele se insere em redes de colaboração, orientação, filiação, citação e reconhecimento. Um artigo não é apenas um documento publicado; ele carrega referências, dialoga com tradições, conecta frentes de pesquisa e pode servir de ponte entre comunidades. Uma palavra-chave não é apenas um descritor textual; ela participa de arranjos conceituais que expressam movimentos de especialização, hibridação e emergência temática. Por essa razão, a análise relacional ocupa posição central na cientometria contemporânea.

A expressão redes científicas designa, neste capítulo, representações formais de relações entre entidades do sistema científico. Essas entidades podem ser autores, documentos, referências, periódicos, instituições, países, palavras-chave, conceitos, áreas ou projetos. As relações podem assumir formas distintas: coautoria, cocitação, citação direta, acoplamento bibliográfico, coocorrência de termos, compartilhamento de afiliação, coparticipação em projetos, cofinanciamento ou coprodução internacional. Ao transformar tais relações em grafos, a cientometria passa a observar não apenas quanto uma unidade produz ou recebe citações, mas como ela se posiciona em estruturas coletivas de comunicação e colaboração. Essa mudança é epistemologicamente relevante porque desloca o olhar da unidade isolada para os padrões de interdependência.

A importância da análise relacional é histórica. O acoplamento bibliográfico foi proposto por Kessler (1963) como critério para identificar proximidade entre documentos a partir de referências compartilhadas. A cocitação foi formulada por Small (1973) como medida da frequência com que dois documentos são citados conjuntamente pela literatura posterior. A análise de co-palavras, associada a Callon, Courtial e Laville (1991), mostrou que a associação entre termos poderia revelar dinâmicas de organização cognitiva e trajetórias temáticas. A análise de redes de colaboração, desenvolvida em diálogo com a ciência das redes, permitiu estudar topologias de coautoria, componentes, centralidade e padrões disciplinares de colaboração. Essas contribuições mostram que a cientometria relacional não é apenas uma extensão técnica da bibliometria; ela é um modo de investigar a Ciência como sistema de conexões.

Com a expansão das bases bibliográficas, dos identificadores persistentes, das ferramentas de visualização e dos pacotes de análise, as redes científicas tornaram-se instrumentos comuns em revisões bibliométricas, estudos de mapeamento científico (ZUPIC; ČATER, 2015), diagnósticos institucionais e relatórios de política científica. Softwares como VOSviewer, CiteSpace, Gephi, Pajek e bibliometrix facilitaram (CHEN, 2006; VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017; BASTIAN; HEYMANN; JACOMY, 2009; COBO *et al.*, 2011) a construção de redes de coautoria, coocorrência, cocitação e acoplamento bibliográfico. Esses ambientes, entretanto, não eliminam a necessidade de interpretação

especializada. Um mapa pode sugerir proximidades e *clusters*, mas não explica por si mesmo porque uma comunidade se formou, se uma relação é substantiva, se uma ausência decorre de lacuna real ou de viés da base de dados, ou se um autor central é líder intelectual, intermediário metodológico ou apenas participante frequente de colaborações extensas.

É nesse ponto que a inteligência artificial, especialmente modelos de linguagem operados por *prompts*, passa a ter papel metodológico promissor e perigoso. *Prompts* podem ajudar a transformar resultados de redes em narrativas interpretativas, comparar tipos de rede, explicar métricas a públicos não especialistas, propor hipóteses para a formação de comunidades e auditar afirmações que extrapolam os dados. Ao mesmo tempo, modelos generativos podem inventar causalidades (JI *et al.*, 2023), atribuir significado excessivo a distâncias visuais, confundir associação com influência e transformar padrões estatísticos em histórias convincentes, porém infundadas. Portanto, a integração entre IA e redes científicas exige um princípio fundamental: a IA deve apoiar a interpretação, mas não substituir a evidência relacional, a validação bibliográfica e o julgamento do pesquisador.

O objetivo deste capítulo é oferecer uma base teórico-metodológica para o uso de redes científicas em estudos cientométricos assistidos por IA. O Capítulo apresenta os principais tipos de rede, suas unidades de análise, seus pressupostos, suas métricas e seus usos interpretativos. Também discute como formular *prompts* que obriguem o modelo a separar descrição estrutural, explicação possível, hipótese interpretativa, evidência necessária e risco de interpretação indevida. Ao final, propõe protocolos e modelos de *prompt* para que a análise relacional seja documentada, reprodutível e epistemicamente prudente. A ambição não é transformar a IA em interprete soberana dos mapas, mas ensina-la a operar como auxiliar critico em um processo conduzido por humanos.

11.2 Redes de coautoria

As redes de coautoria estão entre as formas mais intuitivas de análise relacional da ciência. Nelas, os nós representam autores, grupos, instituições, países ou unidades organizacionais, enquanto as arestas representam publicações em comum. A interpretação básica é aparentemente simples: se dois autores assinam conjuntamente um artigo, eles estão conectados por uma relação de colaboração registrada na literatura. Entretanto, essa simplicidade inicial esconde várias decisões metodológicas. É necessário definir se a rede será construída no nível de autores individuais, instituições, departamentos, programas de pós-graduação, países ou áreas. Também é necessário decidir se cada coautoria terá peso igual ou se será ponderada pelo número de artigos compartilhados, pelo número de autores por artigo, pelo período temporal, pelo tipo documental ou pelo impacto das publicações.

O uso de redes de coautoria permite observar estruturas de colaboração que indicadores simples não revelam. Uma área pode apresentar crescimento expressivo de publicações, mas manter baixa conectividade entre grupos. Outra pode ter produção mais modesta, mas alta densidade de colaboração internacional. Um programa institucional pode publicar muito, mas estar organizado em ilhas pouco conectadas. Um autor pode não ser o mais produtivo, mas atuar como ponte entre comunidades. Esses padrões são relevantes para a sociologia da ciência, para a gestão de pesquisa e para políticas de colaboração. A coautoria, contudo, e um indicador imperfeito de colaboração. Nem toda colaboração resulta em coautoria, e nem toda coautoria expressa colaboração intelectual equivalente. Em alguns campos, grandes consórcios distribuem autoria por centenas de participantes; em outros, a autoria e restrita e hierarquizada. Por isso, a rede de coautoria deve ser interpretada no contexto disciplinar.

A literatura de ciência das redes mostrou que redes de colaboração científica frequentemente apresentam componentes gigantes, alto coeficiente de agrupamento e distribuições assimétricas de grau. Newman (2001, 2004) examinou redes de colaboração em diferentes áreas e mostrou que a estrutura de coautoria varia de acordo com práticas disciplinares e bases de dados utilizadas. Esses achados são importantes para a cientometria porque impedem a leitura ingênua de centralidade como sinônimo automático de excelência. Um autor central em uma área experimental pode estar associado a grandes consórcios; em uma área teórica, a centralidade pode indicar capacidade de articulação intelectual; em uma área emergente, pode revelar poucos atores que concentram conexões iniciais. A mesma métrica assume significados diferentes conforme o sistema de autoria e comunicação.

Em estudos assistidos por IA, as redes de coautoria podem ser exploradas por *prompts* que organizam a interpretação em camadas. Primeiro, o modelo pode descrever a estrutura: número de nós, número de arestas, densidade, componentes, autores ou instituições mais conectados, existência de subgrupos. Segundo, pode propor explicações possíveis: especialização disciplinar, concentração institucional, programas de financiamento, políticas de internacionalização, consórcios temáticos. Terceiro, deve explicitar que essas explicações são hipóteses e indicar evidências adicionais necessárias: leitura de artigos, identificação de projetos, verificação de editais, análise de afiliações e trajetórias históricas. Esse formato evita que a IA transforme a rede em narrativa causal sem base empírica.

Um problema comum na interpretação de redes de coautoria é a invisibilidade das assimetrias de contribuição. A rede mostra que autores publicaram juntos, mas não mostra automaticamente quem concebeu o estudo, quem coletou dados, quem analisou resultados, quem financiou a pesquisa ou quem apenas integrou uma equipe ampla. A taxonomia de contribuições, quando disponível, pode enriquecer a análise, mas ainda e irregular em muitas bases. Do mesmo modo, a coautoria não captura relações informais, revisão por pares, orientação não publicada, influência teórica sem citação ou colaboração técnica não reconhecida. Portanto, a IA deve ser instruída a tratar coautoria como evidência documental de copublicação e não como prova completa de colaboração substantiva.

As redes de coautoria também são importantes para analisar desigualdades e periferias científicas. Em escala geográfica, elas permitem observar dependência de centros dominantes, concentração de parcerias em poucos países, ausência de colaboração Sul-Sul ou baixa conexão entre instituições de uma mesma região. Em escala institucional, podem revelar separação entre departamentos, baixa interação entre campi ou dependência de poucos pesquisadores articuladores. Em escala individual, podem indicar trajetórias de inserção, isolamento relativo ou papel de intermediação. Contudo, a interpretação desses padrões exige cautela: baixa conectividade pode significar isolamento, mas também especialização local, cultura disciplinar de autoria individual, barreiras linguísticas, opções metodológicas ou limitações da base de dados. A pergunta cientométrica deve orientar a leitura da rede.

Quadro 11.1 - Análise de redes de coautoria

O que a rede pode mostrar

Aspecto	O que se observa
Estrutura de colaboração	Densidade, conectividade e componentes.
Atores centrais e influentes	Grau, força e centralidades (grau, intermediação, autovetor).
Comunidades e subgrupos	*Clusters*, modularidade, temas ou áreas articuladas.
Pontes e intermediários	Atores que conectam comunidades diferentes.

Aspecto	O que se observa
Evolução temporal	Formação e dissolução de grupos, entrada de novos atores.
Padrões geográficos	Colaboração internacional, cooperação Sul-Sul, dependência de centros.
A rede revela padrões relacionais, não qualidade intrínseca da produção nem contribuição causal.	

O que a rede não mostra

Limite	O que permanece invisível
Natureza da contribuição	Não indica quem concebeu, executou, financiou ou liderou.
Colaborações não formalizadas	Relações informais, revisão, orientação não publicada e consultorias.
Intensidade real do trabalho	Pode haver coautoria nominal ou participação mínima.
Hierarquias internas	A posição na autoria (primeiro, último) não é capturada na rede simples.
Qualidade e impacto	Atores igualmente centrais podem ter impactos muito diferentes.
Interprete a coautoria como evidência documental de copublicação, não como prova de colaboração substancial.	

Métricas de rede mais comuns

Métrica	O que indica
Grau (degree)	Número de conexões diretas de um nó.
Força (strength)	Soma dos pesos das conexões, em redes ponderadas.
Centralidade de intermediação	Capacidade de atuar como ponte entre comunidades.
Centralidade de autovetor	Conexões com nós que também são bem conectados.
Densidade	Proporção de conexões existentes em relação ao total possível.
Coeficiente de agrupamento	Tendência de os vizinhos de um nó estarem conectados entre si.
Modularidade	Grau de divisão da rede em comunidades.
Componentes conexos	Sub-redes não conectadas entre si.
Escolha métricas alinhadas à pergunta e ao nível da análise.	

Interpretação responsável: perguntas-guia

- Quem são os atores mais conectados e por que podem ser centrais?
- A rede é integrada ou fragmentada?
- Existem comunidades reconhecíveis ou disciplinas distintas?
- Quais atores conectam comunidades diferentes?
- Como a colaboração evoluiu ao longo do tempo?
- A colaboração internacional é diversificada e recíproca?
- A rede reproduz relações de centro e periferia, dependência ou concentração?
- Que fatores contextuais podem explicar os padrões observados?

Fonte: Elaborado pelo autor (2026).

Ao usar *prompts* para interpretar coautoria, recomenda-se fornecer ao modelo dados estruturados, como tabela de métricas, lista de *clusters*, descrição do período, critérios de coleta e limitações conhecidas. O *prompt* deve proibir conclusões sobre qualidade, liderança ou causalidade sem evidência adicional. Também deve exigir que o modelo diferencie padrões observados de interpretações possíveis. Um bom resultado não é aquele que produz a narrativa mais elegante, mas aquele que ajuda o pesquisador a reconhecer o que a rede efetivamente mostra, o que ela sugere e o que permanece desconhecido. Essa distinção é decisiva para transformar a IA em instrumento de análise cientométrica responsável.

11.3 Redes de cocitação

As redes de cocitação representam relações estabelecidas pela literatura posterior. Dois documentos, autores ou periódicos são cocitados quando aparecem juntos nas listas de referências de documentos citantes. A interpretação clássica é que a cocitação indica proximidade intelectual percebida pela comunidade citante. Se dois textos são frequentemente citados em conjunto, é plausível que sejam vistos como complementares, concorrentes, fundacionais ou pertencentes a uma mesma tradição. Small (1973) formulou a cocitação como uma medida do relacionamento entre documentos e abriu caminho para o estudo da estrutura de especialidades científicas por meio de *clusters* de documentos cocitados. Essa abordagem se tornou uma das bases do mapeamento de domínios do conhecimento.

Figura 11.1 - Cocitação X coautoria

Fonte: Imagem gerada por IA, ChatGPT (2026).

A principal diferença entre cocitação e coautoria é que a cocitação não depende da colaboração direta entre os autores. Ela depende do uso que a comunidade faz das referências. Por isso, redes de cocitação são particularmente úteis para identificar estruturas intelectuais, bases teóricas e documentos clássicos de um campo. Em um estudo sobre inteligência artificial na educação, por exemplo, uma rede de cocitação pode revelar conjuntos de textos sobre aprendizagem adaptativa, ética algorítmica, analítica da aprendizagem, processamento de linguagem natural e políticas educacionais. Esses *clusters* não resultam necessariamente de autores colaborando entre si, mas de leitores e pesquisadores que os mobilizam conjuntamente.

A temporalidade é central na cocitação. Um documento precisa ser publicado, lido, citado e cocitado para ganhar posição na rede. Assim, redes de cocitação tendem a favorecer literatura mais consolidada, especialmente quando a janela temporal inclui muitos anos. Documentos recentes podem ser importantes para frentes emergentes, mas ainda não acumularam cocitações suficientes. Essa característica faz da cocitação uma técnica poderosa para mapear bases intelectuais, mas menos sensível para detectar novidades muito recentes. A interpretação com IA deve, portanto, reconhecer que a ausência de um documento em uma rede de cocitação não significa necessariamente irrelevância; pode significar novidade, baixa exposição, viés linguístico, cobertura limitada da base ou período insuficiente de maturação citacional.

As redes de cocitação podem ser construídas no nível de documentos, autores ou periódicos. A cocitação de documentos é adequada para identificar artigos, livros ou Capítulos que estruturam uma frente intelectual. A cocitação de autores permite observar escolas, tradições ou comunidades epistêmicas associadas a nomes recorrentes. A cocitação de periódicos revela proximidades entre campos e canais de comunicação. Cada nível, contudo, implica perdas e agregações. A cocitação de autores pode reunir trabalhos muito distintos de um mesmo pesquisador; a cocitação de periódicos pode mascarar diversidade interna; a cocitação de documentos pode fragmentar a visão geral quando há muitos registros. A IA pode auxiliar comparando níveis de análise e sugerindo qual deles responde melhor a uma pergunta, mas a escolha deve ser metodologicamente justificada.

Na visualização, *clusters* de cocitação costumam ser interpretados como especialidades, correntes teóricas ou bases intelectuais. Essa interpretação é útil, mas não automática. Um *cluster* pode reunir documentos por afinidade teórica, por uso metodológico comum, por controvérsia, por revisão histórica ou por tradição de citação. A IA pode ajudar a rotular *clusters*, analisando títulos, resumos e termos frequentes, mas deve ser impedida de inventar nomes excessivamente precisos quando os dados não sustentam essa rotulagem. Um *prompt* robusto deve exigir que o modelo apresente pelo menos três possibilidades de rótulo, justifique cada uma com evidências dos documentos, indique ambiguidades e proponha verificação por leitura de uma amostra dos textos centrais.

A cocitação também pode ser combinada com análise de *burst*, trajetórias temporais e detecção de pontos de virada. Ferramentas como CiteSpace exploram (CHEN, 2006) a visualização progressiva de domínios do conhecimento, enfatizando documentos com alta centralidade de intermediação e referências com crescimento abrupto de citação. Essas técnicas são valiosas para reconstruir transformações de campos, mas também podem induzir leituras espetaculares. Um documento com *burst* pode indicar emergência de interesse, mas também modismo, controvérsia, mudança terminológica ou efeito de revisões muito citadas. Um documento com alta intermediação pode ser ponte intelectual, mas também resultado de agregação metodológica. A IA deve ser orientada a apresentar interpretações alternativas e evidências necessárias para confirmar cada uma.

Uma rede de cocitação, portanto, não é uma fotografia neutra da verdade intelectual de uma área. Ela é uma representação da memória citacional registrada em uma base, sob um período e um conjunto de regras de seleção. Ao usar IA para interpretar essa rede, o pesquisador deve fornecer informações sobre a base utilizada, critérios de inclusão, tipo documental, idioma, janela temporal, normalização de citações e limiar de ocorrência. Sem esses metadados, a IA tende a preencher lacunas com generalidades. Com eles, pode auxiliar na construção de uma leitura mais disciplinada, separando bases intelectuais consolidadas, frentes de pesquisa recentes e zonas de incerteza interpretativa.

11.4 Redes de acoplamento bibliográfico

O acoplamento bibliográfico ocorre quando dois documentos compartilham referências. Diferentemente da cocitação, que depende de documentos futuros que citam conjuntamente dois textos, o acoplamento pode ser calculado logo após a publicação, desde que as listas de referências estejam disponíveis. Essa propriedade torna o acoplamento bibliográfico particularmente útil para mapear frentes de pesquisa contemporâneas. Se dois artigos recentes citam muitas das mesmas referências, é plausível que partilhem uma base intelectual comum, ainda que ainda não tenham sido citados pela literatura posterior. Kessler (1963) apresentou essa ideia como critério automático de inter-relação entre documentos, e a técnica tornou-se um instrumento central do mapeamento científico.

Figura 11.2 - Acoplamento bibliográfico X cocitação

Fonte: Imagem gerada por IA, ChatGPT (2026).

Em termos interpretativos, o acoplamento bibliográfico evidencia proximidade retrospectiva: documentos se aproximam porque olham para referências semelhantes. A cocitação, ao contrário, evidencia proximidade prospectiva ou comunitária: documentos se aproximam porque são utilizados conjuntamente por citantes posteriores. Essa diferença é fundamental para a escolha metodológica. Quando o objetivo é identificar bases clássicas de um campo, a cocitação costuma ser adequada. Quando o objetivo é mapear a estrutura atual de uma frente de pesquisa, o acoplamento bibliográfico pode ser mais sensível. Boyack e Klavans (2010) compararam abordagens baseadas em cocitação, acoplamento e citação direta para representar frentes de pesquisa, mostrando que a escolha do método afeta significativamente o mapa produzido.

O acoplamento pode ser calculado entre documentos, autores, periódicos, instituições ou países, mas seu significado muda conforme a agregação. No nível de documentos, ele indica compartilhamento de referências entre artigos específicos. No nível de autores, pode

indicar que pesquisadores recorrem a bases bibliográficas semelhantes. No nível de periódicos, pode indicar proximidade de audiência, tema ou tradição de citação. No nível institucional, pode apontar convergência de agendas. Entretanto, quanto maior a agregação, maior o risco de apagar diferenças internas. Duas instituições podem parecer acopladas porque grupos pequenos compartilham referências, mesmo que a maior parte de suas pesquisas seja distinta. O uso de IA deve incluir perguntas sobre granularidade e validade da agregação.

A intensidade do acoplamento depende do número de referências compartilhadas, mas a comparação bruta pode ser enganosa. Artigos com listas de referências muito longas tendem a compartilhar mais itens por simples probabilidade. Por isso, métodos de normalização, como associação, cosseno, Jaccard e *fractional counting*, podem ser empregados para reduzir vieses. A escolha do método afeta a topologia da rede, o peso das arestas e a formação de *clusters*. Um *prompt* para apoio metodológico deve solicitar que a IA compare alternativas de normalização, explicite suas consequências e recomende uma abordagem coerente com o objetivo do estudo. Não se deve pedir apenas que a IA 'interprete o mapa'; deve-se pedir que ela considere como o mapa foi construído.

O acoplamento bibliográfico é especialmente relevante em revisões de literatura e estudos de campos emergentes. Em áreas recentes, como aplicações de IA generativa em educação, saúde ou direito, os artigos ainda podem ter poucas citações, mas suas listas de referências revelam comunidades de fundamentação. Um *cluster* pode se organizar em torno de aprendizagem de máquina, outro em torno de ética, outro em torno de avaliação educacional, outro em torno de regulação. A IA pode ajudar a ler títulos e resumos de artigos de cada *cluster*, sugerindo rótulos temáticos. Contudo, o modelo deve ser instruído a não inferir que artigos acoplados chegaram a conclusões semelhantes; eles podem compartilhar literatura de base e ainda assim divergir em resultados, método ou posição teórica.

A interpretação do acoplamento também requer atenção a práticas de citações. Campos com forte tradição de revisão podem gerar acoplamentos amplos em torno de poucos textos canônicos. Áreas metodológicas podem aproximar estudos de objetos empíricos distintos porque usam as mesmas ferramentas estatísticas ou computacionais. Trabalhos interdisciplinares podem aparecer como pontes por combinarem referências de vários campos. Em cada caso, a proximidade bibliográfica precisa ser analisada junto com conteúdo textual, unidade de análise e contexto disciplinar. A IA pode ser útil para sugerir hipóteses sobre a natureza do acoplamento (teórico, metodológico, empírico ou normativo), mas essas hipóteses devem ser testadas por leitura e por indicadores adicionais.

Um estudo cientométrico rigoroso pode comparar redes de acoplamento, cocitação e coocorrência de palavras-chave para obter uma visão mais robusta. Se os *clusters* de acoplamento coincidem parcialmente com os *clusters* de co-palavras, há indícios de alinhamento entre base bibliográfica e vocabulário temático. Se divergem, pode haver comunidades que compartilham referências, mas usam terminologias diferentes, ou comunidades que usam termos semelhantes com bases teóricas distintas. A IA pode apoiar essa triangulação, organizando convergências e divergências em uma tabela interpretativa. Ainda assim, deve declarar claramente o que é evidência relacional, o que é inferência plausível e o que exige verificação externa.

11.5 Redes de coocorrência de palavras-chave

As redes de coocorrência de palavras-chave, também chamadas de redes de co-palavras, representam a associação entre termos que aparecem conjuntamente em documentos. Os nós podem ser palavras-chave de autores, descritores controlados, termos

extraídos de títulos e resumos ou conceitos normalizados por um processo de limpeza. As arestas indicam coocorrência em um mesmo documento, em um mesmo conjunto de documentos, em uma janela temporal ou em uma unidade agregada. A premissa é que termos que aparecem juntos com frequência expressam proximidade temática, articulação conceitual ou convergência de problemas de pesquisa. Desde os trabalhos de Callon, Courtial e Laville (1991), a análise de co-palavras tornou-se uma técnica importante para descrever redes de interação entre pesquisa básica e tecnológica, acompanhar ciclos de vida de temas e mapear a dinâmica cognitiva de campos científicos.

Quadro 11.2 - Técnicas relacionais para mapeamento científico

Técnica	Relação capturada	Quando usar	Força principal	Limitação principal
Coocorrência de palavras-chave	Termos que aparecem juntos no mesmo documento	Estrutura temática e vocabulário	Sensível a temas e a termos emergentes	Termos amplos e viés de indexação
Cocitação	Documentos citados juntos nas referências de citantes futuros	Bases intelectuais e tradições	Perspectiva retrospectiva e comunitária	Menos sensível a novidades recentes
Acoplamento bibliográfico	Documentos que compartilham referências	Frentes atuais e pesquisas recentes	Detecta a estrutura contemporânea	Sensível a listas longas e a práticas de citação
A triangulação das três técnicas fornece uma visão mais robusta do campo.				

Fonte: Elaborado pelo autor (2026).

A rede de coocorrência de palavras-chave é particularmente sensível ao pré-processamento. Variantes ortográficas, singular e plural, siglas, sinônimos, traduções, termos amplos, erros de indexação e diferenças entre palavras-chave de autores e descritores controlados podem alterar profundamente o mapa. Por exemplo, *'artificial intelligence'*, *'AI'*, *'machine intelligence'* e 'inteligência artificial' podem ser tratados como termos separados ou agrupados, dependendo dos critérios do estudo. Da mesma forma, *'bibliometrics'*, *'scientometrics'* e *'science mapping'* podem ser sinônimos parciais em alguns contextos e conceitos distintos em outros. A IA pode auxiliar na normalização terminológica, mas também pode impor agrupamentos indevidos. Por isso, *prompts* devem exigir justificativa semântica, grau de confiança e marcação de casos ambíguos para revisão humana.

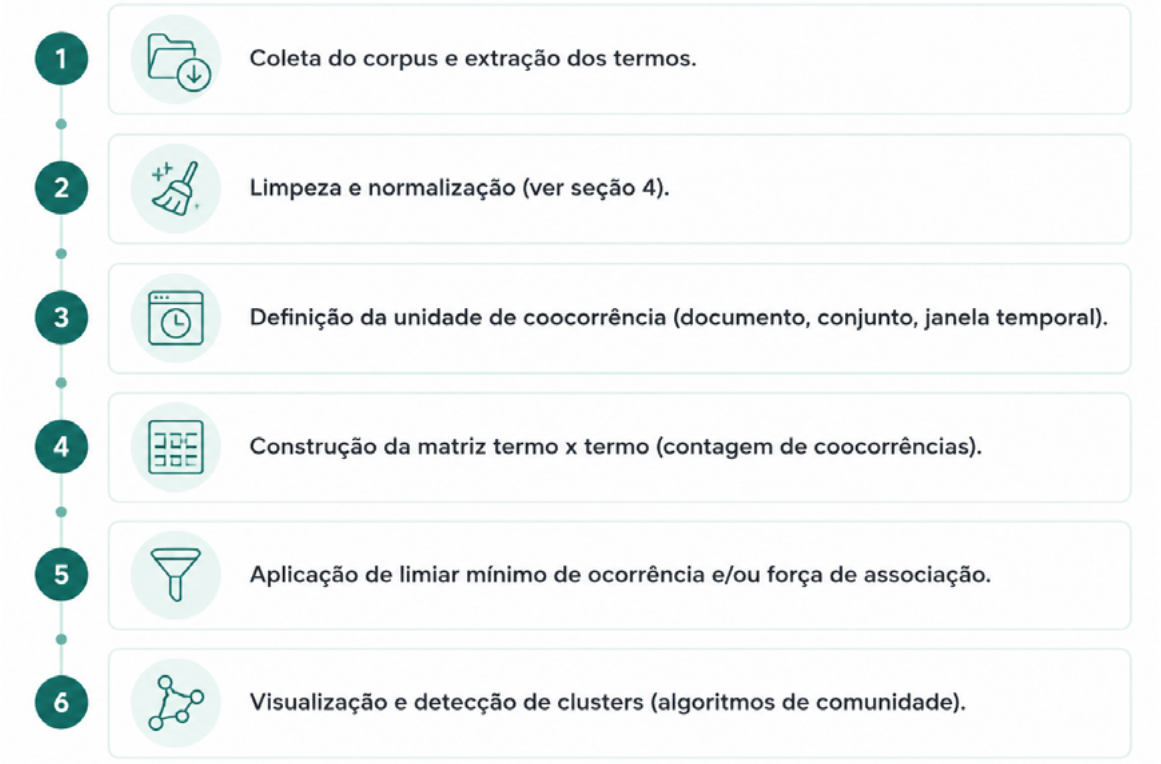
As redes de co-palavras podem ser usadas para identificar temas centrais, periféricos, emergentes e interdisciplinares. Termos com alto grau podem indicar tópicos amplamente conectados; termos em pontes podem revelar articulações entre subcampos; *clusters* podem sugerir famílias temáticas; mudanças temporais podem evidenciar transições conceituais. Contudo, a centralidade de uma palavra não significa, por si só, importância teórica. Termos muito genéricos tendem a conectar muitos documentos sem acrescentar especificidade. Palavras metodológicas podem dominar mapas sem representar objetos de pesquisa. Termos emergentes podem ter baixa frequência mas alta relevância prospectiva. Assim, a interpretação precisa combinar métricas, conhecimento de domínio e leitura qualitativa.

Uma contribuição importante da IA generativa consiste em apoiar a rotulagem de *clusters* de palavras-chave. A partir de uma lista de termos, seus pesos e exemplos

de documentos, o modelo pode sugerir rótulos interpretativos, indicar termos nucleares, diferenciar subtemas e apontar ambiguidades. Entretanto, a rotulagem automática é uma das etapas mais vulneráveis a alucinação conceitual. Modelos de linguagem tendem a produzir nomes elegantes e plausíveis, mesmo quando os termos não sustentam tal interpretação. Um *prompt* adequado deve pedir que o modelo gere rótulos alternativos, explicite quais termos justificam cada rótulo, indique termos que não se encaixam, proponha verificação por amostra e evite categorias excessivamente abrangentes. O rótulo final deve ser decisão do pesquisador, não produto automático da IA.

As palavras-chave também são marcadores sociais de visibilidade. Autores escolhem termos para posicionar seus trabalhos em comunidades, dialogar com bases de indexação e aumentar recuperabilidade. Em alguns campos, palavras-chave refletem vocabulário disciplinar estabilizado; em outros, podem seguir modas, exigências editoriais ou estratégias de busca. A coocorrência, portanto, expressa tanto estrutura cognitiva quanto comportamento comunicacional. A IA pode ajudar a distinguir termos conceituais, metodológicos, empíricos, geográficos e normativos, mas essa classificação deve ser documentada. Em uma análise de temas emergentes, por exemplo, é importante separar termos que indicam tecnologias, problemas sociais, teorias, métodos, populações estudadas e contextos de aplicação.

Figura 11.3 - Como a rede é construída



Fonte: Imagem gerada por IA, ChatGPT (2026).

Em redes de co-palavras, a escolha do *corpus* define a fronteira do mundo observado. Uma busca muito estreita pode produzir *clusters* artificiais e reforçar termos esperados. Uma busca muito ampla pode gerar mapa disperso, com conexões fracas e ruído semântico. Bases diferentes também variam na qualidade das palavras-chave. Algumas indexam descritores controlados; outras dependem dos autores; outras extraem termos automaticamente. Antes de pedir a IA que interprete uma rede, é preciso informar a origem dos termos, os critérios de

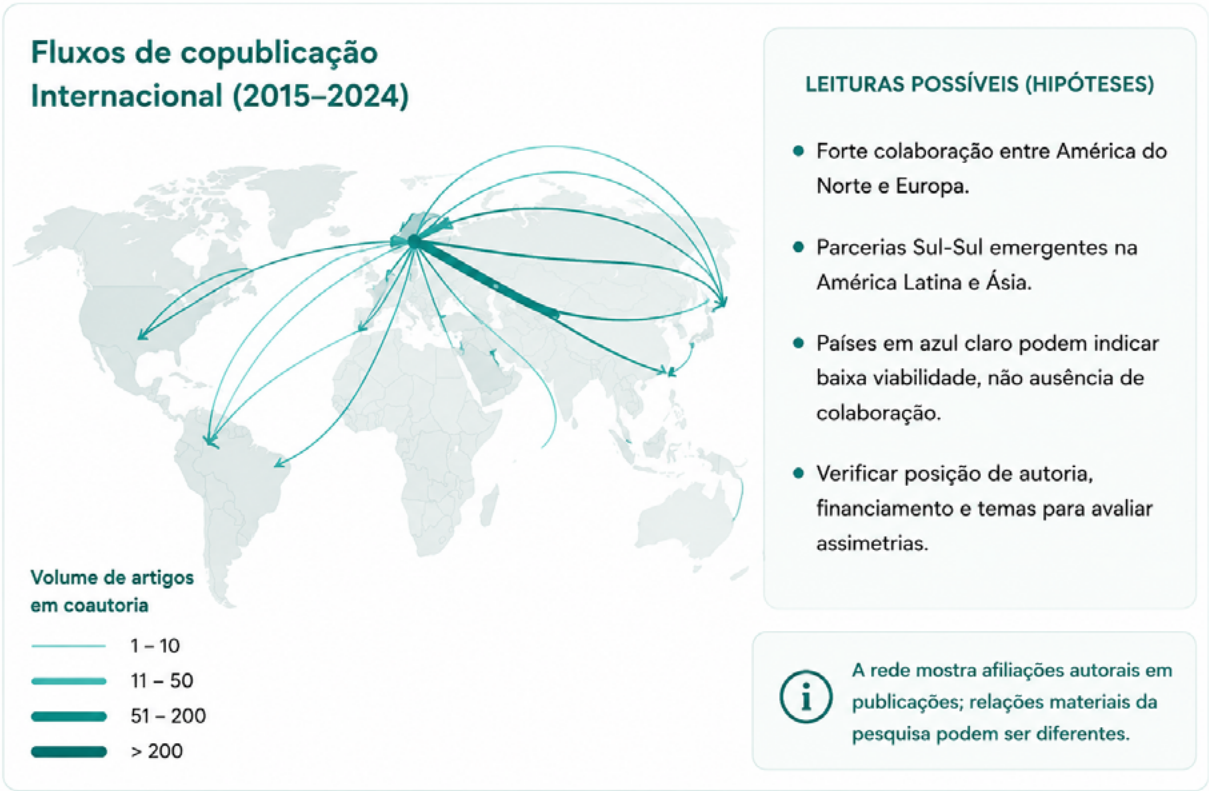
normalização, o limiar mínimo de ocorrência, o período e o idioma. Sem isso, o modelo pode atribuir significado a artefatos técnicos do processamento.

A rede de coocorrência de palavras-chave é uma ponte entre análise quantitativa e análise textual. Ela pode orientar amostragem qualitativa, selecionar documentos representativos de *clusters*, apoiar revisões de escopo e identificar lacunas terminológicas. Em estudos assistidos por IA, recomenda-se combinar *prompts* de classificação com *prompts* de auditoria. O primeiro ajuda a agrupar termos; o segundo questiona os agrupamentos. O *prompt* de auditoria deve perguntar: quais termos foram agrupados de modo discutível? Quais sinônimos podem estar faltando? Quais termos são amplos demais? Quais *clusters* podem refletir viés da estratégia de busca? Quais conclusões não podem ser tiradas apenas da coocorrência? Esse procedimento transforma a IA em instrumento de controle crítico e não apenas de geração de narrativas.

11.6 Redes institucionais e geográficas

As redes institucionais e geográficas ampliam a análise relacional da ciência para além dos autores e documentos, permitindo examinar como universidades, centros de pesquisa, hospitais, empresas, agências governamentais, regiões e países se conectam por meio de publicações, projetos, citações, afiliações e colaborações. Em uma rede institucional de coautoria, os nós podem representar instituições e as arestas indicam artigos assinados por pesquisadores afiliados a ambas. Em uma rede geográfica, os nós podem representar países, estados ou cidades, e as arestas expressam copublicação internacional ou inter-regional. Essas redes são centrais para estudos de internacionalização, diplomacia científica, desigualdade epistêmica, dependência periferia-centro, redes de excelência e políticas de colaboração.

Figura 11.4 - Exemplo ilustrativo de uma rede geográfica



Fonte: Imagem gerada por IA, ChtGPT (2026).

A construção de redes institucionais exige cuidado com a desambiguação de afiliações. Instituições podem aparecer com variantes linguísticas, siglas, nomes antigos, unidades internas, hospitais associados, fundações mantenedoras ou campus diferentes. A Universidade de São Paulo, por exemplo, pode aparecer como USP, Univ São Paulo, University of São Paulo, Faculdade de Medicina da USP ou Hospital das Clínicas. Sem normalização, a rede fragmenta uma instituição em vários nós, reduzindo artificialmente sua centralidade e distorcendo as conexões. A IA pode apoiar a padronização de nomes institucionais, mas deve operar com listas de autoridade, identificadores persistentes e revisão humana. Nunca deve inventar afiliações ou fundir instituições apenas por similaridade superficial.

Em redes geográficas, a escolha do nível de agregação muda a interpretação. Redes por país podem destacar colaboração internacional, mas ocultar desigualdades regionais internas. Redes por cidade podem revelar polos urbanos, mas depender de qualidade de metadados de endereço. Redes por continente podem simplificar demais as relações. Além disso, a afiliação declarada em um artigo não corresponde necessariamente ao local onde a pesquisa foi financiada, realizada ou aplicada. Um estudo clínico pode envolver pacientes em um país, análise estatística em outro e autoria principal em um terceiro. A rede de copublicação internacional captura a relação de afiliação autoral registrada, não todo o circuito material da pesquisa.

As redes institucionais e geográficas também são afetadas por assimetrias de capacidade científica, língua, financiamento e indexação. Países com maior investimento em pesquisa, maior presença em periódicos internacionais e maior cobertura em bases bibliográficas tendem a aparecer mais conectados. Instituições de países periféricos podem ter colaborações intensas localmente, mas baixa visibilidade em bases globais. A IA pode ajudar a formular perguntas sobre dependência, concentração e periferia, mas deve ser instruída a evitar julgamentos simplistas. Baixa centralidade em uma base internacional não significa baixa relevância local; pode refletir idioma, tipo de produção, foco regional ou exclusão de periódicos nacionais.

Uma rede institucional pode ser interpretada em diferentes dimensões. A densidade indica nível geral de conectividade entre instituições. O grau ponderado mostra volume de colaboração. A centralidade de intermediação pode identificar instituições que conectam grupos antes separados. A modularidade pode revelar blocos regionais, temáticos ou políticos. A evolução temporal pode indicar internacionalização crescente, mudança de parceiros ou efeitos de programas de financiamento. A IA pode organizar essas dimensões em uma narrativa, mas deve diferenciar claramente entre padrão observado e possível explicação. Por exemplo, se uma instituição se torna mais central após 2018, isso pode resultar de política institucional, novo convênio, maior produção, mudança de base, crescimento do campo ou alteração na estratégia de busca. A causalidade exige evidências externas.

Redes geográficas podem ser combinadas com mapas cartográficos, fluxos e matrizes de colaboração. Contudo, visualizações geográficas carregam riscos retóricos. Linhas grossas entre países podem sugerir intensidade de cooperação, mas não mostram assimetria de liderança, autoria correspondente, financiamento ou transferência de conhecimento. Um país pode aparecer como parceiro frequente porque fornece dados ou campo empírico, enquanto outro concentra desenho teórico e publicação. Para evitar leituras coloniais ou celebratórias da internacionalização, é importante incorporar indicadores complementares: posição de autoria, financiamento, autores correspondentes, liderança de projeto, idioma, tema e benefício social. A IA pode sugerir esses indicadores em *prompts* de auditoria.

Em política científica, redes institucionais e geográficas são frequentemente usadas para orientar estratégias de colaboração, identificar parceiros, avaliar programas e mapear competências. O risco é converter mapas em rankings ou usar centralidade como critério automático de excelência. Um mapa pode indicar que uma instituição é ponte entre

comunidades, mas isso não significa que ela deva receber mais recursos sem considerar missão institucional, equidade, necessidades regionais e qualidade da colaboração. A IA deve ser usada para ampliar a reflexão, não para automatizar decisões. Um *prompt* responsável deve pedir que o modelo apresente implicações possíveis, limitações, riscos éticos e perguntas adicionais antes de recomendar qualquer ação de gestão.

A análise de redes institucionais e geográficas, por fim, demanda transparência. O relatório deve informar bases usadas, campos de afiliação, regras de desambiguação, tratamento de múltiplas afiliações, contagem inteira ou fracionada, recorte temporal, exclusões e limiares. Sem essa documentação, a rede se torna uma imagem persuasiva sem reprodutibilidade. Em estudos assistidos por IA, os *prompts* usados para padronizar afiliações, interpretar *clusters* e gerar narrativas devem ser anexados ao protocolo ou ao apêndice metodológico. Essa documentação permite que outros pesquisadores compreendam não apenas o mapa final, mas o caminho pelo qual ele foi produzido e interpretado.

11.7 Centralidade, modularidade, clusters e comunidades

A análise de redes científicas depende de conceitos estruturais que permitem descrever posições, conexões e agrupamentos. Entre os mais usados estão centralidade, densidade, grau, intermediação, proximidade, autovetor, modularidade, *cluster*, comunidade, componente e peso das arestas. Esses conceitos são poderosos porque transformam uma rede visualmente complexa em propriedades mensuráveis. Contudo, também são perigosos quando convertidos automaticamente em julgamentos substanti*vos*. Uma centralidade alta não significa necessariamente importância científica; uma comunidade detectada por algoritmo não corresponde necessariamente a uma escola teórica; uma posição periférica não indica irrelevância; uma modularidade alta não prova divisão epistêmica profunda. A interpretação exige vinculação entre métrica, método de construção e pergunta de pesquisa.

A centralidade de grau mede o número de conexões de um nó, ou a soma ponderada dessas conexões quando as arestas possuem peso. Em redes de coautoria, alto grau pode indicar colaboração ampla. Em redes de cocitação, pode indicar documento ou autor frequentemente associado a muitos outros. Em redes de palavras-chave, pode indicar termo transversal. Entretanto, alto grau também pode resultar de gênero documental, consórcios, termos genéricos ou limiares de inclusão. A IA pode explicar o significado de grau, mas deve ser orientada a apresentar interpretações alternativas. O mesmo valor numérico pode significar liderança, interdisciplinaridade, generalidade, ruído ou efeito de escala, dependendo da rede.

A centralidade de intermediação mede a frequência com que um nó aparece em caminhos mais curtos entre outros nós. Em termos intuiti*vos*, aponta possíveis pontes ou mediadores. Em mapas científicos, um artigo, autor ou instituição com alta intermediação pode conectar comunidades que de outra forma estariam separadas. Essa medida e atraente para interpretar transferência de conhecimento, interdisciplinaridade e posições estratégicas. Contudo, sua leitura requer prudência. A intermediação depende da topologia calculada, dos limiares de aresta, da normalização e da completude da rede. Um nó pode parecer ponte porque outros nós foram excluídos por limiar. A IA deve ser instruída a tratar intermediação como indício estrutural e não como prova de agência consciente ou influência causal.

A centralidade de autovetor e medidas relacionadas, como PageRank, valorizam nos conectados a outros nós também bem conectados. Em redes de citação ou cocitação, podem destacar elementos associados a núcleos influentes. Em redes institucionais, podem indicar conexão com parceiros centrais. Essas medidas são úteis para identificar posições em estruturas hierárquicas, mas podem reforçar efeitos cumulati*vos* e invisibilizar atores

periféricos relevantes. Quando a IA interpreta autovetor, deve ser solicitada a considerar que a métrica favorece conexões com centros já estabelecidos. Em estudos de equidade científica, essa característica deve ser explicitada como limite.

Modularidade refere-se a uma função de qualidade usada para detectar divisões da rede em grupos com conexões internas mais densas do que seria esperado por acaso em determinado modelo. Newman (2006) discutiu a modularidade como abordagem central para identificar comunidades em redes complexas. Em cientometria, algoritmos de *cluster*ização baseados em modularidade, VOS ou métodos relacionados são usados para agrupar documentos, autores, termos ou instituições. A existência de *clusters*, porém, não garante que esses agrupamentos tenham significado teórico claro. *Clusters* são resultados de escolhas: tipo de relação, normalização, algoritmo, resolução, limiar e período. A IA deve ser orientada a interpretar *clusters* como agrupamentos produzidos por uma operacionalização específica, e não como categorias naturais do campo.

A noção de comunidade é ainda mais delicada. Em redes sociais, comunidade pode sugerir interação, identidade ou pertencimento. Em redes bibliométricas, comunidade pode ser apenas um conjunto de documentos com referências semelhantes, termos comuns ou padrões de cocitação. Um *cluster* de cocitação pode ser chamado de base intelectual; um *cluster* de acoplamento pode ser frente de pesquisa; um *cluster* de co-palavras pode ser tema; um *cluster* de coautoria pode ser grupo colaborativo. A mesma palavra 'comunidade' não deve ser usada indiscriminadamente para todos esses casos. Uma boa prática é nomear o tipo de agrupamento conforme a rede: comunidade de coautoria, *cluster* de cocitação, frente acoplada, grupo terminológico, bloco institucional. A IA pode ajudar a manter essa precisão terminológica se o *prompt* a exigir explicitamente.

A escolha de limiares é outro elemento decisivo. Para visualizar redes grandes, é comum excluir nos com baixa frequência ou arestas fracas. Essa decisão melhora a legibilidade, mas altera o sistema observado. Ao excluir palavras-chave raras, pode-se perder sinais emergentes. Ao excluir autores com uma única publicação, pode-se apagar periferias. Ao aumentar limiares de cocitação, pode-se privilegiar literatura canônica. Ao manter todos os nós, pode-se produzir um mapa ilegível. A IA pode auxiliar na comparação de cenários de limiar, mas deve receber resultados de sensibilidade: o que muda quando o limiar de ocorrência passa de 3 para 5, ou de 5 para 10? A robustez de uma interpretação depende da estabilidade do padrão diante de parâmetros razoáveis.

Visualizações de rede exercem forte poder persuasivo. Cores, tamanhos, posições e proximidades sugerem ordens interpretativas. Em mapas de distância, nós próximos parecem semelhantes; em grafos de força, a posição pode depender de parâmetros do algoritmo; em layouts dinâmicos, pequenas alterações podem modificar a organização visual. Por isso, a visualização deve ser acompanhada por descrição técnica. Tamanho do nó, cor do *cluster*, espessura da aresta e distância no mapa precisam ser definidos. A IA pode escrever legendas e descrições, mas deve ser instruída a não atribuir significado a elementos gráficos sem conhecer o algoritmo de layout. Em particular, proximidade visual não deve ser confundida automaticamente com causalidade, dependência ou afinidade substantiva.

Um relatório cientométrico responsável deve apresentar as métricas de rede como linguagem de descrição estrutural. A interpretação deve ser gradual: primeiro, informar como a rede foi construída; segundo, apresentar propriedades globais; terceiro, descrever *clusters* e nós centrais; quarto, propor hipóteses com base em conhecimento de domínio; quinto, apontar limitações e validações necessárias. A IA pode apoiar cada uma dessas etapas se o *prompt* for desenhado para conter o entusiasmo interpretativo. O comando 'explique este mapa' é

inadequado. Melhor é solicitar: descreva a estrutura; identifique padrões; proponha hipóteses; indique evidências faltantes; liste riscos de erro; reformule conclusões que parecem causais.

Quadro 11.3 - Glossário de métricas de rede

Métricas de nó (centralidade)	Propriedades estruturais da rede
Grau (degree): número de conexões de um nó, ou a soma dos pesos em redes ponderadas.	**Densidade:** proporção de arestas existentes em relação ao total possível.
Intermediação (betweenness): frequência com que um nó está nos caminhos mais curtos entre outros nós.	**Modularidade (Q):** mede quão bem a rede está dividida em grupos, ou comunidades.
Proximidade (closeness): inverso da distância média de um nó a todos os demais.	**Cluster ou comunidade:** conjunto de nós mais conectados entre si do que com o restante da rede.
Autovetor (eigenvector): valoriza nós conectados a outros nós que também são bem conectados.	**Componente:** subconjunto de nós interconectados entre si.
PageRank: versão ponderada do autovetor, usada em redes direcionadas.	**Peso da aresta:** intensidade da relação entre dois nós.

Fonte: Elaborado pelo autor (2026).

11.8 Como usar IA para interpretar redes sem inventar causalidades

A interpretação de redes científicas é uma tarefa inferencial. Ela parte de dados relacionais observados e busca construir explicações plausíveis sobre estrutura intelectual, colaboração, circulação de conhecimento, emergência de temas ou configuração institucional. A IA generativa pode ajudar nessa tarefa porque sintetiza informações, reorganiza argumentos, compara padrões e explicita incertezas. Contudo, ela também pode produzir narrativas causais sedutoras sem base empírica. O risco é maior quando o *prompt* pede 'explique por que' uma rede apresenta determinado formato sem oferecer evidências externas. Diante de uma comunidade central, o modelo pode afirmar que ela lidera o campo; diante de uma ponte, pode dizer que ela promoveu transferência de conhecimento; diante de um *cluster* emergente, pode sugerir que ele resultou de uma política pública. Todas essas afirmações podem ser plausíveis, mas não decorrem automaticamente da rede.

A primeira regra para usar IA na interpretação de redes é separar descrição, inferência e hipótese. Descrição é aquilo que se observa na rede: há três *clusters* principais; o *cluster* azul concentra autores de uma instituição; o termo '*deep learning*' conecta dois grupos; o documento X possui alta intermediação; a colaboração internacional cresce após 2019. Inferência é uma leitura plausível baseada no padrão e no conhecimento de domínio: o *cluster* azul pode representar uma linha institucional; o termo '*deep learning*' pode funcionar como ponte metodológica; o crescimento após 2019 pode refletir maior internacionalização. Hipótese é uma proposição que exige teste adicional: a política institucional *PROMPT* estimulou a formação do *cluster*; o financiamento Z explica a colaboração; o artigo X mudou a agenda do campo. Um *prompt* responsável deve exigir essa separação de forma explícita.

A segunda regra é exigir evidências necessárias para qualquer explicação causal. Se a IA propuser que um programa de financiamento causou aumento de colaboração, deve indicar quais dados confirmariam isso: datas do programa, participantes financiados, comparação

antes e depois, grupo de comparação, documentos de política, entrevistas ou análise de autoria correspondente. Se sugerir que um autor intermediário integrou comunidades, deve indicar evidências: coautorias com ambos os grupos, publicações interdisciplinares, citações cruzadas, participação em eventos ou projetos. Se afirmar que um tema emergiu por influência de uma tecnologia, deve verificar séries temporais, mudanças de terminologia e documentos-chave. A rede fornece pistas; a causalidade exige desenho analítico mais amplo.

A terceira regra é controlar a falsa autoridade. Modelos de linguagem escrevem com fluidez, o que pode dar aparência de segurança a interpretações fracas. Em relatórios cientométricos, essa fluidez é perigosa porque transforma mapas complexos em explicações lineares. Para reduzir esse risco, *prompts* devem pedir que o modelo atribua nível de confiança a cada afirmação, identifique a evidência que a sustenta e marque afirmações que exigem verificação humana. Também é recomendável solicitar uma versão conservadora da interpretação, na qual conclusões fortes sejam substituídas por formulações prudentes. Por exemplo, 'o autor X lidera a área' pode ser reformulado como 'o autor X apresenta alta conectividade na rede analisada, o que sugere papel relevante de articulação no *corpus* considerado.

A quarta regra é incorporar análise de sensibilidade. Uma interpretação é mais robusta quando se mantém sob diferentes parâmetros razoáveis: outro limiar, outra normalização, outra base, outra janela temporal ou outro algoritmo de *cluster*ização. A IA pode ajudar a comparar cenários, desde que receba os resultados. Um *prompt* pode apresentar duas versões da rede e pedir ao modelo que identifique padrões estáveis e instáveis. Padrões estáveis podem sustentar conclusões mais fortes; padrões instáveis devem ser tratados como dependentes de parâmetro. Esse procedimento é especialmente importante em redes de co-palavras e cocitação, nas quais pequenos ajustes podem alterar *clusters* é centralidades.

A quinta regra é triangulação. Redes de coautoria, cocitação, acoplamento e co-palavras capturam dimensões diferentes da ciência. Uma interpretação responsável deve, quando possível, comparar essas dimensões. Um grupo pode ser central na coautoria, mas periférico na cocitação; um tema pode ser frequente em palavras-chave, mas ainda não ter base intelectual consolidada; duas instituições podem colaborar muito, mas em tópicos pouco relacionados; um documento pode ser amplamente cocitado, mas não estar conectado a uma frente contemporânea. A IA pode construir quadros de triangulação e apontar convergências e tensões, evitando que um único mapa seja tratado como retrato completo do campo.

A sexta regra é considerar vieses de dados. Bases bibliográficas diferem em cobertura, idiomas, periódicos, tipos documentais e metadados. Redes baseadas em Scopus, Web of Science, PubMed, Dimensions, Lens, OpenAlex ou bases institucionais podem produzir resultados diferentes. Erros de desambiguação de autores e instituições podem criar ou apagar conexões. Duplicatas podem aumentar pesos de aresta. Palavras-chave inconsistentes podem fragmentar temas. A IA deve ser sempre informada dessas limitações e convidada a avaliar como elas podem afetar a rede. Um *prompt* de auditoria deve perguntar explicitamente: quais conclusões seriam vulneráveis a viés de base? Quais dependem de limpeza de dados? Quais poderiam mudar com outra fonte?

A sétima regra é preservar a responsabilidade autoral humana. Mesmo que a IA auxilie a interpretar, revisar, rotular e redigir, o pesquisador continua responsável pelo desenho do estudo, pela qualidade dos dados, pela escolha das métricas, pela validação das conclusões e pela comunicação das limitações. A IA não deve ser apresentada como fonte de autoridade interpretativa. Seu papel deve ser descrito metodologicamente: quais *prompts* foram usados, quais dados foram fornecidos, que respostas foram consideradas,

como foram revisadas e que decisões humanas foram tomadas. Essa transparência é essencial para reprodutibilidade e integridade científica.

O *prompt* central deste capítulo foi desenhado exatamente para conter a tendência de causalidade dos modelos generativos. Ao solicitar que a resposta seja separada em descrição estrutural, possíveis explicações, hipóteses interpretativas, evidências necessárias e riscos de interpretação indevida, o pesquisador transforma o modelo em um dispositivo de disciplina epistêmica. A IA deixa de ser convidada a 'contar a história' da rede e passa a ser obrigada a classificar o estatuto de cada afirmação. Essa mudança aparentemente simples é uma das principais contribuições da engenharia de *prompts* para a cientometria relacional.

11.9 Cadeia metodológica para análise relacional assistida por IA

Uma análise relacional assistida por IA deve ser planejada como cadeia metodológica. A primeira etapa é a formulação da pergunta. A pergunta deve definir que tipo de relação é relevante: colaboração, proximidade intelectual, compartilhamento de referências, coocorrência terminológica, coparticipação institucional ou fluxo geográfico. Perguntas diferentes exigem redes diferentes. 'Quem colabora com quem?*prompt* remete a coautoria. 'Quais textos formam a base intelectual de uma área?*prompt* remete a cocitação. 'Quais frentes recentes compartilham fundamentos bibliográficos?*prompt* remete a acoplamento. 'Como os temas se articulam?*prompt* remete a co-palavras. Pedir que a IA escolha o tipo de rede pode ser útil, mas a decisão final deve ser justificada por alinhamento entre pergunta, dados e método.

A segunda etapa é a definição do *corpus*. Redes são construídas sobre conjuntos de documentos. Se o *corpus* e enviesado, a rede também será. A estratégia de busca, as bases utilizadas, os critérios de inclusão, o período e o tratamento de duplicatas determinam a fronteira do sistema analisado. A IA pode auxiliar a revisar se a estratégia de busca é coerente com a pergunta, mas não deve inventar bases, sintaxes ou registros. Em estudos reprodutíveis, o *corpus* bruto deve ser preservado e a versão processada deve ser documentada. Quando redes forem comparadas, deve-se garantir que as diferenças não sejam apenas artefatos de *corpus*.

A terceira etapa é a construção da matriz relacional. Em coautoria, a matriz registra copublicação; em cocitação, documentos citados juntos; em acoplamento, referências compartilhadas; em co-palavras, termos que ocorrem juntos. Essa matriz pode ser binária ou ponderada, simétrica ou assimétrica, normalizada ou bruta. A IA pode ajudar a explicar essas escolhas e suas implicações, mas a matriz deve ser gerada por procedimentos controlados, preferencialmente em ferramentas que permitam reprodutibilidade. Quando a IA participa da limpeza ou classificação textual, seus *prompts* e saídas devem ser armazenados como parte do método.

A quarta etapa é a análise estrutural. Nesse momento, calculam-se métricas globais e locais: número de nós, número de arestas, densidade, componentes, grau, pesos, centralidades, modularidade e *clusters*. A IA pode receber tabelas com essas métricas e produzir uma descrição interpretativa inicial. O *prompt* deve proibir julgamentos de qualidade e exigir que cada afirmação seja associada a uma métrica. Por exemplo, 'o *cluster* 2 é tematicamente central' deve ser substituído por 'o *cluster* 2 apresenta maior número de documentos e contém os termos de maior grau; sua centralidade temática requer confirmação pela leitura dos documentos'.

A quinta etapa é a interpretação substantiva. Aqui entram conhecimento de domínio, leitura de documentos, comparação com literatura, contexto institucional e perguntas de pesquisa. A IA pode propor hipóteses, rótulos e narrativas, mas deve ser submetida a auditoria. Uma prática recomendada é usar dois *prompts*: um gerador, que produz

interpretações, e um crítico, que revisa exageros, causalidades indevidas e evidências ausentes. Esse uso dialético da IA é mais seguro do que aceitar a primeira resposta. Em cientometria, a qualidade interpretativa depende tanto da capacidade de formular hipóteses quanto da capacidade de limitar conclusões.

A sexta etapa é a redação do relatório. Redes científicas podem ser difíceis de comunicar, especialmente para públicos não especialistas. A IA pode ajudar a transformar métricas em texto claro, escrever legendas, explicar tipos de rede e adaptar a linguagem a gestores, pesquisadores ou estudantes. No entanto, a redação final deve preservar o rigor: definir a rede, indicar dados, explicar métricas, apresentar achados, evitar causalidade não sustentada e explicitar limitações. O modelo pode ser instruído a usar verbos prudentes, como 'sugere', 'indica', 'é consistente com', 'pode refletir', em vez de 'prova', 'causa' ou 'demonstra', quando a evidência for apenas relacional.

A sétima etapa é a documentação. Um estudo cientométrico assistido por IA deve registrar versão das bases, data da coleta, estratégia de busca, ferramentas, parâmetros, regras de limpeza, *prompts*, saídas relevantes, critérios de aceitação e revisão humana. Essa documentação é o equivalente metodológico da rastreabilidade laboratorial. Ela permite que outro pesquisador compreenda como a rede foi produzida e como a IA participou do processo. Sem essa transparência, a IA se torna uma caixa-preta adicional sobre um método que já possui muitas decisões técnicas.

Quadro 11.4 - Fluxo da análise relacional assistida por IA

Nº	Etapa e procedimentos	Referência de apoio
01	**Formulação da pergunta** • Defina o objetivo analítico e o fenômeno relacional. • Especifique o tipo de relação relevante. • Delimite unidades de análise (documentos, autores, instituições, países, termos etc.). • Estabeleça o período e o nível geográfico.	**Exemplos de perguntas e redes adequadas** **Quem colabora com quem?:** *Coautoria* **Quais textos formam a base intelectual de uma área?:** *Cocitação* **Quais frentes recentes compartilham fundamentos bibliográficos?:** *Acoplamento bibliográfico* **Como os temas se articulam?:** *Coocorrência de palavras-chave* **Como instituições e países se conectam?:** *Redes institucionais e geográficas*
02	**Definição do** *corpus* • Escolha das bases de dados (ex.: Web of Science, Scopus, Dimensions, OpenAlex, Lens). • Estratégia de busca e campos. • Critérios de inclusão e exclusão. • Tipos documentais e idiomas. • Período temporal. • Remoção de duplicatas. • Registro do número de registros e da cobertura.	**A IA pode ajudar a revisar** • Coerência da estratégia com a pergunta. • Lacunas de cobertura (bases, idioma, período). • Riscos de viés e redundância. • Sugestões de termos e sinônimos (sem executar a busca). *Preserve o corpus bruto (exportação original) e a versão processada.*

Nº	Etapa e procedimentos	Referência de apoio
03	**Construção da matriz relacional** • Defina a relação a ser computada (coautoria, cocitação, acoplamento, coocorrência etc.). • Escolha o nível da unidade de análise. • Defina a ponderação da relação: binária ou ponderada. • Normalização: bruta ou normalizada (ex.: cosseno, Jaccard, associação). • Gere a matriz de adjacência.	**Decisões críticas** • Tipo de relação e direção (simétrica ou assimétrica). • Método de ponderação e normalização. • Limiares mínimos para inclusão de nós e arestas. • Tratamento de múltiplas afiliações e de palavras-chave. *Registre ferramentas, versões, parâmetros e uso de IA, prompts e saídas relevantes.*
04	**Análise estrutural da rede** • Estatísticas globais: nós, arestas, densidade, componentes. • Métricas locais: grau, intermediação, proximidade, autovetor, PageRank. • Detecção de *clusters* ou comunidades e modularidade. • Análise de centralidade por *cluster*. • Sensibilidade: variação de limiares, algoritmos e janelas temporais.	**O que a IA pode fazer (com dados)** • Descrever a estrutura da rede. • Explicar o significado das métricas. • Comparar cenários (limiares, métodos). • Apontar nós e arestas destacados. • Gerar tabelas e resumos estruturados. • Exija que toda afirmação seja vinculada a uma métrica e ao método utilizado.
05	**Interpretação substantiva** • Integre a rede com o conhecimento de domínio. • Leia documentos centrais de cada *cluster* ou comunidade. • Formule hipóteses sobre temas, escolas, frentes, agendas ou políticas. • Considere contexto histórico, disciplinar e institucional. • Triangule com outras técnicas e indicadores.	**Uso dialético da IA** **Gerador (hipóteses):** sugerir rótulos, narrativas, relações plausíveis e explicações alternativas. **Crítico (auditoria):** identificar exageros, causalidades indevidas, vieses e evidências ausentes. Sempre diferencie padrões observados, explicações ou causas e recomendações.
06	**Redação do relatório** • Explique o tipo de rede e o método. • Apresente dados e métricas principais. • Descreva *clusters*, nós centrais e relações relevantes. • Use linguagem prudente e proporcional à evidência. • Indique limitações, incertezas e possíveis vieses.	**A IA pode apoiar** • Redigir opções e legendas de figuras. • Traduzir métricas em linguagem clara. • Adaptar o texto ao público-alvo. • Sugerir estrutura de argumentação. • Revisar consistência interna. *Use verbos prudentes: sugere, indica, é consistente com, pode refletir.*
07	**Documentação e rastreabilidade** • Registre toda a cadeia metodológica. • Preserve dados, scripts, parâmetros, *prompts* e versões. • Descreva decisões, critérios e revisões humanas. • Disponibilize materiais quando possível e permitido. • Permita auditoria e reprodutibilidade.	**O que documentar** • Bases, data de coleta e versão. • Estratégia de busca (sintaxe e campos). • Critérios de inclusão e exclusão. • Limpeza e normalização dos dados. • Matriz relacional e parâmetros. • Métricas e versões utilizadas. • *Prompts* da IA e principais saídas. • Revisões humanas e decisões finais. *Transparência gera confiança e permite replicação e crítica construtiva.*

Fonte: Elaborado pelo autor (2026).

11.10 Prompts do capítulo

Os *prompts* deste capítulo foram elaborados para apoiar pesquisadores em etapas distintas da análise relacional. Eles não devem ser usados como comandos mecânicos, mas como modelos adaptáveis. Em cada caso, recomenda-se inserir informações sobre a pergunta de pesquisa, tipo de rede, base de dados, período, método de normalização, limiares, métricas disponíveis e limitações conhecidas. Quanto mais específico for o contexto, menor a chance de respostas genéricas ou causalidades inventadas.

O *prompt* central do Capítulo deve ser utilizado quando o pesquisador já possui uma rede descrita por métricas e deseja uma interpretação prudente. Ele é especialmente adequado para relatórios, artigos e diagnósticos institucionais. A chave está em obrigar o modelo a separar estatutos epistêmicos. A descrição estrutural deve conter apenas o que está nos dados. As explicações possíveis devem ser formuladas como alternativas. As hipóteses devem indicar evidência necessária. Os riscos devem explicitar vieses de base, parâmetros, normalização, visualização e falta de validação qualitativa.

O *prompt* para redes de coautoria deve enfatizar que coautoria e evidência de copublicação, não medida perfeita de colaboração. O modelo deve avaliar densidade, componentes, autores ou instituições centrais, *clusters* e possíveis assimetrias. Deve também perguntar se há autoria em consórcio, contagem inteira ou fracionada, múltiplas afiliações e diferenças disciplinares. Quando usado em gestão de pesquisa, o *prompt* deve evitar recomendações automáticas de política e solicitar implicações com cautela.

O *prompt* para cocitação deve orientar a IA a identificar bases intelectuais e não frentes necessariamente atuais. Deve pedir que o modelo considere maturação citacional, documentos clássicos, períodos, viés de idioma e rotulagem de *clusters*. Uma boa resposta deve distinguir documentos fundacionais, documentos metodológicos, revisões e textos de controvérsia. Deve também sugerir leitura de uma amostra dos documentos centrais antes de nomear comunidades intelectuais.

O *prompt* para acoplamento bibliográfico deve destacar que proximidade decorre de referências compartilhadas. O modelo deve examinar se o acoplamento parece teórico, metodológico, empírico ou normativo. Também deve considerar o tamanho das listas de referências e o método de normalização. Quando a rede se refere a artigos recentes, a IA deve reconhecer que o acoplamento é útil para frentes atuais, mas não demonstra impacto futuro.

O *prompt* para coocorrência de palavras-chave deve exigir cautela terminológica. O modelo deve classificar termos amplos, específicos, metodológicos, geográficos e emergentes; sugerir agrupamentos; indicar ambiguidades; e propor rótulos alternati*vos* para *clusters*. Esse *prompt* deve ser sempre acompanhado de revisão humana, pois a IA pode agrupar termos por similaridade superficial e perder diferenças conceituais relevantes.

O *prompt* de auditoria causal deve ser usado após qualquer interpretação de rede. Ele deve receber um trecho de texto e identificar afirmações que sugerem causalidade, liderança, influência, excelência ou estratégia sem evidência suficiente. Em seguida, deve reformular as afirmações em linguagem prudente e indicar quais dados adicionais seriam necessários para sustentar conclusões mais fortes. Esse *prompt* é um dos instrumentos mais importantes para evitar o uso retórico de mapas científicos.

Prompt 11.10.1 Prompt central: interpretação prudente de rede científica

Atue como especialista em cientometria e análise de redes científicas.

Interprete a rede descrita abaixo. Separe a resposta em cinco partes:

1) descrição estrutural baseada apenas nos dados fornecidos;

2) possíveis explicações, formuladas como alternativas e não como certezas;

3) hipóteses interpretativas que exigem evidência adicional;

4) evidências necessárias para confirmar ou refutar cada hipótese;

5) riscos de interpretação indevida, incluindo viés de base de dados, parâmetro de corte, normalização, layout visual e inferência causal.

Não afirme causalidade quando houver apenas associação. Não use termos como prova, causa, lidera ou determina sem evidência externa.

Prompt 11.10.2 Prompt para redes de coautoria

A partir das métricas e da descrição da rede de coautoria abaixo, produza uma análise cientométrica prudente.

Considere autores, instituições ou países como nós, conforme indicado.

Explique o significado de grau, peso das arestas, componentes, densidade e *clusters*.

Diferencie copublicação registrada de colaboração substantiva.

Aponte possíveis efeitos de consórcios, múltiplas afiliações, contagem inteira ou fracionada e diferenças disciplinares de autoria.

Finalize com perguntas de validação que o pesquisador deve responder antes de tirar conclusões sobre colaboração.

Prompt 11.10.3 Prompt para redes de cocitação

Analise a rede de cocitação descrita abaixo.

Identifique possíveis bases intelectuais, documentos ou autores centrais, *clusters* e relações entre tradições. Indique se os *clusters* parecem teóricos, metodológicos, empíricos ou mistos.

Considere que cocitação favorece literatura mais consolidada e depende da janela temporal. Sugira rótulos alternati*vos* para cada *cluster* e justifique-os com base em títulos, resumos, palavras-chave ou referências fornecidas.

Marque como hipótese qualquer afirmação sobre escolas, paradigmas ou influência.

Prompt 11.10.4 Prompt para redes de acoplamento bibliográfico

Interprete a rede de acoplamento bibliográfico abaixo, lembrando que a proximidade entre documentos resulta de referências compartilhadas.

Indique possíveis frentes de pesquisa contemporâneas, bases bibliográficas comuns, articulações metodológicas e divergências temáticas.

Explique como tamanho das listas de referências, normalização e limiar de arestas podem afetar os resultados.

Não conclua que documentos acoplados possuem os mesmos resultados ou conclusões; trate a proximidade como compartilhamento de base intelectual.

Prompt 11.10.5 Prompt para redes de coocorrência de palavras-chave

Receba a lista de palavras-chave, seus *clusters* e suas frequências.

Analise a estrutura temática da rede.

Para cada *cluster*, proponha até três rótulos alternati*vos*, indique os termos que sustentam cada rótulo, identifique termos ambíguos ou genéricos, sugira sinônimos ou variantes que podem ter sido separados indevidamente e proponha verificações humanas.

Não elimine termos ambíguos; marque-os para revisão.

Diferencie termos de objeto, método, teoria, população, contexto geográfico e aplicação.

Prompt 11.10.6 Prompt de auditoria causal

Revise o texto interpretativo abaixo e identifique afirmações que extrapolam os dados da rede.

Classifique cada afirmação como: descrição sustentada, inferência plausível, hipótese sem evidência suficiente ou causalidade indevida.

Para cada caso problemático, proponha uma reformulação mais prudente e indique quais dados adicionais seriam necessários para sustentar uma conclusão mais forte.

Preserve o tom científico e evite linguagem excessivamente assertiva.

11.11 Estudo de caso ilustrativo: interpretando uma rede sem causalidades

Considere um estudo hipotético sobre produção científica em inteligência artificial aplicada a saúde pública entre 2015 e 2025. O pesquisador coletou artigos em duas bases bibliográficas, removeu duplicatas, normalizou nomes institucionais e construiu quatro redes: coautoria entre instituições, cocitação de documentos, acoplamento bibliográfico entre artigos recentes e coocorrência de palavras-chave. A rede institucional mostra três *clusters*:

um grupo de universidades norte-americanas e europeias, um grupo de instituições asiáticas e um grupo menor de universidades latino-americanas conectadas a hospitais públicos. A rede de cocitação evidência documentos sobre aprendizado profundo, epidemiologia digital e ética de dados. A rede de acoplamento mostra artigos recentes que compartilham referências sobre vigilância epidemiológica, modelos preditivos e equidade algorítmica. A rede de palavras-chave aproxima '*machine learning*', 'public health', 'surveillance', 'COVID-19', 'bias' e 'electronic health records'.

Uma interpretação apressada poderia afirmar que universidades norte-americanas lideram a área, que a pandemia causou a emergência do *cluster* de vigilância é que a América Latina ocupa posição periférica por menor capacidade científica. Essas conclusões são problemáticas. A rede mostra maior conectividade de certas instituições no *corpus*, mas não mede liderança em sentido amplo. A presença de COVID-19 e vigilância pode estar associada ao período, mas a causalidade exige análise temporal e comparação antes e depois da pandemia. A posição periférica de instituições latino-americanas pode refletir cobertura das bases, idioma, critérios de busca, menor indexação de periódicos regionais ou diferenças de financiamento. A IA, se solicitada de modo genérico, poderia transformar essas leituras em narrativa convincente. O *prompt* adequado deve obrigá-la a classificar tais afirmações como hipóteses.

A resposta responsável com apoio de IA poderia ser estruturada assim: a rede institucional apresenta maior densidade no *cluster* euro-norte-americano e conexões menos númerosas no *cluster* latino-americano; essa configuração sugere assimetria de conectividade no *corpus* analisado; possíveis explicações incluem diferenças de financiamento, cobertura de bases, idioma, histórico de colaboração e concentração de periódicos indexados; para confirmar essas explicações, seriam necessários dados de financiamento, autoria correspondente, idioma das publicações, comparação com bases regionais e análise temporal; não se deve concluir que uma região e cientificamente inferior ou que outra lidera a área sem critérios adicionais de qualidade, impacto, aplicabilidade e relevância local. Essa formulação conserva a utilidade do mapa sem transformar associação em julgamento normativo.

No mesmo caso, a rede de cocitação pode revelar que documentos de aprendizado profundo e ética de dados aparecem em *clusters* próximos. Uma interpretação causal indevida diria que a literatura ética influenciou diretamente os estudos de aprendizado profundo. Uma interpretação prudente diria que, no *corpus*, documentos sobre esses temas são cocitados com alguma frequência, sugerindo que parte da literatura os mobiliza conjuntamente. Para afirmar influência, seria preciso examinar trajetórias de citação, textos que articulam explicitamente os temas, mudança temporal nas referências e possíveis documentos ponte. A IA pode ajudar a identificar documentos candidatos a ponte, mas a leitura qualitativa permanece indispensável.

A rede de acoplamento bibliográfico pode indicar que artigos recentes compartilham referências sobre equidade algorítmica. Isso permite levantar a hipótese de que ha uma frente de pesquisa emergente em IA responsável para saúde pública. Contudo, o acoplamento não prova que esses artigos adotem a mesma perspectiva ética ou cheguem a conclusões convergentes. Alguns podem defender auditorias algorítmicas, outros podem criticar vieses de dados, outros podem propor regulação. O *prompt* deve pedir que a IA diferencie base comum de posicionamento comum. Essa distinção é essencial em temas normativamente carregados.

A rede de palavras-chave pode sugerir que 'COVID-19' funciona como termo articulador entre vigilância, saúde pública e modelos preditivos. Mas a centralidade de COVID-19 pode refletir a estratégia de busca ou o período analisado. Para confirmar emergência temática, seria necessário comparar janelas temporais, verificar a frequência antes e depois de 2020, analisar termos substitutos e examinar documentos de cada *cluster*. A IA pode gerar um

plano de validação temporal, mas não deve concluir emergência apenas com base em uma rede agregada. O estudo de caso mostra que a principal competência do pesquisador não é obter uma interpretação automática, mas desenhar perguntas que transformem o modelo em parceiro crítico.

11.12 Matriz de riscos interpretativos em redes científicas

Quadro 11.5 - Matriz de riscos interpretativos em redes científicas

Risco principal	Descrição	Exemplo comum	Possíveis efeitos	O que fazer
01 Ontologização dos clusters	Tratar agrupamentos algorítmicos como comunidades naturais, escolas teóricas ou áreas plenamente reconhecidas.	Dizer que um *cluster* "é uma escola teórica" sem validação externa.	Reificação de agrupamentos provisórios; decisões de política baseadas em estruturas instáveis; simplificação excessiva.	Tratar *clusters* como hipóteses estruturais. Validar com leitura qualitativa e literatura.
02 Causalidade indevida	Inferir causalidade a partir de conexões, proximidades ou centralidades.	Afirmar que "o autor A influencia B" apenas porque estão conectados.	Conclusões causais sem evidência; decisões estratégicas baseadas em associação.	Usar linguagem de associação, não de causa. Exigir evidências externas para afirmações causais.
03 Avaliação indevida	Usar métricas da rede (tamanho, densidade, centralidade) como *proxy* automático de qualidade.	Dizer que um *cluster* maior é "melhor" ou que um autor central é "o mais importante".	Hierarquias injustificadas; avaliações normativas; concentração de recursos prolongada.	Diferenciar descrição de avaliação. Evitar termos valorativos sem critérios explícitos.
04 Invisibilização	Excluir periferias, documentos recentes, termos raros ou produções locais por causa de limiares e filtros.	Definir limiar alto e remover termos com baixa frequência.	Perda de diversidade; visão parcial do campo; sub-representação de novos temas e regiões.	Testar a sensibilidade com múltiplos limiares. Analisar as periferias explicitamente.
05 Dependência visual	Interpretar distâncias, cores e agrupamentos sem conhecer o algoritmo de layout e os parâmetros usados.	Concluir que nós próximos "são naturalmente relacionados" pelo layout.	Interpretações baseadas em artefatos gráficos; confusão entre algoritmo e estrutura real.	Declarar algoritmo, parâmetros e métricas. Não inferir relações apenas pela posição no mapa.

Perguntas de controle — aplicáveis a todos os riscos, sobretudo em prompts com IA

1. Que dado sustenta esta afirmação?
2. Que parâmetro poderia alterá-la?
3. Que evidência externa seria necessária?
4. Que alternativa explicativa existe?
5. Que grupo ou elemento pode ter sido invisibilizado?
6. Que viés de base pode estar presente?

Fonte: Elaborado pelo autor (2026).

A interpretação de redes científicas envolve riscos recorrentes. O primeiro é o risco de ontologização dos *clusters*: tratar agrupamentos algorítmicos como se fossem comunidades naturais, escolas teóricas ou áreas institucionais plenamente reconhecidas. O segundo é o risco de casualização: transformar conexões, proximidades ou centralidades em causas. O terceiro é o risco de avaliação indevida: usar centralidade, densidade ou tamanho de *cluster* como *proxy* automático de qualidade. O quarto é o risco de invisibilização: excluir periferias, documentos recentes, termos raros ou produções locais por causa de limiares. O quinto é o risco de dependência visual: interpretar distâncias e cores sem conhecer o algoritmo de layout e os parâmetros usados.

A IA pode amplificar esses riscos porque tende a produzir narrativas coerentes. Se receber apenas uma imagem ou uma descrição vaga, pode inferir significados inexistentes. Por isso, a matriz de riscos deve ser incorporada aos *prompts*. Para cada interpretação, o modelo deve responder: que dado sustenta esta afirmação? Que parâmetro poderia alterá-la? Que evidência externa seria necessária? Que alternativa explicativa existe? Que grupo pode ter sido invisibilizado? Que viés de base pode estar presente? Essas perguntas transformam o uso de IA em um processo de controle metodológico.

Outro risco importante é a confusão entre centralidade e poder. Em redes de coautoria, um autor central pode ter alta produtividade colaborativa, mas não necessariamente maior poder intelectual. Em redes institucionais, uma universidade central pode ter ampla rede internacional, mas não necessariamente distribuir benefícios de modo equitativo. Em redes de cocitação, um documento central pode ser texto metodológico de uso transversal, não obra teórica dominante. Em redes de co-palavras, um termo central pode ser genérico. A IA deve ser instruída a substituir substanti*vos* fortes por descrições relacionais: 'central na rede analisada', 'amplamente conectado nó *corpus*', 'frequentemente cocitado', 'termo transversal', e não 'mais importante', 'dominante' ou 'causador'.

Há também riscos técnicos. Erros de desambiguação podem criar autores duplicados ou fundir homônimos. Duplicatas de documentos podem aumentar artificialmente pesos. Citações inconsistentes podem fragmentar referências. Palavras-chave sem normalização podem separar termos equivalentes. Bases incompletas podem excluir periódicos relevantes. Limiar alto pode remover conexões fracas, mas importantes. Layouts podem aproximar elementos por efeito gráfico. A IA não detecta automaticamente esses problemas se eles não forem informados. Assim, *prompts* devem incluir um bloco de metadados sobre tratamento dos dados e pedir ao modelo que avalie como cada decisão técnica afeta a interpretação.

Riscos éticos também precisam ser considerados. Redes geográficas podem reforçar hierarquias centro-periferia. Redes institucionais podem ser usadas para justificar concentração de recursos. Redes de coautoria podem expor indivíduos ou sugerir isolamento sem considerar contexto. Redes de citação podem reproduzir vieses de idioma, gênero, país e disciplina. A IA pode reproduzir esses vieses em sua linguagem. Um *prompt* de revisão ética deve solicitar que o modelo identifique possíveis efeitos normati*vos* da interpretação: quem e valorizado, quem e invisibilizado, que indicadores são usados como critério de legitimidade, que contexto social falta e que cautelas devem acompanhar a comunicação dos resultados.

A matriz de riscos deve resultar em Boas práticas. Toda análise de rede deve declarar unidade de análise, tipo de relação, fonte dos dados, período, critério de inclusão, normalização, algoritmo, limiares, métricas e limitações. Toda interpretação deve diferenciar descrição, inferência e hipótese. Toda conclusão causal deve exigir evidências externas. Todo uso de IA deve ser documentado. Toda visualização deve ser acompanhada por legenda técnica. Toda recomendação de política científica deve incluir análise de equidade e incerteza. Essas regras não reduzem a utilidade das redes; ao contrário, aumentam sua credibilidade científica.

11.13 Redação de resultados de redes com apoio de IA

A redação de resultados é uma etapa em que a IA pode ser particularmente útil. Muitos pesquisadores conseguem calcular métricas e gerar mapas, mas encontram dificuldade para transformar esses achados em texto científico claro, prudente e informativo. Um bom texto de resultados sobre redes deve combinar descrição objetiva, interpretação limitada e transição para discussão. Ele deve evitar dois extremos: a mera listagem de números sem sentido analítico e a narrativa excessiva que ultrapassa os dados. A IA pode ajudar a encontrar esse equilíbrio se o *prompt* fornecer dados e restrições adequadas.

Em uma seção de resultados, recomenda-se iniciar pela caracterização da rede: tipo de rede, número de nós e arestas, período, unidade de análise e critério de construção. Em seguida, apresentar propriedades globais: densidade, componentes, modularidade e distribuição de pesos. Depois, descrever os principais *clusters* e nós centrais, sempre indicando a métrica usada. Por fim, apontar padrões relevantes para a pergunta de pesquisa. A discussão causal, política ou teórica deve ser reservada para a seção de discussão e formulada como interpretação, não como fato bruto. A IA pode gerar uma primeira versão seguindo essa ordem.

A linguagem deve ser precisa. Em vez de escrever 'o mapa mostra que o tema X causou a integração entre áreas', é melhor escrever 'o termo X apresenta alta intermediação entre dois *clusters*, sugerindo possível papel articulador no *corpus* analisado'. Em vez de 'a instituição *PROMPT* lidera a rede', é melhor 'a instituição *PROMPT* apresenta o maior grau ponderado na rede de coautoria institucional'. Em vez de 'o *cluster* vermelho é a escola dominante', é melhor 'o *cluster* vermelho reúne os documentos mais cocitados e parece representar uma base intelectual consolidada, hipótese que requer leitura dos documentos centrais'. A IA deve ser treinada por *prompt* a preferir esse estilo.

Também é útil pedir que a IA produza versões para diferentes públicos. Para um artigo acadêmico, o texto deve incluir métodos e métricas. Para um relatório executivo, pode enfatizar implicações, mas mantendo cautela. Para gestores, pode apresentar perguntas estratégicas em vez de recomendações automáticas. Para estudantes, pode explicar conceitos como grau e modularidade em linguagem acessível. A adaptação de linguagem, contudo, não deve simplificar a ponto de apagar limitações. Uma boa comunicação cientométrica é aquela que torna a complexidade compreensível sem transformá-la em certeza indevida.

A IA também pode auxiliar na revisão de coerência. Um *prompt* pode comparar tabela, mapa e texto para verificar se há afirmações não sustentadas. Outro pode buscar inconsistências terminológicas, como alternar '*cluster*', 'comunidade' e 'tema' sem definição. Outro pode revisar se as limitações mencionam base de dados, limiar, normalização e janela temporal. Outro pode transformar uma conclusão forte em conclusão moderada. Esse uso revisional é talvez mais seguro do que a geração inicial, porque coloca a IA na função de crítica metodológica.

A redação final deve incluir transparência sobre o uso da IA. Quando *prompts* foram usados para interpretar, rotular ou revisar texto, recomenda-se registrar isso em nota metodológica, apêndice ou material suplementar, conforme normas editoriais. O registro deve incluir a finalidade do uso, os dados fornecidos ao modelo, os *prompts* principais e o processo de validação humana. Essa prática fortalece a reprodutibilidade e evita apresentar como leitura puramente humana aquilo que foi parcialmente assistido por modelo generativo.

11.14 Comparação entre tipos de rede e triangulação interpretativa

A escolha entre coautoria, cocitação, acoplamento bibliográfico, coocorrência de palavras-chave e redes institucionais não deve ser tratada como escolha estética, mas

como decisão epistemológica. Cada rede responde a uma pergunta diferente e, por isso, produz uma imagem parcial da ciência. A coautoria privilegia relações de coprodução registrada; a cocitação privilegia a memória intelectual produzida por citantes; o acoplamento privilegia bases bibliográficas compartilhadas; a coocorrência de palavras-chave privilegia vocabulários e articulações temáticas; as redes institucionais e geográficas privilegiam organizações e territórios. Quando um estudo usa apenas uma dessas redes, ele deve reconhecer explicitamente que está observando uma dimensão do sistema científico. Quando combina várias, pode produzir uma interpretação mais robusta, desde que não misture indevidamente seus significados.

A triangulação entre redes permite verificar se uma narrativa se sustenta em diferentes evidências relacionais. Suponha que um campo apresente um *cluster* de co-palavras relacionado a 'sustentabilidade', 'energia renovável' e 'política pública'. Se a rede de acoplamento também revela artigos recentes que compartilham referências sobre transição energética, e a rede de cocitação mostra uma base intelectual consolidada sobre governança ambiental, há uma convergência razoável entre vocabulário, frente contemporânea e base intelectual. A interpretação continua sendo inferencial, mas ganha apoio de diferentes dimensões. Por outro lado, se a rede de co-palavras apresenta o tema como central, mas a cocitação não mostra base consolidada e o acoplamento se dispersa, pode haver um tópico discursivamente frequente, porém ainda pouco estruturado como frente de pesquisa.

A IA pode apoiar a triangulação ao receber tabelas separadas para cada tipo de rede e produzir uma matriz de convergências e divergências. O *prompt* deve solicitar que o modelo identifique padrões confirmados por mais de uma rede, padrões exclusi*vos* de uma rede, contradições aparentes e explicações possíveis. Deve também pedir que o modelo classifique a robustez de cada interpretação. Uma afirmação sustentada por coautoria e rede institucional pode ser robusta para colaboração organizacional, mas não para influência intelectual. Uma afirmação sustentada por cocitação e acoplamento pode ser robusta para proximidade bibliográfica, mas não para coprodução. A robustez deve ser definida em relação à pergunta, não como valor geral.

A triangulação também ajuda a evitar o viés de visualização única. Mapas de redes são muito persuasi*vos* e, quando apresentados isoladamente, podem induzir a crença de que representam o campo em sua totalidade. Comparar redes mostra que um mesmo objeto científico possui várias estruturas sobrepostas. Um autor pode ser pouco colaborativo em coautoria, mas muito central em cocitação. Uma instituição pode ter ampla colaboração internacional, mas pouca presença nas referências centrais. Um termo pode ser frequente em palavras-chave, mas aparecer em documentos pouco conectados por acoplamento. Essas divergências não são necessariamente erros; elas revelam que a ciência é multidimensional.

Um relatório bem elaborado pode dedicar uma subseção a triangulação interpretativa. Nela, o pesquisador apresenta uma tabela com as principais evidências por tipo de rede, o nível de convergência entre redes, a interpretação mais prudente e as evidências adicionais necessárias. A IA pode ajudar a preencher uma primeira versão dessa tabela, mas a decisão final exige conhecimento de domínio. Uma boa regra é nunca permitir que a IA transforme convergência parcial em confirmação plena. Mesmo quando duas redes apontam para o mesmo tema, ainda pode haver viés de busca, base, período ou limpeza de dados. A triangulação aumenta plausibilidade, não elimina incerteza.

11.15 Reprodutibilidade, documentação e governança dos prompts em redes científicas

A reprodutibilidade em análise de redes científicas depende de uma cadeia de documentação que vai muito além do mapa final. É necessário registrar como o *corpus* foi coletado, quais bases foram usadas, em que data os dados foram exportados, quais campos foram selecionados, como duplicatas foram removidas, como autores e instituições foram desambiguados, que tipo de rede foi construída, qual método de normalização foi adotado, quais limiares foram aplicados, que algoritmo de layout e *cluster*ização foi utilizado e como as visualizações foram geradas. Quando a IA participa do processo, a documentação deve incluir também *prompts*, entradas, saídas, revisões humanas e decisões tomadas com base nessas saídas.

A governança de *prompts* é especialmente importante porque modelos generativos podem produzir respostas diferentes (LIU *et al.*, 2023; WEI *et al.*, 2022) para entradas semelhantes, sobretudo quando parâmetros de aleatoriedade, versões do modelo ou contexto conversacional mudam. Uma interpretação de rede gerada hoje pode não ser idêntica a uma interpretação gerada em outro momento. Por isso, o *prompt* usado para interpretar uma rede deve ser salvo junto com a resposta considerada, e não apenas descrito genericamente. Em estudos publicados, recomenda-se incluir os *prompts* principais em apêndice ou repositório, acompanhados dos dados resumidos fornecidos ao modelo. Essa prática permite que leitores avaliem se a IA foi usada de forma orientada, crítica e coerente.

A documentação também deve explicitar quais tarefas foram delegadas à IA e quais permaneceram sob responsabilidade humana. Por exemplo, a IA pode ter sido usada para sugerir rótulos de *clusters*, mas os rótulos finais podem ter sido revisados por especialistas. Pode ter sido usada para reescrever resultados em linguagem mais clara, mas não para calcular métricas. Pode ter sido usada para detectar possíveis causalidades indevidas, mas não para decidir conclusões. Essa separação é importante para evitar opacidade metodológica. O leitor precisa saber quando está diante de uma decisão algoritmicamente assistida e como ela foi validada.

Um protocolo de governança de *prompts* para redes científicas pode conter cinco campos. O primeiro é finalidade: porque o *prompt* foi usado. O segundo é entrada: quais dados, tabelas, descrições ou mapas foram fornecidos. O terceiro é instrução: qual foi o comando completo, incluindo restrições e formato de saída. O quarto é saída: qual resposta foi obtida. O quinto é validação: quais partes foram aceitas, rejeitadas ou modificadas pelo pesquisador. Essa estrutura transforma o uso de IA em processo auditável e reduz o risco de incorporação silenciosa de alucinações.

A reprodutibilidade não exige que todos os leitores obtenham respostas idênticas de um modelo generativo; exige que compreendam o procedimento analítico e possam avaliar sua razoabilidade. Em cientometria, muitas etapas já envolvem escolhas interpretativas: limiares, normalização, rotulagem de *clusters*, seleção de visualizações. A IA adiciona uma camada conversacional a essas escolhas. Documentá-la é uma forma de tornar visível o trabalho interpretativo que, muitas vezes, ficaria implícito. Assim, a governança de *prompts* não é burocracia; é parte da integridade científica em estudos assistidos por inteligência artificial.

11.16 Validação humana de clusters, rótulos e narrativas geradas por IA

A validação humana é indispensável em redes científicas porque os algoritmos produzem agrupamentos e a IA produz linguagem, mas o significado científico depende de julgamento situado. Um *cluster* de documentos pode ser estatisticamente coerente e,

ainda assim, receber um rótulo inadequado. Um conjunto de palavras-chave pode parecer temático, mas reunir termos de natureza distinta. Uma comunidade de coautoria pode parecer grupo de pesquisa, mas resultar de consórcio temporário. Uma narrativa gerada por IA pode ser formalmente elegante, mas ignorar debates disciplinares, história institucional ou desigualdades de campo. Portanto, a validação deve ser concebida como etapa planejada, e não como revisão superficial ao final.

Uma primeira estratégia de validação é a leitura amostral. Para cada *cluster*, selecionam-se documentos centrais, documentos periféricos e documentos ponte. A leitura de títulos, resumos e, quando necessário, textos completos permite verificar se o rótulo sugerido realmente corresponde ao conteúdo. A IA pode ajudar a criar uma ficha de leitura, mas o julgamento deve ser humano. A ficha pode incluir problema de pesquisa, método, objeto, base teórica, principais conceitos e relação com o *cluster*. Se os documentos centrais não convergem com o rótulo, o *cluster* deve ser renomeado ou descrito como heterogêneo.

Uma segunda estratégia é a validação por especialistas. Pesquisadores da área podem avaliar se os *clusters* fazem sentido, se os rótulos são reconhecíveis, se autores centrais foram interpretados corretamente e se ausências importantes decorrem de falha de busca. A IA pode sintetizar comentários de especialistas e comparar versões de rótulos, mas não substitui o conhecimento tácito do campo. Em áreas interdisciplinares, é recomendável consultar especialistas de mais de uma tradição, pois o mesmo *cluster* pode ser lido de formas diferentes. Essa pluralidade ajuda a evitar que a IA imponha uma classificação única e excessivamente ordenada.

Uma terceira estratégia é a comparação temporal. *Clusters* agregados podem esconder mudanças internas. Um tema pode ter significado diferente em 2010 e em 2025. Um termo pode migrar de uma comunidade para outra. Uma tecnologia pode inicialmente aparecer como método e depois como objeto de reflexão ética. A IA pode ajudar a descrever essas transições se receber dados por janela temporal. A validação temporal impede que o pesquisador trate como estrutura fixa aquilo que é processo histórico.

Uma quarta estratégia é a verificação cruzada com fontes externas. Políticas públicas, programas de financiamento, eventos científicos, chamadas especiais de periódicos e crises sociais podem explicar alterações de rede, mas tais explicações não estão contidas automaticamente no grafo. Se a IA sugerir que um *cluster* emergiu por causa de uma política ou evento, o pesquisador deve buscar documentos externos que sustentem essa relação. Sem essa verificação, a narrativa permanece hipótese. A função da IA deve ser indicar quais evidências procurar, não declarar que elas existem.

A validação de narrativas também inclui revisão de linguagem. Termos como 'dominante', 'periférico', 'líder', 'influente', 'isolado' e 'emergente' devem ser usados com definição operacional. Um *cluster* grande pode ser dominante em tamanho, mas não em impacto; um nó periférico pode ser emergente; uma instituição isolada na base pode ser central em redes locais não indexadas. A IA pode revisar esses termos e sugerir substituições mais precisas. Esse processo de edição conceitual é parte da qualidade cientométrica do texto.

11.17 Boas práticas para figuras, legendas e comunicação visual de redes

A comunicação visual é uma das fortalezas da análise de redes, mas também uma fonte de mal-entendidos. Figuras de rede condensam grande quantidade de informação e, por isso, precisam de legendas tecnicamente completas. Uma legenda adequada deve informar o tipo de rede, a unidade de análise, a fonte dos dados, o período, o critério de inclusão dos nós, o significado das arestas, o significado do tamanho dos nós, o significado das cores, o método

de normalização, o algoritmo de *cluster*ização e o layout utilizado. Sem esses elementos, o mapa pode ser visualmente atraente, mas metodologicamente opaco.

A IA pode auxiliar na redação de legendas, mas somente se receber os parâmetros técnicos. Um *prompt* que pede 'crie uma legenda para este mapa' tende a gerar frases genéricas. Um *prompt* adequado deve fornecer: rede de coocorrência de palavras-chave; *corpus* de 1.200 artigos; período 2015-2025; base Scopus; nós com pelo menos 10 ocorrências; arestas por coocorrência em documentos; normalização por associação; *clusters* pelo algoritmo de VOSviewer; tamanho do nó proporcional a frequência; cor indicando *cluster*; espessura da aresta indicando força de coocorrência. Com esses dados, a IA pode produzir legenda precisa e texto interpretativo adequado.

As figuras também devem evitar excesso visual. Redes muito densas podem se tornar ilegíveis e induzir interpretações arbitrárias. Em alguns casos, é melhor apresentar sub-redes, tabelas complementares, mapas por período ou visualizações interativas. A IA pode sugerir alternativas de visualização a partir do objetivo comunicacional. Se o público é técnico, pode-se mostrar métricas detalhadas. Se o público é gestor, pode-se destacar padrões principais e limitações. Se o objetivo é publicação acadêmica, a figura deve equilibrar legibilidade e informação suficiente para reprodutibilidade.

A legenda deve evitar afirmar mais do que a figura mostra. Uma legenda não deve dizer 'rede demonstra a liderança da instituição X' se o mapa mostra apenas alto grau ponderado. Deve dizer 'a instituição X apresenta o maior grau ponderado na rede de copublicação institucional'. A discussão pode explorar possíveis implicações, mas a legenda deve permanecer descritiva. A IA deve ser instruída a escrever legendas descritivas e separar comentário interpretativo em parágrafo próprio. Essa separação melhora a clareza e reduz risco de causalidade visual.

Por fim, a comunicação visual deve ser acessível e ética. Cores devem ser distinguíveis, rótulos devem ser legíveis, figuras devem ter explicação textual, e mapas geográficos devem evitar reforçar hierarquias sem contexto. Quando a rede envolve indivíduos, deve-se considerar privacidade, exposição e uso institucional dos resultados. A IA pode ajudar a revisar acessibilidade e linguagem, mas a responsabilidade final é do autor. Um mapa cientométrico é uma representação parcial de relações documentadas, não um veredito sobre valor científico.

11.18 Integração entre análise quantitativa de redes e leitura qualitativa assistida por IA

A análise relacional em cientometria atinge maior maturidade quando combina medidas quantitativas com leitura qualitativa orientada. Redes mostram padrões de conexão, mas não substituem a leitura dos documentos que compõem esses padrões. Um *cluster* de cocitação pode indicar uma base intelectual; a leitura de seus documentos centrais permite identificar se essa base é teórica, metodológica, normativa ou histórica. Uma rede de acoplamento pode indicar uma frente atual; a leitura dos artigos permite saber se a frente compartilha problemas, métodos, objetos ou apenas referências gerais. Uma rede de co-palavras pode indicar tema; a leitura permite verificar se os termos são usados com o mesmo significado. A IA pode apoiar essa ponte entre estrutura e conteúdo ao resumir documentos, extrair conceitos, comparar resumos e organizar fichas, desde que o pesquisador controle as entradas e valide as saídas.

Um procedimento recomendado é selecionar documentos por posição estrutural. Em vez de ler aleatoriamente, o pesquisador pode escolher documentos de maior centralidade, documentos representativ*os* de cada *cluster*, documentos ponte e documentos periféricos que parecem emergentes. Para cada documento, pode solicitar a IA que produza uma ficha padronizada: objetivo, problema, método, base teórica, resultados, termos-chave, relação com

o *cluster* e possível papel na rede. Essa ficha não deve ser aceita sem verificação; ela é um instrumento de triagem. Ao comparar fichas de vários documentos, o pesquisador consegue avaliar se a estrutura relacional corresponde à coerência substantiva. Esse uso da IA reduz tempo de organização, mas não elimina leitura crítica.

A integração entre rede e texto também ajuda a revisar rótulos. Suponha que um *cluster* de co-palavras seja inicialmente rotulado como 'aprendizagem de máquina em saúde'. A leitura assistida por IA pode mostrar que metade dos documentos trata de predição clínica, parte discute governança de dados e outra parte aborda vigilância epidemiológica. O rótulo inicial pode então ser dividido, refinado ou substituído por 'modelagem preditiva e governança de dados em saúde pública'. Esse refinamento ilustra que rótulos não devem ser meras etiquetas estéticas; eles são hipóteses sobre a estrutura cognitiva do campo. A IA pode gerar possibilidades, mas a boa cientometria exige que os rótulos sejam sustentados por documentos.

Outra aplicação é a construção de narrativas históricas controladas. Redes temporais indicam mudanças na conectividade, mas a explicação dessas mudanças depende de contexto. A IA pode ajudar a sintetizar documentos de diferentes períodos, identificar mudanças terminológicas e comparar agendas. Contudo, deve ser instruída a separar transformação observada de causa histórica. Por exemplo, se um termo cresce após determinado ano, o modelo pode sugerir eventos possíveis, mas deve indicar evidências necessárias: publicações seminais, chamadas de financiamento, crises sociais, mudanças regulatórias ou desenvolvimento tecnológico. Esse procedimento evita que a narrativa histórica se torne ficção retrospectiva.

A leitura qualitativa assistida por IA também contribui para identificar ausências. Um mapa pode mostrar os temas presentes, mas a leitura revela o que não aparece: populações negligenciadas, regiões pouco estudadas, métodos ausentes, debates éticos marginalizados, conceitos usados de modo superficial. *Prompts* podem pedir que a IA identifique lacunas a partir de amostras de resumos, mas tais lacunas devem ser tratadas como indícios. A ausência em um *corpus* pode resultar da estratégia de busca, da base utilizada ou dos termos escolhidos. Por isso, qualquer lacuna deve ser acompanhada da pergunta: ela é lacuna do campo ou lacuna do *corpus*? Essa distinção é fundamental em estudos cientométricos.

Por fim, a integração qualiquantitativa reforça o papel da IA como mediadora metodológica. Ela não deve apenas produzir texto final, mas apoiar ciclos de análise: rede, seleção de documentos, resumo, comparação, revisão de rótulos, auditoria de conclusões e retorno aos dados. Esse ciclo é iterativo. Uma interpretação inicial pode levar a nova limpeza de termos, a construção de sub-redes ou a comparação temporal. A IA pode acompanhar essa iteração, desde que cada passo seja registrado. Assim, a análise relacional deixa de ser procedimento linear e se torna processo reflexivo, no qual mapas, textos, métricas e *prompts* se corrigem mutuamente sob responsabilidade humana.

11.19 Conclusão

As redes científicas constituem uma das dimensões mais ricas da cientometria porque permitem analisar a ciência como sistema de relações. Elas mostram que conhecimento, colaboração, influência, circulação e visibilidade não se distribuem apenas em unidades isoladas, mas em estruturas conectadas. Redes de coautoria revelam padrões de copublicação; redes de cocitação iluminam bases intelectuais; redes de acoplamento bibliográfico indicam frentes que compartilham referências; redes de coocorrência de palavras-chave ajudam a mapear vocabulários e temas; redes institucionais e geográficas evidenciam circuitos de

colaboração e desigualdades de conectividade. A análise dessas redes amplia a capacidade de compreender campos científicos, mas exige rigor.

A inteligência artificial, quando operada por *prompts* bem desenhados, pode apoiar esse rigor. Ela pode organizar descrições, comparar tipos de rede, sugerir rótulos, formular hipóteses, revisar texto e auditar causalidades indevidas. Entretanto, sua utilidade depende do desenho do *prompt*, da qualidade dos dados e da supervisão humana. Modelos generativos não sabem, por si mesmos, se uma proximidade visual e teoricamente relevante, se um *cluster* corresponde a uma escola intelectual, se uma parceria é substantiva ou se uma mudança temporal decorre de política pública. Eles produzem linguagem plausível. Cabe ao pesquisador transformar essa linguagem em instrumento controlado de análise.

O princípio central deste capítulo é que redes científicas devem ser interpretadas em camadas: descrição estrutural, inferência plausível, hipótese interpretativa, evidência necessária e risco de erro. Essa arquitetura impede que mapas sejam convertidos em histórias deterministas. Também ajuda a comunicar resultados de forma mais responsável, especialmente quando redes são usadas em avaliação institucional e política científica. Em um contexto de crescente automatização analítica, a competência essencial não é apenas saber gerar redes, mas saber perguntar a elas o que podem e o que não podem responder.

A cientometria com IA deve, portanto, combinar capacidade computacional, conhecimento bibliométrico, ciência das redes, teoria social da ciência e ética da interpretação. *Prompts* não substituem métodos; eles os tornam mais dialógicos, documentáveis e auditáveis quando bem usados. A promessa deste capítulo é mostrar que a IA pode enriquecer a análise relacional sem usurpar a responsabilidade científica. Uma rede bem interpretada não é aquela que produz a narrativa mais impressionante, mas aquela que respeita os limites dos dados e torna explícitas as condições de suas inferências.

Posicionamento do capítulo no percurso do livro

No percurso da obra, este Capítulo ocupa uma posição de transição entre indicadores e mapas temáticos. Os Capítulos anteriores apresentam fundamentos, coleta, limpeza e indicadores; os Capítulos seguintes aprofundam estrutura intelectual, temporalidade, visualização e aplicações. As redes científicas conectam esses elementos porque transformam metadados limpos e indicadores isolados em estruturas relacionais. Sem coleta cuidadosa, as redes serão frágeis. Sem limpeza, seus nós e arestas serão distorcidos. Sem indicadores, suas propriedades serão descritas de modo impressionista. Sem interpretação, seus mapas serão apenas imagens. Sem consciência epistemológica, suas narrativas poderão converter associações em causalidades. Por isso, a análise relacional é uma espécie de eixo metodológico que articula o livro inteiro.

A perspectiva do livro - Cientometria com Inteligência Artificial: *Prompts* para Analisar, Mapear e Interpretar a ciência - exige que redes sejam ensinadas não apenas como técnicas de software, mas como objetos de diálogo metodológico. O pesquisador precisa aprender a perguntar a uma rede o que ela pode responder. Precisa também aprender a perguntar à IA o que ela pode apoiar. Uma rede de coautoria pode responder sobre copublicação, mas não sobre toda a cooperação intelectual. Uma rede de cocitação pode responder sobre uso conjunto de referências, mas não sobre causalidade teórica direta. Uma rede de co-palavras pode responder sobre associação terminológica, mas não sobre significado profundo sem leitura textual. A IA pode ajudar a lembrar esses limites quando o *prompt* a obriga a fazê-lo.

O valor pedagógico deste capítulo está em formar uma competência dupla. A primeira é técnica: compreender tipos de redes, métricas, *clusters*, parâmetros e visualizações. A

segunda é interpretativa: reconhecer o estatuto das inferências, formular hipóteses prudentes, validar rótulos, comunicar incertezas e impedir exageros. Em muitos estudos cientométricos, o problema não está na produção do mapa, mas na passagem do mapa para o argumento. Essa passagem é exatamente o ponto em que a IA pode ser tanto aliada quanto risco. Aliada, quando organiza raciocínio, explicita lacunas e revisa causalidades. Risco, quando oferece narrativas fluentes que parecem substituir a validação. O Capítulo, portanto, ensina uma prática de uso crítico da IA: pedir interpretações, mas também pedir limites; pedir hipóteses, mas também pedir evidências; pedir redação, mas também pedir auditoria.

Ao final, o pesquisador deve estar preparado para construir e interpretar redes científicas de modo mais consciente. Deve saber que um mapa é resultado de decisões; que uma comunidade é produto de algoritmo e dados; que centralidade não é sinônimo automático de qualidade; que proximidade visual não é causalidade; que bases diferentes podem produzir mundos diferentes; que *prompts* precisam ser documentados; e que a interpretação cientométrica é sempre uma atividade situada. Essa formação é essencial para que a IA não seja incorporada como atalho retórico, mas como instrumento de ampliação da reflexividade metodológica. A análise relacional assistida por *prompts* será tanto melhor quanto mais claro for o compromisso do pesquisador com transparência, verificação e prudência epistêmica.

Síntese operacional

Em termos operacionais, uma análise relacional cientométrica assistida por IA deve iniciar com uma pergunta bem delimitada, seguir com *corpus* rastreável, aplicar limpeza documentada, escolher o tipo de rede conforme a relação investigada, calcular métricas coerentes, interpretar *clusters* com validação e redigir resultados com linguagem prudente. A IA pode participar de todas essas etapas como assistente de planejamento, classificação, revisão e comunicação, mas não deve assumir a função de juiz metodológico. O pesquisador deve manter controle sobre dados, parâmetros, inferências e conclusões.

A regra prática mais importante é simples: toda afirmação interpretativa deve poder ser reconduzida a uma evidência. Se a evidência é a métrica, a afirmação deve permanecer estrutural. Se a evidência inclui leitura documental, a afirmação pode avançar para interpretação substantiva. Se a evidência inclui dados externos e desenho comparativo, a afirmação pode se aproximar de explicação causal. Sem essa cadeia, a conclusão deve ser formulada como hipótese. Essa disciplina é o fundamento da cientometria relacional responsável em tempos de inteligência artificial generativa.

Quadro 11.6 - Checklist para análise relacional assistida por IA

Etapa	Pergunta de controle	Evidência/documentação esperada
Corpus	A estratégia de busca corresponde à pergunta relacional?	String, bases, período, critérios de inclusão e exclusão.
Limpeza	Autores, instituições, documentos e termos foram normalizados?	Regras de desambiguação, dicionários e amostras auditadas.
Tipo de rede	A relação escolhida responde à pergunta?	Justificativa para coautoria, cocitação, acoplamento, co-palavras ou rede institucional.
Parâmetros	Os limiares e normalizações foram explicitados?	Limiar de ocorrência, método de normalização, contagem inteira ou fracionada.

Etapa	Pergunta de controle	Evidência/documentação esperada
Métricas	As métricas foram interpretadas no contexto correto?	Definição de grau, intermediação, modularidade, *clusters* e pesos.
IA	Os *prompts* e saídas relevantes foram preservados?	Registro dos *prompts*, versão do modelo, dados fornecidos e revisão humana.
Causalidade	Conclusões causais foram evitadas ou justificadas?	Separação entre descrição, inferência, hipótese e evidência externa.
Comunicação	O relatório apresenta limites e riscos?	Seção de limitações, matriz de riscos e linguagem prudente.

Fonte: Elaborado pelo autor (2026).

Capítulo 12

Mapas temáticos e estrutura conceitual da ciência

12.1 Introdução

Os mapas temáticos ocupam uma posição estratégica nos estudos cientométricos porque permitem transitar da contagem de documentos para a interpretação da organização cognitiva de uma área. Em vez de perguntar apenas quantos artigos foram publicados, quantas citações foram recebidas ou quais autores aparecem com maior frequência, os mapas temáticos procuram representar como os assuntos se articulam, quais conceitos estruturam o campo, quais linhas de investigação se consolidam, quais frentes emergem e quais combinações terminológicas indicam mudanças no vocabulário científico. No contexto da trilha deste livro, dedicada à cientometria com inteligência artificial e *prompts*, esse deslocamento é decisivo: a IA pode apoiar a identificação, a nomeação e a narrativa de padrões temáticos, mas tais padrões continuam dependendo de decisões metodológicas sobre coleta, pré-processamento, unidade de análise, medida de associação, algoritmo de agrupamento e validação especializada.

A expressão estrutura conceitual da ciência refere-se ao conjunto de relações conceituais, teóricas, metodológicas e documentais que sustentam a produção de conhecimento em uma comunidade científica. Essa estrutura pode ser observada por meio de citações, cocitações, acoplamento bibliográfico, coocorrência de palavras-chave, termos extraídos de títulos e resumos, classificações disciplinares, trajetórias de autores, periódicos e instituições. Ela não é idêntica ao organograma administrativo da ciência nem se confunde com a distribuição formal de departamentos, programas ou áreas de avaliação. Trata-se, antes, de uma arquitetura cognitiva e socialmente produzida, na qual conceitos, problemas, métodos e comunidades se estabilizam temporariamente e se reorganizam diante de novas evidências, tecnologias, políticas de financiamento e controvérsias.

Os mapas temáticos são particularmente relevantes porque situam o pesquisador diante de uma dupla exigência: representar a complexidade e, ao mesmo tempo, simplificá-la de modo interpretável. Toda visualização cientométrica reduz dimensões; todo *cluster* temático agrega documentos heterogêneos; todo rótulo sintetiza um conjunto de termos ou publicações; todo quadrante de um diagrama estratégico transforma relações contínuas em categorias discretas. A utilidade do mapa depende justamente da consciência dessa redução. Um mapa não é uma fotografia neutra da ciência, mas uma construção analítica baseada em escolhas explícitas. Por isso, este capítulo insiste na diferença entre visualizar temas e concluir automaticamente que eles constituem escolas, paradigmas, especialidades ou consensos.

A incorporação de modelos de linguagem e *prompts* amplia as possibilidades de análise temática, principalmente em quatro frentes. Em primeiro lugar, a IA pode auxiliar na leitura de *clusters*, oferecendo interpretações preliminares sobre termos, documentos e relações. Em segundo lugar, pode sugerir rótulos alternativos para comunidades temáticas, explicitando graus de confiança e ambiguidades. Em terceiro lugar, pode apoiar a construção de narrativas sobre a estrutura conceitual de uma área, conectando mapas, indicadores e literatura substantiva. Em quarto lugar, pode ajudar a produzir instrumentos de validação, como checklists, matrizes de comparação e protocolos de revisão por especialistas. Entretanto, esses benefícios só

são cientificamente aceitáveis quando acompanhados por rastreabilidade, documentação e controle dos riscos de alucinação, viés e falsa coerência.

A tradição do mapeamento científico fornece a base metodológica desse debate (GARFIELD, 1955; KESSLER, 1963; SMALL, 1973; CALLON; COURTIAL; LAVILLE, 1991; BÖRNER; CHEN; BOYACK, 2003; CHEN, 2006; VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017). A cocitação, proposta por Small, permitiu observar a proximidade intelectual entre documentos citados conjuntamente; o acoplamento bibliográfico, formulado por Kessler, mostrou a aproximação entre documentos que compartilham referências; a análise de co-palavras, desenvolvida em estudos como o de Callon, Courtial e Laville, abriu caminho para investigar dinâmicas conceituais por meio de termos; ferramentas como VOSviewer, CiteSpace, SciMAT e bibliometrix difundiram procedimentos para extração, normalização, visualização e análise de redes bibliométricas. A contribuição contemporânea da IA não substitui essa tradição; ela a reorganiza ao introduzir uma camada de mediação textual e interpretativa baseada em instruções conversacionais.

Neste capítulo, os *prompts* são tratados como instrumentos metodológicos para tornar a interpretação de mapas temáticos mais explícita e auditável. Um *prompt* para nomear *clusters* não é apenas uma pergunta ao modelo; é uma forma de delimitar o papel do analista, o tipo de dado permitido, o formato da resposta, os critérios de rotulagem, as incertezas esperadas e as condições de validação. Um *prompt* para comparar rótulos alternativos não deve simplesmente escolher o nome mais elegante; deve justificar a escolha com base em termos recorrentes, documentos centrais, medidas de densidade, centralidade e coerência semântica. Um *prompt* para construir uma narrativa da estrutura conceitual precisa separar descrição, interpretação e hipótese, evitando transformar proximidade lexical em relação causal.

A organização do capítulo segue seis movimentos. Primeiro, discutem-se os temas motores, básicos, emergentes e declinantes, com atenção ao uso de diagramas estratégicos. Segundo, examinam-se as frentes de pesquisa e sua relação com documentos recentes, termos em aceleração e agrupamentos temporais. Terceiro, analisam-se escolas teóricas e comunidades epistêmicas, destacando a necessidade de evidências além da simples vizinhança em rede. Quarto, apresentam-se mapas conceituais produzidos a partir de títulos, resumos e palavras-chave. Quinto, discute-se o uso da IA para rotular *clusters* temáticos. Sexto, propõe-se um modelo de validação dos rótulos por especialistas. Ao longo do texto, são oferecidos *prompts* e procedimentos que podem ser incorporados a protocolos de pesquisa cientométrica.

Quadro 12.1 - Funções analíticas dos mapas temáticos

Função	Pergunta orientadora	Dados usuais	Risco interpretativo
Organizar o campo	Quais temas estruturam a literatura?	Palavras-chave, termos, citações e documentos	Confundir *cluster* algorítmico com categoria teórica natural
Detectar movimento	Que temas emergem, declinam ou se consolidam?	Séries temporais, *burst* terms, acoplamento recente	Confundir crescimento bibliográfico com inovação cognitiva
Nomear comunidades	Como descrever grupos de documentos ou termos?	Termos centrais, documentos representativos e autores	Aceitar rótulos automáticos sem validação
Construir narrativa	Como os temas se conectam na estrutura conceitual?	Mapas, métricas, literatura substantiva e contexto	Transformar associação em causalidade ou hierarquia

Fonte: Elaborado pelo autor (2026).

12.2 Temas motores, básicos, emergentes e declinantes

A classificação de temas em motores, básicos, nicho, emergentes e declinantes deriva, em grande parte, da tradição dos diagramas estratégicos de análise temática (COBO *et al.*, 2011; COBO *et al.*, 2012). Nessa abordagem, cada tema ou *cluster* é posicionado em função de duas dimensões interpretativas: centralidade e densidade. A centralidade indica o grau de conexão do tema com outros temas do campo; a densidade expressa o grau de desenvolvimento interno do próprio tema. Quando essas dimensões são combinadas, surgem quadrantes que ajudam a interpretar a função relativa de cada tema na estrutura do domínio analisado. Trata-se de uma simplificação útil, mas não automática: centralidade e densidade dependem da rede construída, da medida de associação, dos parâmetros de agrupamento e do recorte temporal adotado.

Temas motores são aqueles que apresentam alta centralidade e alta densidade. Em termos analíticos, são temas bem desenvolvidos internamente e fortemente conectados ao restante do campo. Em um mapa temático, eles costumam representar áreas de investigação com vocabulário consolidado, massa crítica de publicações, relações estáveis com subtemas e capacidade de organizar agendas adjacentes. No entanto, a expressão tema motor não deve ser entendida como sinônimo de tema mais importante em sentido absoluto. Ela indica uma posição relacional dentro do conjunto de dados e do período analisado. Um tema pode ser motor em uma base de dados e periférico em outra; pode ser central em um recorte disciplinar e marginal em um recorte interdisciplinar.

Temas básicos, por sua vez, tendem a apresentar alta centralidade e baixa densidade. Eles se conectam a muitos outros temas, mas ainda não formam uma comunidade interna tão coesa. Essa condição pode indicar conceitos transversais, métodos compartilhados, problemas amplos ou vocabulários genéricos usados por múltiplas subáreas. Em alguns casos, um tema básico corresponde a uma infraestrutura conceitual indispensável; em outros, pode refletir excesso de generalidade terminológica. Por essa razão, a interpretação precisa verificar se os termos do *cluster* são substantivos ou apenas expressões amplas, se os documentos centrais compartilham de fato uma agenda, e se a baixa densidade decorre de dispersão conceitual ou de uma etapa inicial de desenvolvimento.

Temas emergentes e declinantes ocupam, tradicionalmente, o quadrante de baixa centralidade e baixa densidade. Essa posição é ambígua por definição. Um *cluster* pouco conectado e internamente pouco desenvolvido pode representar um tema novo que ainda não se articulou ao núcleo do campo; mas também pode representar uma linha de investigação em declínio, um tópico residual, um artefato de busca ou um conjunto de termos pouco coerente. A distinção entre emergência e declínio exige análise temporal, comparação entre períodos e leitura substantiva. A IA pode ajudar a levantar hipóteses, mas não deve concluir emergência apenas porque um *cluster* aparece no quadrante inferior esquerdo.

Há ainda os temas de nicho, que são altamente desenvolvidos e isolados, frequentemente associados à combinação de alta densidade e baixa centralidade. Eles possuem forte coesão interna, mas pouca conexão com os demais temas do campo. Podem representar especialidades autônomas, nichos metodológicos, comunidades fechadas, tradições teóricas específicas ou subdomínios que dialogam pouco com a área mapeada. Em estudos de política científica, tais temas podem ser interpretados como capacidades especializadas ou ilhas de excelência; em estudos epistemológicos, podem indicar diferenciação cognitiva. Todavia, o isolamento também pode resultar de escolhas metodológicas, como vocabulário restrito, fragmentação terminológica ou sub-representação de certas fontes.

O uso de IA nesse tipo de análise pode ser particularmente produtivo quando o modelo recebe, como entrada, a lista de termos do *cluster*, os documentos mais representativos, as medidas de centralidade e densidade, o período de ocorrência e o contexto do estudo. O *prompt* deve solicitar uma resposta que diferencie descrição estrutural, interpretação temática e hipótese temporal. Essa distinção é fundamental porque os quadrantes do diagrama estratégico não carregam significado substantivo por si mesmos. O rótulo tema motor, por exemplo, descreve uma posição no mapa; a interpretação de que esse tema orienta uma agenda de pesquisa depende de evidências adicionais.

Em uma análise cientométrica assistida por *prompts*, recomenda-se que cada *cluster* temático seja acompanhado por um pequeno dossiê: termo preferencial, termos associados, documentos centrais, autores ou periódicos predominantes, período de maior crescimento, posição no diagrama estratégico, indicadores de desempenho e notas de validação. Esse dossiê evita que o mapa se torne apenas uma figura ornamental. Ele transforma o *cluster* em objeto interpretável e auditável. A IA pode auxiliar na elaboração desse dossiê, desde que seja instruída a não inventar informações ausentes e a declarar quando os dados fornecidos são insuficientes.

Uma limitação recorrente dos diagramas estratégicos é a tendência de naturalizar rótulos. Ao classificar um tema como emergente, básico ou motor, o pesquisador pode induzir o leitor a imaginar uma hierarquia estável de valor científico. Essa leitura é inadequada. Um tema básico pode ser mais relevante para a sustentabilidade de uma área do que um tema motor de alta visibilidade. Um tema emergente pode jamais se consolidar. Um tema isolado pode ser extremamente inovador, mas pouco integrado ao vocabulário dominante. Portanto, a interpretação deve ser formulada em termos de função relacional e hipótese analítica, não de julgamento normativo.

A comparação entre períodos temporais fortalece a análise. Um tema que se desloca de emergente para motor pode indicar consolidação; um tema que passa de motor para básico pode sinalizar difusão conceitual; um tema que permanece isolado pode revelar especialização persistente; um tema que desaparece pode indicar substituição terminológica, declínio substantivo ou efeito de mudança na estratégia de busca. A IA pode ser usada para narrar essas transições, mas precisa receber uma tabela longitudinal e instruções claras para não atribuir causalidade a simples movimento no mapa.

Do ponto de vista pedagógico, a análise de temas motores, básicos, emergentes e declinantes é uma excelente porta de entrada para ensinar que cientometria não é sinônimo de ranking. O objetivo não é descobrir o melhor tema, mas compreender como os temas se organizam, se conectam e se transformam. Esse princípio é ainda mais importante quando *prompts* entram no processo, pois modelos de linguagem tendem a produzir narrativas coerentes mesmo quando a evidência é fraca. O *prompt* cientométrico deve, portanto, solicitar prudência interpretativa, identificação de incertezas e indicação das evidências necessárias para sustentar cada afirmação.

Quadro 12.2 - Interpretação prudente dos quadrantes do diagrama estratégico

Quadrante	Condição relacional	Leitura possível	Cautela metodológica
Tema motor	Alta centralidade e alta densidade	Tema desenvolvido e conectado ao campo	Não equivale automaticamente a tema mais importante ou melhor
Tema básico	Alta centralidade e baixa densidade	Tema transversal ou infraestrutura conceitual	Pode refletir generalidade terminológica ou heterogeneidade

Quadrante	Condição relacional	Leitura possível	Cautela metodológica
Tema emergente/ declinante	Baixa centralidade e baixa densidade	Tema novo, residual, em transformação ou com baixa coesão	Exige análise temporal para distinguir emergência e declínio
Tema de nicho	Baixa centralidade e alta densidade	Especialidade autônoma, nicho ou tradição relativamente separada	Isolamento pode decorrer de recorte da busca ou limiar de rede

Fonte: Elaborado pelo autor (2026).

A leitura de mapas temáticos precisa considerar que os *clusters* são sensíveis à escala. Em escalas amplas, um *cluster* pode corresponder a uma disciplina inteira; em escalas estreitas, pode representar uma técnica, um objeto ou uma subpergunta. A escolha da escala deve estar subordinada à pergunta de pesquisa. Um estudo sobre a estrutura de uma área nacional talvez precise de *clusters* amplos; um estudo sobre evolução de um método específico pode exigir granularidade fina. A IA pode auxiliar a comparar interpretações em diferentes escalas, mas o analista deve explicitar porque uma escala foi escolhida como adequada ao objetivo.

A normalização das relações é outro ponto relevante. Medidas de associação, como força de associação, cosseno ou índice de Jaccard, podem alterar a proximidade entre termos ou documentos. Diferentes softwares adotam padrões distintos. A escolha não é apenas técnica; ela influencia a forma do mapa e, consequentemente, a narrativa interpretativa. Em um protocolo assistido por *prompts*, a IA pode ser solicitada a explicar as implicações de cada medida para leitores não especialistas, mas os valores e decisões devem ser extraídos do software e documentados pelo pesquisador.

A visualização também tem retórica própria. Cores sugerem comunidades; tamanhos sugerem importância; distâncias sugerem proximidade; linhas sugerem relação. O leitor tende a interpretar esses elementos de maneira intuitiva, mesmo quando a legenda é incompleta. Por isso, todo mapa deve vir acompanhado de explicação sobre o que cada elemento visual representa. A IA pode ajudar a redigir legendas analíticas, mas deve ser instruída a não afirmar que distância gráfica equivale diretamente a distância teórica sem informar a técnica de layout utilizada.

Um problema comum é a sobreinterpretação de *clusters* pequenos. Em alguns mapas, *clusters* minoritários podem ser visualmente destacados por cores fortes ou posições periféricas, levando o leitor a superestimar sua relevância. Em outros casos, *clusters* pequenos podem representar sinais importantes de emergência. A decisão exige contexto. O *prompt* de interpretação deve solicitar que o modelo avalie o tamanho do *cluster*, sua coerência interna, sua trajetória temporal e sua relação com a pergunta de pesquisa antes de propor uma leitura.

12.3 Frentes de pesquisa

As frentes de pesquisa constituem uma categoria central para compreender a dinâmica da ciência (CHEN, 2006; BOYACK; KLAVANS, 2010; KLAVANS; BOYACK, 2017). De modo geral, uma frente de pesquisa pode ser entendida como um conjunto de problemas, métodos, conceitos, documentos e comunidades que se articulam em torno de questões ativas e relativamente recentes. Enquanto a estrutura conceitual remete frequentemente às bases cognitivas de uma área, as frentes de pesquisa apontam para sua zona de movimento, isto é, para os tópicos em torno dos quais se concentram esforços atuais. Em estudos de cocitação, a distinção entre base intelectual e frente de pesquisa foi discutida por Chen ao desenvolver

abordagens para detectar tendências emergentes e padrões transitórios na literatura científica (CHEN, 2006; CHEN; IBEKWE-SANJUAN; HOU, 2010).

A identificação de frentes de pesquisa pode ser realizada por diferentes caminhos. Redes de acoplamento bibliográfico tendem a capturar proximidades entre documentos recentes, pois artigos publicados no mesmo período podem compartilhar referências semelhantes antes de acumular citações. Redes de coocorrência de palavras-chave e termos de resumo podem revelar mudanças vocabulares e temas em expansão. Análises de citações recentes, *burst detection* e crescimento temporal de termos podem indicar aceleração de atenção. Nenhum desses procedimentos, isoladamente, garante que uma frente de pesquisa foi identificada. A robustez aumenta quando múltiplos sinais convergem.

Figura 12.1 - Exemplo de matriz de interpretação de frentes

Frente (tema)	Evidência quantitativa	Evidência textual	Evidência relacional	Interpretação provisoria	Verificações necessarias
IA generativa na educação		Termos: "LLM", "generative AI", "prompt", "GPT" em títulos/resumos.	Acoplamento alto entre docs (2022–2024); novos autores.	Frente emergente focada em ensino superior e automatização de tarefas.	Ler amostras; mapear métodos; comparar com financiamento.
Ética algorítmica e educação		Termos: "biased", "ethics", "fairness", "algorithmic accountability".	Cocitação crescente com autores de ética e educação; interdisciplinar.	Frente consolidando ética aplicada à IA educacional.	Ver diretrizes e políticas; checar aplicação empírica real.

Fonte: Imagem gerada por IA, ChatGPT (2026).

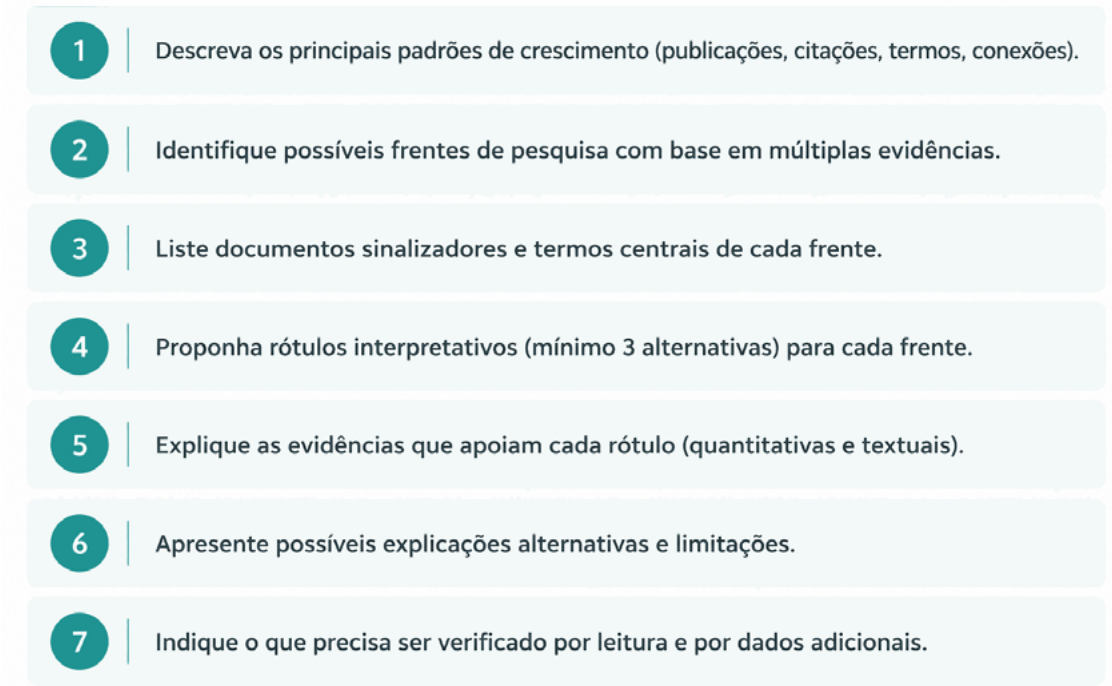
Em um estudo assistido por IA, a noção de frente de pesquisa exige especial cuidado, porque modelos de linguagem podem confundir novidade textual com relevância científica. Um conjunto de termos recentes pode parecer emergente apenas por causa de modismos terminológicos, mudanças na indexação ou expansão de uma base de dados. Da mesma forma, um tema consolidado pode reaparecer com nova nomenclatura e ser erroneamente interpretado como inovação radical. Por isso, *prompts* voltados à detecção de frentes devem solicitar a comparação entre evidências: crescimento absoluto, crescimento relativo, documentos-chave, presença em periódicos centrais, ligação com autores influentes, continuidade temporal e relação com políticas ou tecnologias externas.

A temporalidade é a chave para diferenciar frentes de pesquisa de *clusters* estáticos. Um *cluster* pode ser coeso e central, mas não necessariamente representar uma frente atual; pode ser uma tradição consolidada, uma base teórica ou um método amplamente utilizado.

Por outro lado, uma frente de pesquisa pode ser pequena e ainda assim relevante, desde que apresente crescimento, novidade conceitual e capacidade de atrair atenção. A IA pode apoiar a construção de narrativas temporais, mas deve ser instruída a apresentar evidências em camadas: o que os dados mostram, o que eles sugerem e o que precisaria ser investigado para confirmar a interpretação.

A análise de frentes também se beneficia da distinção entre emergência bibliográfica e emergência cognitiva. A emergência bibliográfica ocorre quando há aumento de publicações, citações ou termos associados a um tópico. A emergência cognitiva ocorre quando há reorganização de problemas, métodos ou conceitos de modo a produzir nova agenda intelectual. A primeira pode ser medida com relativa facilidade; a segunda exige leitura especializada. Um *prompt* bem desenhado deve reconhecer essa diferença e solicitar que o modelo não infira mudança cognitiva apenas a partir de crescimento quantitativo.

Figura 12.2 - Fluxo de instrução de IA



Fonte: Imagem gerada por IA, ChatGPT (2026).

A identificação de documentos sinalizadores é uma prática útil para interpretar frentes de pesquisa. Documentos sinalizadores são aqueles que concentram termos emergentes, conectam subcampos, recebem citações rápidas, aparecem como referências compartilhadas ou ocupam posições estruturais relevantes. Esses documentos não devem ser automaticamente tratados como marcos fundadores; eles funcionam como pontos de entrada para leitura qualitativa. A IA pode resumir esses documentos, comparar seus objetivos e extrair termos recorrentes, mas a seleção final da frente depende da consistência entre os dados e o julgamento do pesquisador.

Em áreas fortemente interdisciplinares, frentes de pesquisa podem surgir nas bordas dos mapas, entre *clusters* previamente separados. Essa condição é comum em domínios nos quais novas tecnologias, problemas sociais ou políticas públicas aproximam comunidades distintas. O mapa temático, nesse caso, deve ser lido não apenas por seus *clusters*, mas também por suas zonas de contato. Termos ponte, documentos de fronteira e autores que transitam entre comunidades podem revelar movimentos de recombinação. A IA pode ajudar

a identificar esses elementos, desde que o *prompt* solicite explicitamente a análise de relações entre *clusters*, e não apenas a descrição isolada de cada grupo.

A relação entre frente de pesquisa e agenda de financiamento também merece atenção. Em muitos campos, o surgimento de termos emergentes acompanha chamadas públicas, programas internacionais, prioridades tecnológicas ou crises sociais. Contudo, a presença de financiamento não implica automaticamente maturidade científica, e a visibilidade política de um tema pode produzir crescimento bibliográfico sem estabilidade conceitual. O analista deve, portanto, evitar narrativas lineares do tipo financiamento gera frente ou frente gera impacto. O procedimento mais adequado é formular hipóteses contextuais e submetê-las à verificação por dados adicionais.

Quando a IA é utilizada para apoiar a detecção de frentes, recomenda-se que o *prompt* exija uma matriz com cinco colunas: evidência quantitativa, evidência textual, evidência relacional, interpretação provisória e verificação necessária. Essa estrutura obriga o modelo a separar os níveis de inferência e reduz o risco de produzir um texto excessivamente assertivo. Além disso, convém solicitar que o modelo liste possíveis explicações alternativas para o mesmo padrão, como mudança terminológica, expansão da base, ruído de indexação ou influência de um periódico temático.

Em síntese, frentes de pesquisa são objetos analíticos que emergem da combinação entre movimento temporal, densidade temática, conectividade relacional e interpretação substantiva. Mapas temáticos ajudam a localizá-las, mas não as definem sozinhos. A IA pode ampliar a capacidade de exploração, leitura e síntese, mas também pode gerar narrativas de novidade sem base empírica suficiente. A solução metodológica consiste em transformar cada afirmação sobre emergência em uma proposição verificável, associada a evidências, limites e grau de confiança.

Outro risco é a invisibilidade de temas ausentes. Mapas representam apenas o que entrou na base, passou pela busca, foi indexado, sobreviveu ao pré-processamento e atingiu limiar de frequência ou conexão. Ausências podem ser metodológicas, linguísticas ou políticas. Um tema local, publicado em língua portuguesa ou em periódicos não indexados, pode desaparecer do mapa internacional. A IA não corrige automaticamente essa ausência; no máximo, pode ajudar a formular perguntas sobre possíveis lacunas, desde que o analista forneça contexto.

A construção de narrativa deve evitar teleologia. É tentador descrever a evolução de uma área como trajetória linear de maturação, na qual temas emergentes se tornam motores e depois se estabilizam. Embora isso possa ocorrer, a ciência também apresenta rupturas, abandonos, ciclos de atenção, apropriações e renomeações. A IA tende a produzir narrativas ordenadas; o *prompt* deve solicitar explicitamente alternativas não lineares e explicitar quando a evolução observada pode ser resultado de artefatos de dados.

A comparação entre bases é indispensável em áreas multidisciplinares. Web of Science, Scopus, PubMed, Dimensions, OpenAlex e bases nacionais têm coberturas distintas. Um mapa temático baseado em uma única fonte pode privilegiar certas comunidades. A IA pode ajudar a comparar mapas derivados de fontes diferentes, mas não deve harmonizar conclusões sem considerar diferenças de cobertura, campos de metadados e critérios de indexação. Quando possível, o relatório deve incluir uma seção de sensibilidade sobre base de dados.

Quadro 12.3 - Checklist de qualidade

CHECKLIST DE QUALIDADE *(antes de fechar o estudo relacional assistido por IA)*
☐ Pergunta de pesquisa clara e alinhada ao método.
☐ *Corpus* bem definido, reproduzível e documentado.
☐ Construção da matriz relacional transparente.
☐ Parâmetros, limiares e algoritmos informados.
☐ Análises de sensibilidade realizadas.
☐ Interpretações separadas de evidências.
☐ Limitações e vieses explicitados.
☐ Documentação dos *prompts* e das saídas da IA.
☐ Revisão humana e validação qualitativa.
☐ Conclusões proporcionais às evidências.

Registro da verificação

Estudo ou etapa verificada	Responsável	Data

Fonte: Elaborado pelo autor (2026).

12.4 Escolas teóricas e comunidades epistêmicas

Escolas teóricas e comunidades epistêmicas são categorias mais exigentes do que *clusters* temáticos (HAAS, 1992). Um *cluster* pode reunir documentos que compartilham palavras, referências ou relações de citação; uma escola teórica pressupõe algum grau de continuidade conceitual, tradição interpretativa, repertório metodológico, autores de referência e mecanismos de transmissão intelectual. Uma comunidade epistêmica, por sua vez, envolve atores que compartilham crenças causais, critérios de validade, vocabulário especializado e formas de intervenção em problemas. Portanto, a passagem de *cluster* para escola ou comunidade não é automática. Ela requer evidências bibliométricas, textuais e qualitativas.

A cientometria oferece indícios para identificar escolas teóricas. Redes de cocitação podem revelar autores ou obras frequentemente mobilizados em conjunto; acoplamento bibliográfico pode indicar grupos de documentos que partem de bases semelhantes; coocorrência de termos pode mostrar vocabulários compartilhados; redes de coautoria podem evidenciar colaboração entre pesquisadores; análise de periódicos pode indicar espaços de circulação. Esses indícios, contudo, são insuficientes sem leitura substantiva. Dois autores podem ser cocitados por oposição teórica, por comparação metodológica ou por pertencimento a tradições diferentes. A IA, se não for orientada, pode interpretar toda cocitação como afinidade.

O estudo de escolas teóricas demanda atenção à polissemia dos termos. Uma expressão pode ter sentidos distintos em comunidades diferentes; um método pode ser apropriado por tradições

teóricas divergentes; um conceito pode circular de modo superficial em áreas próximas. Mapas conceituais baseados apenas em palavras podem ocultar essas diferenças. Por isso, a IA deve ser utilizada para levantar possibilidades interpretativas, não para decidir sozinha o pertencimento teórico. Um *prompt* adequado deve solicitar que o modelo identifique usos possíveis do termo, documentos que sustentam cada uso e ambiguidades que exigem leitura humana.

Comunidades epistêmicas também se manifestam em padrões institucionais e geográficos. Um tema pode estar associado a redes de pesquisa concentradas em certos países, laboratórios ou programas; outro pode circular por meio de consórcios internacionais, conferências e projetos colaborativos. A análise de afiliações e coautorias pode complementar o mapa temático ao mostrar onde os conceitos são produzidos, difundidos e apropriados. No entanto, a proximidade institucional não garante unidade intelectual. A unidade deve ser demonstrada por convergência entre produção, referências, vocabulário e práticas de pesquisa.

A IA pode contribuir para a análise de escolas e comunidades ao comparar textos representativos. Dado um conjunto de resumos, títulos e palavras-chave, o modelo pode sugerir quais documentos compartilham premissas, métodos ou problemas. Dado um conjunto de referências centrais, pode indicar tradições teóricas prováveis. Dado um *cluster* de termos, pode propor hipóteses sobre sua orientação conceitual. Mas todas essas operações devem ser acompanhadas de solicitação de incerteza. O modelo precisa declarar quando a evidência é lexical, quando é bibliográfica e quando é apenas inferencial.

Uma armadilha frequente é confundir comunidade algorítmica com comunidade sociológica. Algoritmos de *cluster*ização identificam grupos segundo critérios matemáticos de similaridade ou modularidade. Tais grupos podem coincidir com comunidades de pesquisa, mas também podem reunir documentos por propriedades formais, como palavras comuns, referências metodológicas ou periódicos compartilhados. Uma comunidade sociológica, por outro lado, envolve reconhecimento mútuo, circulação de pesquisadores, espaços de debate e continuidade institucional. O mapa fornece pistas, não prova. A IA deve ser instruída a preservar essa diferença.

Na prática, recomenda-se que a investigação de escolas teóricas siga um protocolo em três níveis. O primeiro nível é relacional: quais documentos, autores, termos e referências aparecem conectados? O segundo é textual: que conceitos, problemas e métodos aparecem nos títulos, resumos e palavras-chave? O terceiro é contextual: há periódicos, conferências, instituições, países ou redes de colaboração associados ao grupo? A IA pode preencher uma matriz preliminar desses níveis, mas a validação requer leitura de documentos centrais e consulta a especialistas.

A nomeação de uma escola ou comunidade deve ser ainda mais cautelosa do que a nomeação de um *cluster* temático. Rótulos como escola crítica, abordagem evolucionária, paradigma sociotécnico ou comunidade de aprendizagem de máquina carregam implicações fortes. Quando utilizados sem evidência, podem impor ao campo uma interpretação externa. Por isso, é preferível usar rótulos descritivos em fases exploratórias, como *cluster* de governança algorítmica, literatura sobre redes de inovação ou estudos de avaliação responsável. Rótulos teóricos mais fortes devem aparecer somente quando o conjunto de evidências sustenta tal leitura.

A IA pode auxiliar a comparar rótulos fortes e fracos. Um *prompt* pode pedir três alternativas: um rótulo descritivo, um rótulo conceitual e um rótulo teórico. Para cada alternativa, o modelo deve indicar vantagens, riscos e evidências necessárias. Essa estratégia é útil porque evita que a primeira sugestão do modelo seja aceita como denominação definitiva. Além disso, permite documentar o processo de escolha do rótulo, o que aumenta a transparência da análise.

Em síntese, mapas temáticos podem iluminar escolas teóricas e comunidades epistêmicas, mas não as substituem. A identificação dessas formações exige triangulação

entre redes, textos, contextos e validação especializada. A IA amplia a capacidade de organizar evidências, gerar hipóteses e produzir sínteses comparativas, mas não deve transformar similaridade algorítmica em pertencimento intelectual. O analista deve manter a diferença entre *cluster*, tema, frente, escola, comunidade e paradigma como princípio de rigor conceitual.

A análise temática também pode incorporar indicadores de desempenho, mas deve fazê-lo com cautela. Um *cluster* com alta média de citações não é necessariamente teóricamente mais relevante; pode ser mais antigo, internacionalizado ou associado a periódicos de maior visibilidade. Um *cluster* recente pode ter baixa citação por efeito temporal. A IA pode ajudar a redigir resultados que combinem estrutura e desempenho sem transformar impacto bibliométrico em qualidade intrínseca. Essa distinção é fundamental para evitar uso avaliativo indevido.

Em estudos aplicados à gestão de pesquisa, mapas temáticos podem orientar decisões sobre colaboração, investimento, formação de equipes e prospecção. Todavia, decisões institucionais baseadas em mapas devem considerar missão, contexto local, capacidade instalada e prioridades sociais. A IA pode produzir briefings executivos, mas o documento científico deve preservar limitações. Um *cluster* emergente não deve ser apresentado como recomendação automática de investimento; ele pode indicar oportunidade, mas a decisão exige análise estratégica mais ampla.

Por fim, a ética da interpretação exige reconhecer que rótulos podem afetar comunidades. Nomear uma área como declinante, periférica ou isolada pode ter consequências simbólicas e institucionais. Mesmo quando essas categorias são técnicas, sua recepção pode ser avaliativa. Por isso, recomenda-se linguagem prudente: tema posicionado no quadrante de baixa centralidade e baixa densidade no período analisado é mais preciso do que tema irrelevante. A IA deve ser instruída a evitar adjetivos normativos e a preservar a natureza relacional das categorias.

12.5 Mapas conceituais a partir de títulos, resumos e palavras-chave

Mapas conceituais construídos a partir de títulos, resumos e palavras-chave constituem uma das formas mais diretas de acessar o vocabulário de uma área (CALLON; COURTIAL; LAVILLE, 1991; BÖRNER; CHEN; BOYACK, 2003; RAFOLS; PORTER; LEYDESDORFF, 2010). Diferentemente de mapas baseados apenas em citações, que dependem das relações entre documentos citados e citantes, mapas textuais exploram termos utilizados pelos próprios autores, indexadores ou bases de dados. Eles são especialmente úteis para identificar temas, métodos, objetos empíricos, populações estudadas e mudanças terminológicas. Entretanto, também são sensíveis a ruído linguístico, sinonímia, polissemia, variações de grafia, modismos e diferenças entre palavras-chave de autor e descritores controlados.

Títulos tendem a oferecer sinais de alta precisão, mas baixa riqueza contextual. Como são curtos, privilegiam o tema principal e frequentemente omitem métodos, teorias ou objetos secundários. Resumos oferecem maior densidade semântica, permitindo extrair objetivos, métodos, resultados e lacunas, mas também introduzem variação estilística e termos genéricos. Palavras-chave de autor expressam a autodescrição do artigo, embora possam refletir estratégias de visibilidade. Descritores controlados, quando disponíveis, reduzem variação terminológica, mas podem demorar a incorporar temas emergentes. A decisão sobre que campo textual usar altera substantivamente o mapa.

A construção de mapas conceituais exige pré-processamento cuidadoso. É necessário remover termos vazios, normalizar singular e plural, padronizar siglas, agrupar sinônimos, preservar expressões compostas, decidir se termos muito frequentes serão mantidos e avaliar se palavras genéricas devem ser excluídas. A IA pode apoiar essas tarefas por meio de *prompts* de normalização e agrupamento semântico, mas deve receber instruções claras

para não eliminar termos ambíguos. Em estudos cientométricos, uma palavra ambígua pode ser teoricamente relevante; removê-la automaticamente pode apagar uma tensão conceitual.

A análise de coocorrência de palavras parte da premissa de que termos que aparecem juntos em documentos podem indicar proximidade temática. Essa premissa é útil, mas limitada. Coocorrência não significa equivalência, causalidade ou dependência teórica. Dois termos podem aparecer juntos porque são comparados, porque pertencem a um mesmo método, porque representam objeto e contexto, ou porque um artigo discute sua oposição. Portanto, a interpretação deve considerar a natureza dos termos, os documentos em que aparecem e o nível de granularidade da análise.

Uma vantagem dos mapas textuais é sua sensibilidade a temas emergentes. Antes que uma literatura acumule citações suficientes para produzir padrões robustos de cocitação, seus termos podem aparecer em títulos e resumos. Isso torna a análise textual útil para prospecção e monitoramento. Contudo, a mesma sensibilidade aumenta o risco de capturar ruído. Termos de moda, expressões midiáticas e palavras associadas a chamadas especiais podem gerar *clusters* efêmeros. A IA pode ajudar a distinguir sinais de ruído ao comparar frequência, contexto de uso e persistência temporal, mas a decisão final requer interpretação substantiva.

Figura 12.3 - Indicadores úteis em mapas textuais

Indicador	O que mede	Uso principal	Limitações
Frequência	Ocorrência do termo no corpus	Identificar termos relevantes e dominantes	Favorece termos genéricos populares
Grau (degree)	Nº de conexões de um termo	Importância estrutural e transversalidade	Pode inflar termos muito amplos
Centralidade de intermediação	Papel de ponte entre conjuntos de termos	Interfaces entre subcampos	Sensível a limiares e ruído
Autovetor / PageRank	Conexão com termos também centrais	Prestígio dentro da rede	Reflete centro: pode invisibilizar periferias
Modularidade (Q)	Qualidade dos clusters (comunidades)	Validar divisão da rede	Depende de algoritmo, resolução e dados

Fonte: Imagem gerada por IA, ChatGPT (2026).

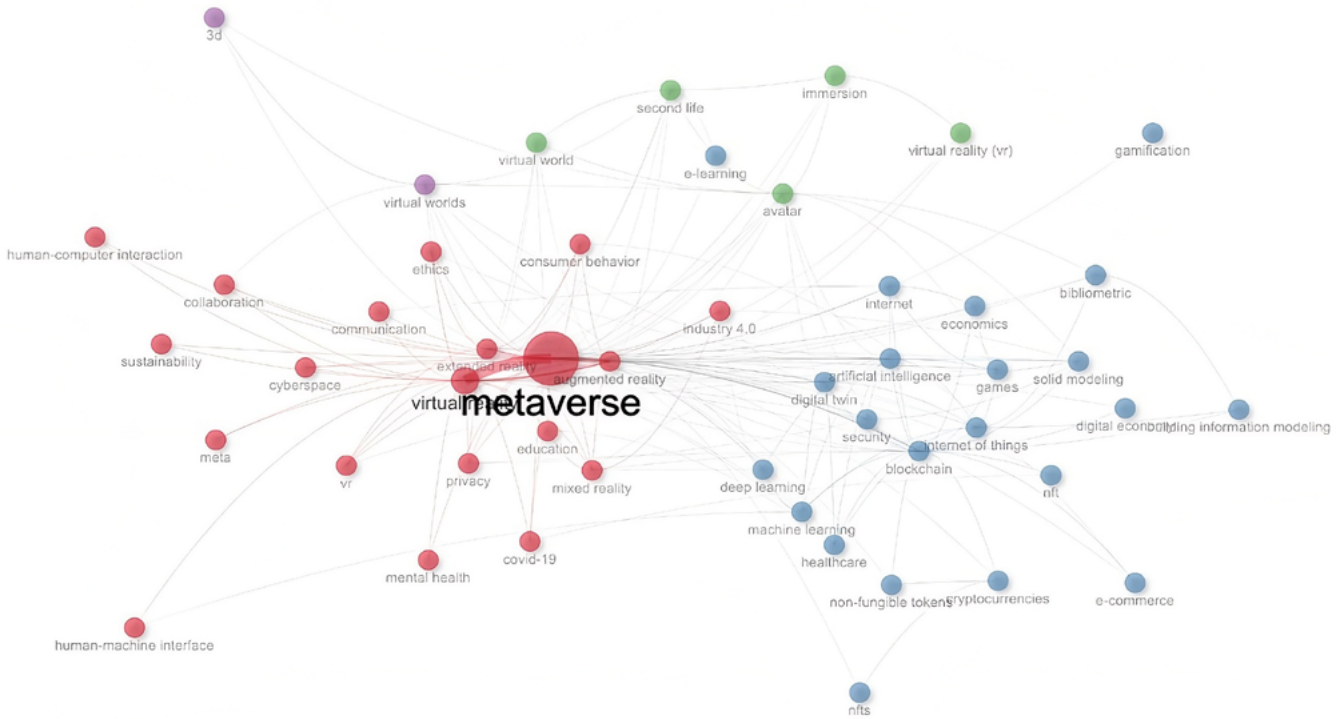
A escolha da granularidade é decisiva. Um mapa construído com termos unigramas pode destacar palavras amplas, como inovação, modelo, sistema ou impacto. Um mapa baseado em bigramas e trigramas pode preservar expressões mais informativas, como inovação aberta, modelo de linguagem, ciência cidadã ou avaliação responsável. Técnicas de extração de termos podem melhorar a qualidade do mapa, mas também introduzem parâmetros. A IA pode sugerir expressões compostas relevantes a partir de uma amostra de títulos e resumos, desde que as sugestões sejam confrontadas com a frequência real nos dados.

Mapas conceituais também podem ser combinados com redes de citação. Essa integração permite comparar estrutura textual e estrutura conceitual. Um *cluster* textual pode reunir documentos que compartilham vocabulário, mas não referências; um *cluster* de cocitação pode reunir documentos que compartilham base teórica, mas utilizam vocabulários diferentes. Quando as duas estruturas convergem, a interpretação ganha robustez. Quando divergem, surge uma oportunidade analítica: talvez o campo esteja passando por tradução terminológica, apropriação interdisciplinar ou recombinação de métodos.

A IA pode apoiar a construção de narrativas comparativas entre mapas textuais e mapas citacionais. Um *prompt* pode solicitar que o modelo identifique convergências, divergências e hipóteses explicativas, mantendo distinção entre evidência lexical e evidência de citação. Essa prática é útil para evitar que o analista atribua a um mapa textual uma profundidade histórica que ele não possui, ou que atribua a um mapa citacional uma atualidade que ele talvez não capture.

A visualização de mapas conceituais deve ser acompanhada de tabelas interpretativas. A figura mostra proximidades e *clusters*, mas a tabela torna explícitos os termos dominantes, documentos centrais, medidas de frequência, período de crescimento e rótulos propostos. Essa combinação entre figura e tabela favorece a reprodutibilidade e facilita a validação por especialistas. A IA pode gerar a primeira versão da tabela analítica, desde que os dados de entrada sejam fornecidos e que o *prompt* exija rastreabilidade das justificativas.

Figura 12.4 - Exemplo de mapa conceitual



Fonte: Imagem gerada por IA, ChatGPT (2026).

Em síntese, mapas conceituais a partir de títulos, resumos e palavras-chave são instrumentos poderosos para representar o vocabulário de um campo, identificar temas e observar mudanças. Eles se tornam ainda mais úteis quando combinados com *prompts* que organizam a leitura dos *clusters*, mas permanecem dependentes de pré-processamento, parâmetros e validação. A IA

não deve ser tratada como substituta da análise terminológica; deve ser incorporada como apoio à curadoria, comparação e explicitação das decisões de interpretação.

Quadro 12.4 - Fontes textuais para mapas conceituais

Fonte textual	Força analítica	Limitação	Uso recomendado com IA
Títulos	Alta precisão temática	Pouco contexto e baixa densidade semântica	Identificar temas centrais e expressões compostas
Resumos	Maior riqueza conceitual e metodológica	Variação estilística e ruído linguístico	Extrair objetivos, métodos, objetos e subtemas
Palavras-chave de autor	Autodescrição da pesquisa	Pode refletir estratégia de visibilidade	Normalizar termos e agrupar variantes
Descritores controlados	Padronização terminológica	Demora para incorporar temas emergentes	Comparar vocabulário controlado e vocabulário emergente

Fonte: Elaborado pelo autor (2026).

12.6 IA para rotular clusters temáticos

A rotulagem de *clusters* temáticos é uma das tarefas mais promissoras e, ao mesmo tempo, mais arriscadas para o uso de IA em cientometria (KOJIMA *et al.*, 2022; WANG *et al.*, 2023; WHITE *et al.*, 2023). Um *cluster* é um conjunto de itens agrupados por similaridade, modularidade ou coocorrência; o rótulo é uma síntese linguística atribuída a esse conjunto. Entre o agrupamento e o rótulo há um salto interpretativo. Modelos de linguagem podem realizar esse salto com fluência, produzindo nomes aparentemente convincentes. Essa fluência, entretanto, pode ocultar incertezas, ambiguidades e extrapolações. A função do *prompt* científico é controlar esse salto, exigindo evidências, alternativas e limites.

A rotulagem manual de *clusters* costuma considerar termos mais frequentes, termos mais distintivos, documentos mais citados, documentos mais centrais, autores de referência, periódicos predominantes e conhecimento especializado sobre a área. A IA pode combinar esses elementos em uma proposta preliminar. Para isso, o *prompt* deve fornecer ao modelo não apenas uma lista de termos, mas também informações contextuais: tipo de rede, período, fonte dos dados, critérios de seleção, tamanho do *cluster*, medidas de centralidade, documentos representativos e objetivo da análise. Quanto mais pobre for a entrada, maior o risco de rótulos genéricos.

Um bom rótulo de *cluster* deve ser informativo, conciso, fiel aos dados e útil para a pergunta de pesquisa. Ele não precisa ser o nome perfeito da área, mas deve permitir que o leitor compreenda o conteúdo dominante do grupo. Rótulos excessivamente amplos, como tecnologia, inovação ou educação, raramente são suficientes. Rótulos excessivamente específicos podem ignorar a diversidade interna do *cluster*. Rótulos avaliativos, como pesquisa avançada, tema promissor ou abordagem superior, devem ser evitados, pois introduzem julgamento não derivado do mapa. A IA deve ser orientada a preferir rótulos descritivos e a explicitar alternativas.

A comparação de rótulos alternativos é uma estratégia importante. Em vez de pedir ao modelo que forneça um único nome, o analista pode solicitar três ou cinco opções, cada uma com justificativa, escopo, risco e grau de confiança. Essa abordagem revela a ambiguidade do *cluster* e evita a naturalização da primeira resposta. Em campos interdisciplinares, é comum que um *cluster* possa ser rotulado por objeto, método, teoria ou aplicação. Por exemplo, um

mesmo grupo pode ser chamado de aprendizagem profunda em saúde, diagnóstico assistido por IA ou análise de imagens médicas. A escolha depende da pergunta cientométrica.

A IA também pode ser usada para detectar rótulos inadequados. Um *prompt* de revisão crítica pode apresentar o rótulo proposto e solicitar que o modelo identifique termos ou documentos que não se encaixam, possíveis vieses, alternativas mais precisas e evidências necessárias para validação. Esse uso crítico é mais seguro do que o uso puramente gerativo, porque transforma o modelo em interlocutor de revisão, não em autoridade final. Ainda assim, a revisão do modelo deve ser verificada por especialistas.

A rotulagem de *clusters* deve respeitar a diferença entre rótulo operacional e rótulo interpretativo. O rótulo operacional descreve o conteúdo do *cluster* no contexto do mapa; o rótulo interpretativo propõe uma leitura teórica ou histórica. Por exemplo, governança de dados em saúde pode ser um rótulo operacional; institucionalização da medicina algorítmica seria uma interpretação mais forte. O segundo rótulo pode ser interessante, mas requer evidências adicionais. *Prompts* devem solicitar que os dois níveis sejam apresentados separadamente.

Um procedimento robusto de rotulagem assistida por IA pode seguir sete etapas. Primeiro, extrair os termos, documentos e métricas do *cluster*. Segundo, solicitar rótulos descritivos preliminares. Terceiro, pedir comparação entre rótulos alternativos. Quarto, solicitar uma revisão crítica do rótulo escolhido. Quinto, confrontar a proposta com leitura dos documentos centrais. Sexto, submeter os rótulos a especialistas. Sétimo, registrar a decisão final no protocolo. Esse fluxo torna o uso da IA documentável e reduz a arbitrariedade.

A rotulagem também deve considerar o público-alvo. Um relatório para gestores de pesquisa pode precisar de rótulos mais comunicativos; um artigo científico pode exigir rótulos mais técnicos; um *dashboard* institucional pode usar nomes curtos e padronizados. A IA pode adaptar rótulos a diferentes públicos, mas a versão analítica deve permanecer documentada. Caso contrário, corre-se o risco de confundir simplificação comunicacional com interpretação científica.

A documentação do processo de rotulagem é parte da reprodutibilidade. O pesquisador deve registrar o *prompt* utilizado, a versão do modelo quando disponível, os dados fornecidos, as opções geradas, os critérios de escolha, as alterações humanas e as justificativas finais. Essa documentação não garante replicação perfeita, pois modelos generativos podem variar suas respostas, mas permite auditoria metodológica. Quando possível, recomenda-se congelar as entradas e guardar as saídas completas em apêndice ou repositório.

Em síntese, a IA é particularmente útil para nomear *clusters*, comparar rótulos e construir narrativas temáticas, mas essa utilidade vem acompanhada de risco. O modelo pode produzir rótulos sedutores, exagerar coerência, ignorar minorias internas do *cluster* ou importar conhecimento externo não presente nos dados. A resposta metodológica é desenhar *prompts* que obriguem o modelo a justificar, relativizar, comparar e indicar incertezas. A rotulagem cientométrica não deve ser delegada; deve ser assistida, auditada e validada.

12.7 Validação dos rótulos por especialistas

A validação dos rótulos por especialistas é etapa essencial em estudos que utilizam mapas temáticos (HICKS *et al.*, 2015). Ela reconhece que a estrutura algorítmica do mapa e a fluência interpretativa da IA não bastam para assegurar validade substantiva. Especialistas podem identificar termos anacrônicos, usos equivocados, relações históricas, diferenças entre escolas, problemas de tradução e lacunas que o algoritmo não percebe. A validação não é um ritual de confirmação; é uma etapa crítica que pode levar à revisão de rótulos, reagrupamento de temas, exclusão de *clusters* incoerentes ou reformulação da narrativa.

Existem diferentes formatos de validação. Uma validação individual pode solicitar que especialistas avaliem a adequação de cada rótulo em escala de concordância. Uma validação qualitativa pode pedir comentários sobre ambiguidade, termos ausentes e interpretações alternativas. Uma validação coletiva, em oficina ou painel, pode promover discussão entre especialistas de subáreas distintas. Uma validação por pares pode ser incorporada ao processo de revisão de artigo. Em estudos institucionais, a validação pode envolver gestores e pesquisadores, desde que a diferença entre interesse estratégico e interpretação científica seja preservada.

A IA pode apoiar a preparação da validação ao gerar formulários, sintetizar justificativas, organizar comentários e identificar pontos de discordância. Contudo, não deve substituir o julgamento dos especialistas. Um uso adequado consiste em pedir ao modelo que transforme os dados de cada *cluster* em uma ficha de validação: rótulo proposto, termos principais, documentos representativos, interpretação preliminar, riscos e perguntas ao avaliador. Essa ficha facilita o trabalho humano e padroniza o processo.

A escolha dos especialistas deve ser documentada. É preciso indicar critérios de seleção, áreas de competência, número de avaliadores e eventuais conflitos de interesse. Em campos interdisciplinares, recomenda-se incluir avaliadores com formações diferentes, pois um rótulo adequado para uma subárea pode ser inadequado para outra. A validação também deve reconhecer assimetrias linguísticas: termos em inglês podem ter traduções diferentes em português, e expressões consolidadas em bases internacionais podem não corresponder ao vocabulário local.

A concordância entre especialistas não precisa ser absoluta para que a validação seja útil. Divergências podem revelar heterogeneidade interna do *cluster*, fronteiras disciplinares ou disputas interpretativas. Em vez de ocultar discordâncias, o relatório pode registrá-las como parte dos resultados. A IA pode ajudar a sintetizar comentários divergentes, desde que o *prompt* instrua o modelo a não resolver artificialmente conflitos. Em alguns casos, a conclusão mais rigorosa é admitir que um *cluster* não possui rótulo único satisfatório.

A validação deve diferenciar precisão, abrangência e utilidade. Um rótulo preciso captura o núcleo dominante do *cluster*; um rótulo abrangente inclui sua diversidade; um rótulo útil comunica a função do *cluster* para a pergunta de pesquisa. Esses critérios podem entrar em tensão. Um rótulo muito preciso pode excluir subtemas; um rótulo muito abrangente pode tornar-se vago; um rótulo muito comunicativo pode simplificar demais. O formulário de validação deve permitir que especialistas avaliem esses critérios separadamente.

A amostragem de documentos para validação é outro ponto crítico. Especialistas não precisam ler todos os documentos de um *cluster*, mas devem ter acesso a uma amostra representativa: documentos mais centrais, mais citados, mais recentes, mais típicos e possivelmente casos de fronteira. A IA pode ajudar a selecionar e resumir essa amostra, mas a seleção deve ser baseada em critérios explícitos. É inadequado validar um rótulo apenas com termos de alta frequência se os documentos centrais apontam uma direção diferente.

Quadro 12.5 - Critérios de validação de rótulos temáticos

Critério	Pergunta ao especialista	Evidência esperada	Possível decisão
Fidelidade	O rótulo representa os termos e documentos centrais?	Convergência entre rótulo, termos e documentos	Manter ou ajustar rótulo
Precisão	O rótulo evita generalidade excessiva?	Delimitação conceitual clara	Especificar objeto, método ou teoria
Abrangência	O rótulo contempla subtemas relevantes?	Inclusão dos subgrupos principais	Adicionar subtítulo ou rótulo composto

Critério	Pergunta ao especialista	Evidência esperada	Possível decisão
Prudência	O rótulo evita inferências causais ou normativas?	Ausência de termos avaliativos indevidos	Reformular em linguagem descritiva
Comunicabilidade	O rótulo é compreensível ao público-alvo?	Clareza sem perda de rigor	Adaptar versão executiva e versão analítica

Fonte: Elaborado pelo autor (2026).

A validação também pode ocorrer em ciclos. No primeiro ciclo, especialistas avaliam rótulos preliminares; no segundo, revisam rótulos ajustados; no terceiro, examinam a narrativa final. Esse procedimento é especialmente recomendado quando a análise será usada em política científica, avaliação institucional ou tomada de decisão. A IA pode organizar as versões dos rótulos e registrar mudanças, criando uma trilha de auditoria. Essa trilha é valiosa porque mostra que o resultado final não foi produzido por um ato isolado de nomeação.

Do ponto de vista ético, a validação por especialistas protege contra falsa autoridade. Mapas coloridos, *clusters* numerados e rótulos gerados por IA podem produzir aparência de objetividade. A validação explicita que a interpretação é situada, revisável e dependente de conhecimento contextual. Também protege contra enviesamentos de base, língua, disciplina e visibilidade. Um *cluster* formado por literatura internacional dominante pode invisibilizar contribuições locais; especialistas podem apontar essa limitação e recomendar cautela na narrativa.

Em síntese, validar rótulos por especialistas não é uma etapa opcional, mas componente central da cientometria interpretativa. A IA pode acelerar e padronizar a preparação da validação, mas a legitimidade dos rótulos depende do confronto entre dados, métodos e conhecimento especializado. O produto final deve registrar não apenas o rótulo escolhido, mas também alternativas consideradas, discordâncias, ajustes e limites. A qualidade de um mapa temático não reside apenas na beleza da visualização, mas na transparência do processo que transforma dados em interpretação.

12.8 Prompts do capítulo

Os *prompts* a seguir são modelos adaptáveis. Devem ser utilizados com dados reais extraídos do estudo, e não como substitutos da coleta, limpeza, visualização ou validação. Recomenda-se registrar integralmente o *prompt*, a entrada fornecida, a saída do modelo e as decisões humanas tomadas após a resposta.

Prompt 12.8.1 Prompt para nomear clusters temáticos

Atue como especialista em cientometria e análise temática.

Receba os dados do *cluster* abaixo, compostos por termos principais, documentos representativos, autores, período e medidas de centralidade e densidade.

Proponha cinco rótulos possíveis para o *cluster*.

Para cada rótulo, apresente: justificativa baseada nos termos e documentos fornecidos; escopo semântico; vantagens; limitações; termos que não se encaixam bem; grau de confiança de 0 a 100.

Não use conhecimento externo para inventar autores, documentos ou tendências ausentes dos dados.

Ao final, recomende o rótulo mais prudente e explique por que ele é preferível aos demais.

Prompt 12.8.2 Prompt para comparar rótulos alternativos

Compare os rótulos alternativos [RÓTULO A], [RÓTULO B] e [RÓTULO C] para o *cluster* descrito abaixo.

Avalie cada rótulo segundo quatro critérios: fidelidade aos dados, precisão conceitual, abrangência temática e utilidade comunicacional.

Identifique riscos de interpretação indevida e indique qual rótulo é mais adequado para um artigo científico, qual é mais adequado para relatório executivo e qual deve ser evitado.

Apresente a resposta em tabela e acrescente uma síntese final com incertezas.

Prompt 12.8.3 Prompt para identificar tema central e subtemas

A partir da lista de títulos, resumos e palavras-chave abaixo, identifique o tema central do conjunto e até seis subtemas.

Para cada subtema, indique termos evidenciais, documentos exemplares, possíveis fronteiras com outros subtemas e grau de coesão.

Separe o que é observável nos dados do que é hipótese interpretativa.

Não transforme frequência de termos em prova de importância científica sem justificar a inferência.

Prompt 12.8.4 Prompt para construir narrativa da estrutura conceitual

Com base nos *clusters*, rótulos, documentos centrais, relações entre *clusters* e séries temporais fornecidas, construa uma narrativa da estrutura conceitual da área.

Organize a resposta em: base intelectual consolidada, temas motores, temas básicos, frentes emergentes, temas isolados ou especializados, tensões conceituais, lacunas e hipóteses de evolução.

Para cada afirmação interpretativa, indique a evidência que a sustenta e o nível de incerteza.

Prompt 12.8.5 Prompt para auditar causalidade em interpretações de mapas

Revise a interpretação abaixo e identifique afirmações que extrapolam os dados.

Classifique cada afirmação como descrição, inferência plausível, hipótese especulativa ou causalidade indevida. Sugira reescritas mais prudentes e indique que evidências seriam necessárias para sustentar uma conclusão causal.

Não elimine a interpretação; torne-a metodologicamente defensável.

Prompt 12.8.6 Prompt para preparar ficha para validação por especialistas

Transforme os dados de cada *cluster* em uma ficha de validação para especialistas.

Cada ficha deve conter: rótulo proposto, rótulos alternativos, termos principais, documentos representativos, período de maior ocorrência, interpretação preliminar, ambiguidades, perguntas de validação e espaço para comentários.

Use linguagem clara, mas mantenha precisão científica. Não esconda incertezas.

12.9 Protocolo recomendado para mapas temáticos assistidos por IA

Quadro 12.6 - Protocolo recomendado para mapas temáticos assistidos por IA

Nº	Descrição da etapa	A IA pode ajudar com...	Entregas principais
01	**Definir a pergunta e o objetivo do mapa** Delimitar claramente a pergunta de pesquisa e a finalidade do mapa. Exemplos de objetivos: exploração inicial, análise de estrutura conceitual, prospecção de frentes, diagnóstico institucional ou comunicação executiva.	Sugerir formulações alternativas de perguntas. Alinhar a pergunta ao tipo de mapa e aos indicadores relevantes. Indicar limitações esperadas para cada objetivo.	Pergunta de pesquisa final. Objetivo do mapa definido. Justificativa da escolha.
02	**Selecionar a fonte de dados e registrar informações** Escolher a base bibliográfica e registrar cobertura, data da coleta, campos disponíveis e limitações. A escolha deve ser justificada em relação ao domínio estudado.	Comparar bases (cobertura, qualidade, tipos documentais). Sugerir campos mais adequados ao objetivo. Listar limitações e vieses da base.	Base selecionada e justificada. Data da coleta. Campos utilizados. Limitações registradas.
03	**Realizar limpeza e pré-processamento dos metadados** Normalizar palavras-chave, termos compostos, nomes de autores e instituições; remover duplicatas e inconsistências. Documentar regras e decisões.	Sugerir regras de normalização. Unificar variantes de termos e instituições. Detectar duplicatas e inconsistências.	Tabela de regras de limpeza. Base limpa (versão 1). Relatório de alterações.

Nº	Descrição da etapa	A IA pode ajudar com...	Entregas principais
04	**Construir a rede temática ou documental** Definir unidade de análise, tipo de relação, medida de associação, limiares e algoritmo de agrupamento. Ex.: coocorrência de palavras-chave, cocitação de referências, acoplamento bibliográfico.	Sugerir medidas de normalização (força da associação, Jaccard, cosseno). Recomendar limiares e algoritmos (Louvain, Leiden, Infomap). Explicar implicações das escolhas.	Especificação metodológica completa. Matriz de rede. Parâmetros utilizados.
05	**Gerar visualizações e tabelas de apoio** Produzir mapas, indicadores de centralidade, densidade, frequência, documentos centrais e distribuição temporal. Combinar diferentes níveis de análise.	Interpretar métricas da rede. Sugerir visualizações complementares. Criar tabelas e gráficos descritivos.	Mapas temáticos. Tabelas de indicadores. Ranking de itens centrais. Séries temporais.
06	**Utilizar prompts para nomear clusters e formular narrativa preliminar** Pedir à IA para nomear *clusters*, comparar rótulos, identificar subtemas e elaborar narrativa inicial, sempre exigindo evidências e incertezas. *Prompts* devem pedir evidências, incertezas, termos representativos e documentos-chave.	Propor rótulos para *clusters*. Comparar alternativas de rótulos. Identificar subtemas e relações. Gerar resumo preliminar com fontes.	Tabela de rótulos candidatos. Evidências (termos e documentos-chave). Narrativa preliminar. Incertezas identificadas.
07	**Revisar criticamente as respostas da IA** Verificar se não há extrapolações, rótulos genéricos, causalidades indevidas e afirmações não sustentadas pelos dados. Distinguir descrição estrutural de interpretação.	Sinalizar afirmações sem evidência. Detectar causalidades indevidas. Apontar generalizações excessivas. Sugerir reformulações cautelosas.	Lista de problemas identificados. Reformulações sugeridas. Versão revisada da narrativa.
08	**Submeter rótulos e narrativa à validação por especialistas** Obter avaliação de especialistas e registrar concordâncias, discordâncias e alterações. Métodos: Delphi, reuniões ou feedback estruturado com anotação colaborativa.	Preparar materiais para validação (rótulos, evidências, mapas). Consolidar o feedback dos especialistas. Comparar atribuições (ex.: kappa).	Registros de validação. Mudanças adotadas e justificadas. Nível de concordância.

Nº	Descrição da etapa	A IA pode ajudar com...	Entregas principais
09	**Redigir a interpretação final** Produzir a interpretação final distinguindo descrição estrutural, inferência plausível, hipótese e limitação metodológica. Usar linguagem relacional e cautelosa.	Estruturar o texto por níveis de afirmação. Sugerir redação clara e equilibrada. Destacar limitações e implicações.	Interpretação final. Quadro de níveis de afirmação. Limitações e implicações.
10	**Arquivar dados, parâmetros, prompts, saídas e decisões** Organizar e armazenar todos os materiais para auditoria, reuso e transparência. Usar padrões de nomeação e metadados.	Gerar documentação completa. Organizar *prompts* e saídas. Criar README com instruções de reuso.	Repositório estruturado. Scripts e parâmetros. Logs de decisões. Metadados da documentação.

Fonte: Elaborado pelo autor (2026).

12.10 Estudo de caso ilustrativo: IA, avaliação da pesquisa e ciência aberta

Para ilustrar o uso integrado dos procedimentos discutidos, imagine-se um estudo cientométrico sobre inteligência artificial na avaliação da pesquisa. A base foi composta por artigos indexados entre 2015 e 2026, extraídos de uma base internacional e complementados por registros abertos. Após limpeza de duplicatas, normalização de palavras-chave e construção de uma rede de coocorrência de termos, o pesquisador identifica seis *clusters* principais. O primeiro reúne termos como *bibliometrics*, *research evaluation*, *citation analysis* e *responsible metrics*. O segundo agrega *artificial intelligence*, *machine learning*, *natural language processing* e automated classification. O terceiro reúne *peer review*, *publication ethics, reproducibility* e research integrity. O quarto articula *altmetrics*, *social media, public engagement* e online *attention*. O quinto contém *open science, open data, open citations* e transparency. O sexto aproxima *policy, funding, institutional evaluation* e *science policy*.

Um uso ingênuo de IA poderia nomear esses *clusters* de modo rápido e aparentemente satisfatório. Entretanto, a análise metodologicamente robusta exige maior cautela. O primeiro *cluster* talvez seja mais bem rotulado como avaliação bibliométrica e métricas responsáveis, mas esse rótulo deve ser confrontado com documentos centrais para verificar se a literatura é crítica, normativa ou técnica. O segundo *cluster* poderia receber o rótulo automação e aprendizagem de máquina em metaciência, mas talvez o termo metaciência não apareça nos dados. O terceiro *cluster* poderia ser interpretado como integridade da comunicação científica, mas a presença de revisão por pares pode indicar uma sublinha específica. A IA deve ser solicitada a comparar essas possibilidades, não a fixar um nome definitivo.

Em seguida, o pesquisador calcula centralidade e densidade. O *cluster* de avaliação bibliométrica aparece como tema motor; o de ciência aberta como tema básico; o de IA aplicada à metaciência como emergente; o de altmetria aparece como tema desenvolvido, mas com centralidade reduzida. Essa configuração não autoriza conclusões automáticas. O fato de a IA aplicada à metaciência aparecer como emergente pode refletir crescimento recente, mudança terminológica ou efeito da estratégia de busca. O fato de altmetria aparecer mais isolada pode indicar especialização, estabilização ou redução de integração com debates atuais. O relatório deve apresentar essas leituras como hipóteses.

A partir dos *clusters*, a IA é usada para gerar rótulos alternativos e fichas de validação. Três especialistas são convidados: um de cientometria, um de ciência aberta e um de inteligência artificial aplicada à informação. Eles avaliam fidelidade, abrangência e precisão dos rótulos. O especialista em IA rejeita o rótulo automação de avaliação científica por considerá-lo normativo demais; o especialista em cientometria sugere incluir o termo modelos de linguagem apenas quando os documentos centrais de fato tratam de LLMs; o especialista em ciência aberta aponta que *open citations* e *open data* não devem ser agregados sem distinção. A narrativa final é ajustada a partir dessas observações.

Esse exemplo mostra que o valor da IA está menos em entregar o mapa final e mais em organizar o processo interpretativo. Ela ajuda a produzir alternativas, expor ambiguidades, sistematizar evidências e redigir sínteses. O pesquisador, contudo, continua responsável por decidir o escopo, avaliar a qualidade dos dados, interpretar o contexto e responder pelos efeitos do relatório. Em mapas temáticos, a aparência de clareza pode ser enganosa; por isso, a triangulação entre algoritmo, IA e especialista é condição de prudência científica.

12.11 Notas metodológicas complementares

A primeira nota refere-se ao estatuto dos rótulos. Rótulos são instrumentos de comunicação e análise, não propriedades naturais dos *clusters*. Eles devem ser suficientemente estáveis para orientar a leitura, mas suficientemente provisórios para aceitar revisão. Em relatórios extensos, recomenda-se indicar quando um rótulo é descritivo, quando é conceitual e quando é teórico. Essa transparência reduz conflitos entre leitores de diferentes formações.

A segunda nota refere-se à compatibilidade entre métodos. Mapas temáticos baseados em co-palavras, cocitação e acoplamento bibliográfico podem produzir estruturas diferentes porque respondem a perguntas diferentes. Copalavras representam proximidade lexical; cocitação representa proximidade na base intelectual; acoplamento bibliográfico captura proximidade entre documentos que compartilham referências. A triangulação entre esses métodos deve buscar complementaridade, não identidade perfeita.

A terceira nota diz respeito à validação temporal. Um tema não é emergente apenas porque aparece no mapa; é emergente quando há evidências de crescimento, diferenciação e inserção em debates recentes. Da mesma forma, um tema não é declinante apenas porque apresenta baixa densidade em um período; pode estar migrando de vocabulário ou sendo incorporado por outros *clusters*. A análise temporal precisa observar frequências, relações e mudanças de linguagem.

A quarta nota envolve a tradução terminológica. Em estudos conduzidos em português, termos internacionais nem sempre têm equivalentes diretos. A decisão entre manter um termo em inglês, traduzi-lo ou usar rótulo bilíngue deve considerar a literatura da área, o público-alvo e a finalidade do estudo. A IA pode sugerir traduções, mas especialistas devem validar termos que possuem tradição consolidada.

A quinta nota trata do risco de homogeneização. Modelos de linguagem tendem a produzir sínteses coerentes e podem suavizar tensões internas. Em mapas temáticos, entretanto, tensões são informativas. Um *cluster* heterogêneo pode revelar debate, disputa teórica ou recombinação interdisciplinar. O *prompt* deve solicitar que o modelo identifique heterogeneidade, e não apenas produza uma narrativa harmonizada.

A sexta nota refere-se à responsabilidade na comunicação. Mapas temáticos usados em gestão científica podem influenciar decisões. Por isso, a versão executiva deve preservar

as limitações essenciais. A IA pode ajudar a simplificar a linguagem, mas não deve apagar condições de validade. Um resultado resumido deve continuar distinguindo o que foi medido, o que foi inferido e o que permanece incerto.

A sétima nota destaca a necessidade de guardar saídas intermediárias. Em estudos assistidos por IA, o resultado final é apenas a ponta visível de um processo iterativo. As versões rejeitadas de rótulos, as críticas produzidas pelo modelo e os comentários de especialistas ajudam a compreender por que a interpretação final foi escolhida. Essa trilha de decisão é parte da qualidade científica do estudo.

A oitava nota trata da escala institucional. Ao aplicar mapas temáticos a universidades, programas ou grupos de pesquisa, é preciso evitar comparações diretas sem normalização de escopo. Uma instituição pode aparecer associada a tema periférico no mapa global e, ainda assim, desempenhar papel estratégico em contexto regional. A leitura deve ser proporcional à pergunta e ao universo analisado.

A nona nota refere-se às limitações dos metadados. Palavras-chave ausentes, resumos incompletos, afiliações ambíguas e referências mal formatadas afetam a construção dos mapas. A IA não deve mascarar essas limitações com linguagem segura. Pelo contrário, pode ser usada para gerar uma seção de limitações com base nas falhas documentadas durante a coleta e o pré-processamento.

A décima nota ressalta que a estética do mapa não substitui rigor. Layouts visualmente atraentes podem ocultar parâmetros frágeis; mapas simples podem ser metodologicamente robustos. O analista deve priorizar clareza, interpretabilidade e documentação. A IA pode sugerir formas de apresentação, mas não deve orientar a análise segundo critérios meramente visuais.

12.12 Integração entre mapas temáticos, indicadores e interpretação qualitativa

A interpretação de mapas temáticos ganha consistência quando é articulada a indicadores bibliométricos e a leitura qualitativa de documentos. O mapa oferece a arquitetura relacional; os indicadores ajudam a dimensionar produção, visibilidade, colaboração e crescimento; a leitura qualitativa fornece sentido substantivo aos padrões encontrados. Nenhuma dessas camadas é suficiente isoladamente. Um *cluster* pode parecer central no mapa e, ainda assim, corresponder a literatura metodologicamente frágil; pode ter baixa visibilidade citacional e, ainda assim, representar uma frente socialmente relevante; pode reunir termos semelhantes e, ainda assim, conter debates teóricos divergentes. A IA pode apoiar a integração dessas camadas ao organizar evidências em matrizes analíticas, mas o pesquisador deve preservar a distinção entre cada tipo de evidência.

Figura 12.5 - Exemplo de matriz integrada de interpretação

Cluster	Nº de Documentos	Crescimento (%) a.a.	Citações Normalizadas (média)	Autores Recorrentes	Instituições Predominantes	Países Principais	Periódicos Principais	Termos Centrais (Top 5)	Docs. Centrais (Exemplos)	Posição no Mapa*	Síntese Interpretativa (Evidências + Incertezas)
1	342	18,7	1,42	Autor A; Autor B	Univ. X; Inst. Y	EUA; China; Alemanha	Journal Z; IEEE Access	IA generativa; LLM; educação; personalização; tutores inteligentes	D1, D2, D3		Forte crescimento e alta visibilidade. Foco em aplicações de IA generativa para personalização. Incerteza: viés de cobertura por idioma.
2	278	12,1	0,98	Autor C; Autor D	Univ. M; Inst. N	Reino Unido; Espanha; Brasil	Computers & Education; J. Learning Analytics	Aprendizagem adaptativa; analytics; dados educacionais; feedback	D4, D5, D6		Tema consolidado em analytics e adaptação. Menor impacto citacional que cluster 1. Incerteza: heterogeneidade conceitual interna.
3	189	24,3	0,61	Autor E; Autor F	Inst. P; Univ. Q	Índia; Coreia do Sul; Turquia	Sustainability; Education and Information Technologies	Sustentabilidade; competências; IA ética; inclusão digital; currículo	D7, D8, D9		Emergente e interdisciplinar. Menor visibilidade ainda, possível fronteira de pesquisa. Incerteza: cobertura desigual em bases selecionadas.
4	121	6,8	1,15	Autor G; Autor H	Univ. R; Inst. S	EUA; Canadá; Austrália	J. Educational Psychology; Review of Education Research	Avaliação; aprendizagem automática; predição; desempenho	D10, D11		Núcleo maduro com teorias estabelecidas. Impacto citacional relativamente alto. Incerteza: pode misturar abordagens teóricas distintas.

Fonte: Imagem gerada por IA, ChatGPT (2026).

Uma matriz integrada de interpretação pode conter, para cada *cluster*, o número de documentos, a taxa de crescimento, a média de citações normalizada, os autores recorrentes, as instituições predominantes, os países, os periódicos, os termos centrais, os documentos representativos, a posição no mapa e uma síntese interpretativa. Essa matriz transforma a visualização em objeto de análise. Em vez de observar apenas a figura, o analista passa a examinar a coerência entre posição estrutural, desempenho bibliométrico e conteúdo textual. A IA pode ser solicitada a preencher a síntese interpretativa com base nos dados fornecidos, mas deve ser proibida de inventar valores ausentes.

A combinação entre mapa e indicador também evita leituras impressionistas. *Clusters* grandes tendem a chamar atenção visual; *clusters* bem-posicionados no centro do mapa podem parecer mais importantes; *clusters* com nomes sedutores podem parecer mais inovadores. Indicadores ajudam a relativizar essas impressões. Entretanto, também podem produzir novos vieses. Citações favorecem áreas mais antigas e internacionalizadas; produção favorece campos com maior volume de publicação; colaboração favorece domínios organizados em grandes equipes. Portanto, a integração deve ser analítica e não hierárquica. O objetivo não é ranquear *clusters*, mas compreender suas funções relativas.

Figura 12.6 - Mapa temático, indicadores bibliométricos e interpretação qualitativa

Fonte: Imagem gerada por IA, ChatGPT (2026).

A leitura qualitativa de documentos centrais é indispensável para calibrar a interpretação temática. Documentos podem ser centrais por frequência de citação, por posição de rede, por representatividade textual ou por papel temporal. A escolha da amostra de leitura deve combinar esses critérios. A IA pode produzir resumos estruturados dos documentos selecionados, identificando objetivo, método, objeto, teoria e contribuição. No entanto, quando a análise depende de nuance conceitual, o resumo automático deve ser conferido no texto original. A interpretação da estrutura conceitual exige contato com a literatura, e não apenas com seus metadados.

A triangulação também permite distinguir tema, método e objeto. Em mapas textuais, termos metodológicos e objetos empíricos podem formar *clusters* mistos. Por exemplo, uma comunidade pode reunir aprendizado de máquina, imagens médicas e diagnóstico precoce. O rótulo poderia enfatizar o método, o objeto ou a aplicação. A decisão depende da pergunta do estudo. Se a pergunta é sobre tecnologias analíticas, o método ganha centralidade; se é sobre saúde pública, a aplicação pode ser mais relevante. Um *prompt* adequado deve pedir que o modelo identifique qual eixo organiza melhor o *cluster* e quais alternativas permanecem plausíveis.

A interpretação qualitativa também ajuda a reconhecer conceitos guarda-chuva. Termos como sustentabilidade, inovação, governança, inteligência artificial e transformação digital podem aparecer em muitos *clusters*. Às vezes funcionam como temas reais; em outras, como etiquetas amplas que ocultam subdomínios. A IA pode ajudar a decompor conceitos guarda-chuva em subtemas, desde que tenha acesso a trechos textuais suficientes. Sem essa decomposição, o mapa pode apresentar uma falsa unidade conceitual.

A integração entre mapas e leitura qualitativa é particularmente relevante quando a análise envolve áreas emergentes. Em domínios recentes, a rede de citações pode ser instável, a terminologia pode estar em disputa e a cobertura das bases pode ser desigual. Mapas temáticos podem fornecer hipóteses iniciais, mas a confirmação exige leitura de documentos-chave, análise de eventos científicos, observação de chamadas de financiamento e acompanhamento de periódicos especializados. A IA pode organizar esses elementos em uma linha de investigação, mas não deve declarar consolidação sem evidência temporal.

Em estudos comparativos, como comparação entre países, instituições ou períodos, a integração deve considerar diferenças de volume e cobertura. Um tema pode parecer ausente em uma instituição porque seus pesquisadores publicam em periódicos não indexados ou usam vocabulário distinto. Um país pode aparecer associado a uma frente por causa de poucos documentos altamente conectados. A IA pode ajudar a detectar discrepâncias e sugerir verificações, mas a interpretação deve ser acompanhada de cautela sobre representatividade.

A integração metodológica também depende da preservação dos dados intermediários. Termos originais, termos normalizados, lista de documentos por *cluster*, parâmetros de rede e saídas da IA devem ser armazenados. Sem esses elementos, o mapa final não pode ser auditado. A reprodutibilidade em estudos com IA não exige apenas guardar o código; exige guardar as instruções conversacionais e as decisões humanas que transformaram saídas probabilísticas em resultados científicos.

Por fim, a integração entre mapa, indicador e interpretação qualitativa ajuda a produzir relatórios mais robustos. A seção de resultados pode apresentar o mapa; a seção analítica pode explicar os *clusters* com base em dados; a discussão pode conectar os achados à literatura e às implicações. A IA pode apoiar a redação, mas o texto final deve manter linguagem prudente. Frases como os dados sugerem, no conjunto analisado, ou uma hipótese interpretativa são preferíveis a formulações categóricas que atribuem causalidade ou universalidade.

Figura 12.7 - Riscos de interpretação e como mitigar

Fonte: Imagem gerada por IA, ChatGPT (2026).

12.13 Erros comuns na leitura de mapas temáticos e estratégias de mitigação

O primeiro erro comum é confundir frequência com centralidade intelectual. Um termo frequente pode aparecer em muitos documentos porque é genérico, obrigatório ou associado ao vocabulário da área, e não porque organiza teoricamente o campo. A mitigação consiste em comparar frequência com distintividade, relações de coocorrência e presença em documentos centrais. A IA pode ser solicitada a classificar termos em genéricos, específicos, metodológicos, teóricos e aplicados, sempre com justificativa e revisão humana.

O segundo erro é aceitar *clusters* como categorias naturais. Algoritmos de agrupamento produzem divisões a partir de parâmetros; pequenas mudanças de limiar, medida de associação ou período podem alterar os grupos. A mitigação exige análise de sensibilidade. O pesquisador pode gerar mapas alternativos e pedir à IA que compare estabilidade dos temas, sem concluir que a versão mais visualmente atraente seja a mais verdadeira. A estabilidade de um *cluster* deve ser evidenciada, não presumida.

O terceiro erro é inferir causalidade a partir de proximidade. A proximidade no mapa indica relação segundo os dados e o método utilizados, mas não prova influência, dependência ou origem comum. Dois temas podem estar próximos porque compartilham vocabulário, porque aparecem nos mesmos periódicos ou porque são comparados em uma literatura crítica. A mitigação consiste em redigir interpretações como hipóteses relacionais e indicar evidências adicionais necessárias para conclusões causais.

O quarto erro é ignorar a dimensão temporal. Um mapa agregado de vinte anos pode esconder mudanças profundas. Temas que aparecem juntos no mapa geral podem pertencer a períodos diferentes; termos antigos e recentes podem ser artificialmente aproximados. A mitigação consiste em produzir mapas por janelas temporais, diagramas de evolução temática ou séries de frequência. A IA pode ajudar a narrar transições, mas deve receber dados temporais explícitos.

O quinto erro é nomear *clusters* com rótulos muito amplos. Rótulos genéricos facilitam comunicação, mas reduzem precisão analítica. Um *cluster* chamado educação pode conter formação docente, tecnologias educacionais, políticas curriculares ou aprendizagem online. A mitigação consiste em pedir à IA rótulos em três níveis: amplo, intermediário e específico. O pesquisador pode então escolher o nível mais adequado ao relatório e registrar o motivo.

O sexto erro é usar conhecimento externo do modelo para preencher lacunas. Modelos de linguagem possuem padrões aprendidos de textos amplos e podem trazer associações não presentes no conjunto analisado. Isso pode ser útil para gerar hipóteses, mas é perigoso para descrever resultados. A mitigação é formular *prompts* que restrinjam a resposta aos dados fornecidos e peçam que qualquer conhecimento externo seja marcado separadamente como contextualização, não como achado.

O sétimo erro é apagar termos minoritários. Durante a normalização e rotulagem, a IA pode enfatizar termos dominantes e ignorar subtemas menores. No entanto, subtemas minoritários podem representar fronteiras emergentes ou perspectivas críticas. A mitigação consiste em pedir que o modelo identifique termos periféricos relevantes e explique se eles indicam ruído, exceção ou sublinha analítica. Especialistas podem avaliar essa classificação.

O oitavo erro é tratar visualização como validação. Um mapa bonito não prova que o agrupamento é conceitualmente adequado. Cores, layout e resolução podem tornar padrões aparentes mais convincentes do que são. A mitigação é acompanhar a figura de tabelas, métricas, critérios e validação. A IA pode produzir legendas e comentários, mas a validação deve vir de dados e especialistas.

O nono erro é esquecer que mapas são produtos de bases de dados específicas. Uma área pode aparecer dominada por literatura anglófona porque a base indexa majoritariamente periódicos internacionais. Uma comunidade local pode desaparecer. A mitigação consiste em discutir cobertura, idioma, tipo documental e fontes complementares. *Prompts* de limitação devem solicitar explicitamente vieses de base e de idioma.

O décimo erro é transformar quadrantes estratégicos em recomendações normativas diretas. Dizer que um tema é motor, básico ou emergente não define automaticamente prioridade de pesquisa, investimento ou avaliação. A mitigação é separar diagnóstico cientométrico de decisão estratégica. A IA pode ajudar a redigir cenários, mas decisões devem considerar missão institucional, capacidade, relevância social e critérios externos ao mapa.

12.14 Do mapa ao relatório: redação científica assistida por prompts

A passagem do mapa temático para o relatório científico é uma etapa frequentemente subestimada. Muitos estudos apresentam figuras sofisticadas, mas oferecem interpretação limitada. O relatório precisa explicar por que o mapa foi construído, que dados o sustentam, que parâmetros foram utilizados, como os *clusters* foram nomeados, quais evidências apoiam a interpretação e quais limitações permanecem. A IA pode auxiliar essa passagem ao transformar tabelas e descrições em texto articulado, mas deve ser orientada por um roteiro metodológico claro.

Quadro 12.7 - Estrutura recomendada do relatório

Etapa	O que incluir	Pergunta-Guia	Como a IA pode ajudar *(com prompts)*
1. Descrição do Mapa	• Quantos *clusters* foram identificados? • Quais são os maiores (por documentos)? • Quais estão mais centrais? Quais mais densos? • Quais termos e documentos caracterizam cada *cluster*? • Quais são as principais conexões entre *clusters*?	"O que os dados mostram?" (descrição neutra)	*Prompt*: "Descreva objetivamente o mapa temático abaixo (com base na tabela/figura fornecida). Liste número de *clusters*, tamanho, posição relativa, termos centrais, documentos representativos e conexões principais. Não interprete, apenas descreva."
2. Interpretação dos padrões	• Quais temas motores e maduros? • Quais frentes emergentes ou em crescimento? • Quais comunidades epistêmicas se formam? • Que tensões ou sobreposições conceituais aparecem? • Que mudanças temporais podem estar associadas?	"O que esses padrões podem significar no contexto do campo?" (interpretação)	*Prompt*: "Interprete os padrões observados para cada *cluster* usando as evidências da tabela. Associe cada interpretação às evidências (termos, documentos, métricas). Identifique frentes emergentes, temas motores e possíveis tensões. Indique grau de confiança."

Etapa	O que incluir	Pergunta-Guia	Como a IA pode ajudar *(com prompts)*
3. Discussão com a literatura	• Como os achados se alinham ou divergem da literatura? • Há estudos prévios que observaram fenômeno similar? • Que fatores externos podem explicar as tendências? • A base e o período permitem sustentar as conclusões?	"Como esses achados dialogam com o que já se sabe?" (contextualização)	*Prompt*: "Compare as interpretações dos *clusters* com a literatura indicada (lista de referências/trechos fornecidos). Identifique convergências, divergências e lacunas. Cite evidências relevantes e limite-se às fontes fornecidas (não invente referências)."
4. Limitações	• Cobertura da base de dados e vieses de indexação. • Qualidade dos metadados e ambiguidade terminológica. • Sensibilidade a parâmetros (limiares, algoritmos, layout). • Instabilidade de *clusters* pequenos ou periféricos. • Ausência de literatura não indexada e viés linguístico.	"Quais são as limitações que delimitam a validade dos achados?"	*Prompt*: "Liste as limitações metodológicas e de dados deste estudo de mapa temático. Explique como cada limitação pode afetar a interpretação."
5. Tradução para o público	• Explicar o significado das categorias e métricas. • Usar linguagem clara e exemplos concretos. • Manter precisão sem jargões desnecessários. • Evidenciar implicações práticas, se houver.	"Como explicar esses achados para um público não especialista?"	*Prompt*: "Reescreva a interpretação dos *clusters* em linguagem acessível para gestores/tomadores de decisão. Explique o que o mapa representa, o que não representa e quais são as implicações práticas."
6. Revisão crítica de extrapolações	• Evitar causalidade indevida e determinismo. • Substituir verbos categóricos por verbos prudentes. • Conferir se cada afirmação tem evidência. • Marcar hipóteses como tais (não como conclusões).	"Onde posso ter extrapolado ou afirmado sem evidência?"	*Prompt*: "Revise o texto abaixo e identifique: (1) afirmações sem evidência, (2) linguagem determinista, (3) causalidades indevidas. Reescreva com linguagem prudente e clareza sobre evidências e limitações."
7. Documentação e arquivo	• Guardar dados, scripts, parâmetros e versões. • Registrar *prompts*, respostas da IA e decisões humanas. • Documentar critérios, exclusões e normalizações. • Disponibilizar materiais para auditoria e reuso.	"Como garantir transparência, auditoria e reprodutibilidade?"	*Prompt*: "Organize a documentação do estudo: base, data de coleta, parâmetros, scripts, *prompts* usados, saídas da IA, decisões e validações. Sugira estrutura de arquivo e nomes de pastas/arquivos."

Fonte: Elaborado pelo autor (2026).

Uma boa seção de resultados deve começar pela descrição objetiva do mapa. Quantos *clusters* foram identificados? Quais são os maiores? Quais aparecem mais centrais? Quais apresentam maior densidade? Quais termos e documentos caracterizam cada grupo? Essa

descrição deve evitar conclusões explicativas prematuras. A IA pode ser instruída a escrever uma descrição neutra, baseada apenas em dados observáveis, antes de produzir qualquer interpretação.

Depois da descrição, o relatório pode apresentar interpretação. Nesse nível, o pesquisador conecta os padrões a debates do campo, teorias, métodos e mudanças temporais. É aqui que se discutem temas motores, frentes emergentes, comunidades epistêmicas e tensões conceituais. A IA pode sugerir narrativas, mas o *prompt* deve exigir que cada interpretação seja associada a evidências. Interpretações sem evidência devem ser marcadas como hipóteses.

A seção de discussão deve relacionar os mapas à literatura existente. Isso evita que o mapa seja lido como evidência isolada. Se o *cluster* sugere emergência de determinado tema, a discussão deve perguntar se estudos anteriores observaram fenômeno semelhante, se há fatores externos que explicam a tendência e se a base de dados permite sustentar a conclusão. A IA pode ajudar a organizar a discussão, mas referências devem ser verificadas, pois modelos podem produzir citações inexistentes ou incorretas.

O relatório também deve explicitar limitações. Limitações não enfraquecem o estudo; elas delimitam sua validade. Em mapas temáticos, as principais limitações incluem cobertura da base, qualidade dos metadados, ambiguidade terminológica, sensibilidade a parâmetros, instabilidade de *clusters* pequenos, ausência de literatura não indexada e dependência de validação humana. Um *prompt* específico para limitações pode ajudar a evitar omissões.

Quando o público do relatório é não especialista, a IA pode auxiliar na tradução da linguagem técnica. Contudo, simplificação não deve eliminar precisão. Em vez de dizer que o mapa mostra os temas mais importantes, o relatório pode explicar que ele mostra temas mais conectados ou mais desenvolvidos no conjunto de dados. Essa diferença é crucial. O texto deve educar o leitor sobre o significado das categorias, não apenas entregar uma conclusão pronta.

A redação de mapas também deve evitar linguagem determinista. Expressões como comprova, demonstra definitivamente ou revela a causa devem ser substituídas por indica, sugere, é compatível com ou pode ser interpretado como. A IA pode ser usada para revisar o texto e identificar extrapolações. Esse *prompt* de revisão crítica é especialmente útil antes da submissão de artigo ou entrega de relatório institucional.

Por fim, o relatório deve incluir um apêndice metodológico. Nele podem constar *strings* de busca, datas de coleta, softwares, versões, parâmetros, termos excluídos, critérios de normalização, *prompts*, respostas da IA e síntese da validação especializada. Esse apêndice transforma o uso de IA em procedimento científico documentado. Sem ele, a análise se torna difícil de auditar e reproduzir.

Quadro 12.8 - Checklist de qualidade para mapas temáticos assistidos por IA

CHECKLIST DE QUALIDADE *(antes de fechar o estudo relacional assistido por IA)*	
☐	Pergunta de pesquisa clara e alinhada ao método.
☐	*Corpus* bem definido, reproduzível e documentado.
☐	Construção da matriz relacional transparente.
☐	Parâmetros, limiares e algoritmos informados.
☐	Análises de sensibilidade realizadas.
☐	Interpretações separadas de evidências.
☐	Limitações e vieses explicitados.

CHECKLIST DE QUALIDADE *(antes de fechar o estudo relacional assistido por IA)*
☐ Documentação dos *prompts* e das saídas da IA.
☐ Revisão humana e validação qualitativa.
☐ Conclusões proporcionais às evidências.

Registro da verificação

Estudo ou etapa verificada	Responsável	Data

Fonte: Elaborado pelo autor (2026).

12.15 Aplicações em política científica, gestão de pesquisa e organização do conhecimento

Mapas temáticos da estrutura conceitual podem apoiar políticas científicas quando são utilizados como instrumentos de diagnóstico, e não como mecanismos automáticos de decisão (PORTER; RAFOLS, 2009; RAFOLS; PORTER; LEYDESDORFF, 2010; HICKS *et al.*, 2015). Uma agência de fomento pode empregar mapas para identificar concentração de competências, lacunas temáticas, frentes emergentes e possibilidades de cooperação interinstitucional. Uma universidade pode utilizá-los para compreender como suas publicações se distribuem em relação a temas globais e regionais. Uma biblioteca universitária pode usar mapas para orientar desenvolvimento de coleções, serviços de informação e curadoria de repositórios. Em todos esses casos, a IA pode gerar sumários e perguntas estratégicas, mas a decisão deve considerar critérios sociais, éticos, econômicos e institucionais que extrapolam o mapa.

Na gestão de pesquisa, mapas temáticos permitem observar a relação entre capacidades internas e agendas externas. Um programa de pós-graduação pode mapear suas teses, dissertações e artigos para verificar se há coerência entre linhas declaradas e produção efetiva. Um grupo de pesquisa pode comparar seus temas com a literatura internacional para identificar oportunidades de colaboração. Uma instituição pode descobrir que possui competências dispersas em diferentes departamentos, mas relacionadas por temas comuns. A IA pode ajudar a nomear essas competências, construir painéis executivos e redigir diagnósticos, desde que os limites dos dados sejam preservados.

Em organização do conhecimento, mapas temáticos dialogam com tesauros, taxonomias, ontologias e sistemas de classificação. A coocorrência de termos pode revelar vocabulários emergentes que ainda não foram incorporados a instrumentos controlados. A comparação entre descritores oficiais e palavras-chave de autor pode indicar desalinhamentos entre linguagem institucional e linguagem de pesquisa. A IA pode sugerir equivalências, hierarquias e relações associativas, mas essas sugestões devem ser avaliadas por especialistas em organização do conhecimento. O risco é confundir similaridade semântica superficial com relação conceitual legítima.

Em estudos de prospecção tecnológica, mapas temáticos podem mostrar aproximações entre ciência básica, desenvolvimento tecnológico e aplicações industriais. A análise de co-palavras, patentes, artigos e projetos pode revelar zonas de convergência. No entanto, a interpretação deve distinguir maturidade científica de prontidão tecnológica. Um tema muito citado pode ainda estar distante de aplicação; um tema pouco citado pode ter relevância

estratégica em nichos específicos. A IA pode auxiliar a construir matrizes de evidências, mas não deve converter presença temática em previsão de mercado ou recomendação de investimento sem análise complementar.

Em avaliação institucional, mapas temáticos devem ser usados com prudência. Eles podem revelar diversidade, especialização, interdisciplinaridade e evolução temática de uma instituição, mas não medem qualidade de maneira direta. Um *cluster* institucional com alta densidade pode indicar foco; pode também indicar isolamento. Um mapa amplo pode indicar diversidade; pode também indicar dispersão. A interpretação depende da missão da instituição, do tamanho do corpo docente, da área de atuação e dos objetivos estratégicos. *Prompts* destinados a relatórios institucionais devem solicitar alternativas interpretativas, e não uma conclusão única.

Na comunicação pública da ciência, mapas temáticos podem tornar visíveis relações entre temas complexos. Um relatório para sociedade pode mostrar como pesquisas sobre clima, saúde, energia e políticas públicas se conectam. A IA pode traduzir mapas técnicos para linguagem acessível, mas deve evitar simplificações enganosas. A legenda deve explicar que o mapa representa relações no conjunto de publicações analisado, e não a totalidade da realidade científica. Essa distinção é fundamental para comunicação responsável.

Em ambientes educacionais, mapas temáticos podem apoiar formação de pesquisadores. Estudantes de pós-graduação podem usá-los para compreender a estrutura de um campo, localizar autores centrais, identificar frentes de pesquisa e formular perguntas. A IA pode funcionar como tutor metodológico, explicando categorias e sugerindo leituras. Contudo, o uso pedagógico deve enfatizar que mapas são pontos de partida para leitura crítica, não atalhos para substituir a familiarização com a literatura. Um mapa não dispensa leitura; ele orienta a leitura.

Aplicações em políticas científicas devem enfrentar ainda o problema da performatividade. Quando mapas e indicadores passam a orientar decisões, pesquisadores podem adaptar estratégias de publicação, palavras-chave e colaboração para aparecerem em determinados temas. A própria visualização pode influenciar comportamento. Por isso, a governança de mapas temáticos exige transparência, revisão periódica e participação das comunidades avaliadas. A IA pode ajudar a produzir relatórios, mas não deve tornar opacos os critérios de classificação e rotulagem.

A utilização de mapas em contextos decisórios requer uma camada de governança metodológica. Essa governança deve definir quem pode alterar rótulos, como divergências são registradas, como atualizações são feitas, como especialistas são selecionados e como limitações são comunicadas. A IA deve operar dentro desse arranjo, como ferramenta documentada. Sem governança, mapas temáticos assistidos por IA podem gerar decisões rápidas, mas pouco responsáveis.

Em síntese, mapas temáticos têm grande valor aplicado, desde que sejam lidos como instrumentos interpretativos situados. Eles ajudam a formular perguntas, orientar leituras, organizar diagnósticos e comunicar estruturas complexas. A IA amplia sua utilidade ao acelerar síntese, rotulagem e redação. Contudo, quanto maior o impacto decisório do mapa, maior deve ser o rigor da validação. A regra prática é simples: mapas usados para explorar podem tolerar mais incerteza; mapas usados para decidir exigem documentação, validação e responsabilidade institucional ampliadas.

12.16 Modelos avançados de prompts para análise temática reprodutível

Os *prompts* avançados para mapas temáticos devem ser pensados como instrumentos de protocolo (WHITE *et al.*, 2023; KOJIMA *et al.*, 2022; WANG *et al.*, 2023). Eles não servem apenas para obter respostas; servem para estruturar uma interação analítica que possa ser examinada posteriormente. Um *prompt* robusto informa o papel atribuído ao modelo, os dados permitidos, a tarefa, o formato da resposta, os critérios de julgamento e os limites da inferência. Em mapas temáticos, isso significa declarar se a resposta deve nomear *clusters*, comparar rótulos, revisar causalidade, identificar subtemas, construir narrativa ou preparar validação especializada.

Um primeiro modelo avançado é o *prompt* de leitura estratificada do *cluster*. Nele, o pesquisador fornece termos, documentos centrais, medidas e período. O modelo deve responder em quatro camadas: descrição lexical, descrição documental, interpretação temática e incertezas. Essa estratificação evita que a primeira impressão semântica se transforme em conclusão. Também facilita a auditoria, pois permite verificar se a interpretação decorre dos termos, dos documentos ou de conhecimento externo.

Um segundo modelo é o *prompt* de confronto entre evidências. O pesquisador pode pedir que a IA compare sinais textuais, citacionais e temporais. Por exemplo: o *cluster* parece emergente nos termos, mas suas referências são antigas; ou o *cluster* apresenta vocabulário novo, mas documentos centrais pertencem a uma tradição consolidada. O objetivo é produzir tensão interpretativa, não apenas síntese. Bons *prompts* não buscam respostas confortáveis; buscam tornar explícitas as ambiguidades.

Um terceiro modelo é o *prompt* de auditoria de rótulos. Após escolher um rótulo, o pesquisador solicita que o modelo tente refutá-lo. A resposta deve indicar termos que ficam de fora, documentos que sugerem outro foco, riscos de generalização, alternativas mais restritas e evidências necessárias. Esse procedimento aproxima o uso da IA de uma revisão crítica. Ele ajuda a reduzir vieses de confirmação, principalmente quando o analista já tem expectativa sobre o tema.

Um quarto modelo é o *prompt* de tradução para públicos. Depois de validar o rótulo analítico, o pesquisador pode pedir uma versão técnica, uma versão executiva e uma versão pedagógica. A versão técnica preserva precisão; a executiva comunica implicações; a pedagógica explica conceitos. Essa adaptação é útil, mas deve manter rastreabilidade. O rótulo executivo não deve substituir o rótulo científico no protocolo, pois simplificação comunicacional e classificação analítica possuem finalidades diferentes.

Um quinto modelo é o *prompt* de narrativa controlada. O modelo recebe uma tabela de *clusters* e deve produzir uma narrativa separada em parágrafos de descrição, interpretação e limitação. Cada frase interpretativa deve conter referência ao tipo de evidência: frequência, centralidade, densidade, crescimento temporal, documentos centrais ou validação especializada. Essa exigência torna o texto menos fluente, mas mais científico. Em relatórios formais, a clareza da evidência é mais importante do que o brilho retórico.

Um sexto modelo é o *prompt* de sensibilidade metodológica. A IA recebe duas versões de mapa, produzidas com parâmetros diferentes, e deve comparar quais *clusters* permanecem, quais mudam, quais rótulos são estáveis e quais conclusões dependem dos parâmetros. Essa prática é especialmente importante quando os resultados serão usados em avaliação ou planejamento. Um achado que desaparece com pequena mudança de limiar deve ser apresentado com cautela.

Um sétimo modelo é o *prompt* de preparação de apêndice. O modelo organiza informações metodológicas em formato de reprodutibilidade: base, período, consulta,

campos, software, parâmetros, limiares, normalizações, *prompts* e validações. Esse uso é pouco glamoroso, mas extremamente relevante. A credibilidade de estudos assistidos por IA dependerá da capacidade de documentar o processo com detalhe suficiente para auditoria.

Um oitavo modelo é o *prompt* de comparação interdisciplinar. Em áreas híbridas, o pesquisador pode pedir que a IA identifique qual parte do *cluster* se relaciona a cada tradição disciplinar. A resposta deve evitar reduzir o *cluster* a uma disciplina dominante e deve indicar possíveis zonas de tradução. Essa abordagem é útil para campos como sustentabilidade, saúde digital, humanidades digitais, ciência aberta e inteligência artificial responsável.

Um nono modelo é o *prompt* de identificação de lacunas. O modelo analisa *clusters* e pergunta quais temas esperados estão ausentes ou sub-representados, mas deve marcar essas lacunas como hipóteses dependentes da estratégia de busca e da base. Essa distinção evita afirmar que uma área não pesquisa determinado tópico quando, na verdade, o tópico pode ter sido excluído pelo recorte. A IA pode gerar perguntas de verificação para ampliar a busca ou consultar especialistas.

Um décimo modelo é o *prompt* de revisão ética. O modelo examina se a interpretação usa linguagem depreciativa, hierarquiza temas de modo inadequado, invisibiliza comunidades ou transforma categorias técnicas em juízos de valor. Essa revisão é relevante porque mapas temáticos podem circular em espaços institucionais. A linguagem do relatório deve ser proporcional à evidência e respeitosa com a diversidade de áreas e tradições.

Esses modelos avançados mostram que engenharia de *prompts*, em cientometria, não é uma habilidade periférica. Ela integra o desenho metodológico. O *prompt* define o que a IA pode ver, o que pode inferir, como deve justificar e quando deve declarar incerteza. A qualidade do resultado depende menos de comandos isolados e mais da construção de uma cadeia de *prompts* compatível com a lógica do estudo. Mapas temáticos assistidos por IA são mais confiáveis quando a interação com o modelo é planejada como parte do método.

12.17 Reprodutibilidade, versionamento e preservação das decisões interpretativas

A reprodutibilidade de mapas temáticos assistidos por IA deve ser entendida de modo ampliado (ARIA; CUCCURULLO, 2017; MORAL-MUÑOZ *et al.*, 2020; HICKS *et al.*, 2015). Em estudos tradicionais, costuma-se registrar base de dados, consulta, data de coleta, software e parâmetros. Em estudos que utilizam *prompts*, é necessário registrar também as interações com o modelo, as versões dos rótulos, as respostas descartadas, as justificativas humanas e os critérios de validação. Isso ocorre porque a interpretação final não deriva apenas do algoritmo de *cluster*ização, mas de uma cadeia de decisões que inclui linguagem natural, inferência probabilística e julgamento especializado.

O versionamento dos rótulos é especialmente importante. Um *cluster* pode receber inicialmente um rótulo gerado pela IA, depois um rótulo ajustado pelo pesquisador, depois uma terceira versão após validação por especialistas. Cada versão revela uma etapa do raciocínio. Guardar apenas o rótulo final apaga o processo de deliberação e dificulta auditoria. Em relatórios mais robustos, recomenda-se incluir uma tabela com rótulo inicial, rótulos alternativos, rótulo validado e razão da mudança.

A preservação das decisões interpretativas também protege contra instabilidade dos modelos. Um mesmo *prompt* pode gerar respostas diferentes em momentos distintos, dependendo do modelo, da configuração e do contexto. Essa variação não inviabiliza o uso científico da IA, mas exige documentação. Quando possível, o pesquisador deve registrar data,

ferramenta, configuração, dados de entrada e saída completa. Em contextos sensíveis, pode-se repetir o *prompt* e comparar a estabilidade das respostas, usando a convergência como indício de robustez, não como prova definitiva.

Outro aspecto da reprodutibilidade é a separação entre dados públicos e dados restritos. Em estudos institucionais, alguns metadados, comentários de especialistas ou informações estratégicas podem não ser divulgáveis. Ainda assim, é possível documentar o procedimento de modo suficiente, indicando quais elementos foram preservados internamente, quais foram anonimizados e quais limitações isso impõe à replicação. A IA não deve receber dados sensíveis sem avaliação ética e institucional adequada.

A documentação também deve contemplar decisões negativas: termos excluídos, *clusters* rejeitados, rótulos abandonados e interpretações recusadas por falta de evidência. Essas decisões são cientificamente relevantes porque mostram o controle exercido pelo analista. Em muitos casos, a qualidade de um estudo depende tanto do que foi incluído quanto do que foi deliberadamente excluído. *Prompts* podem ajudar a gerar um registro dessas decisões, transformando a curadoria em evidência metodológica.

Por fim, a reprodutibilidade deve ser comunicada de modo legível. Um apêndice excessivamente técnico pode afastar leitores; uma descrição curta demais impede auditoria. A solução é combinar uma seção metodológica narrativa com tabelas de parâmetros, arquivos suplementares e, quando possível, repositórios. A IA pode ajudar a adaptar a documentação para diferentes públicos, mas o conteúdo técnico essencial deve permanecer acessível. Mapas temáticos confiáveis são aqueles cujos caminhos de produção podem ser reconstruídos, questionados e aprimorados por outros pesquisadores.

A preservação das decisões também favorece aprendizagem cumulativa. Em vez de cada estudo reiniciar a discussão sobre rótulos, parâmetros e validação, registros bem-organizados permitem comparar escolhas, identificar padrões de erro e aperfeiçoar protocolos. Esse aspecto é especialmente relevante para equipes que produzirão séries de relatórios ou capítulos ao longo do tempo. A IA pode ajudar a transformar registros anteriores em memória metodológica, desde que essa memória seja revisada e atualizada por pesquisadores responsáveis, com critérios explícitos, governança documental, revisão periódica, responsabilidade autoral, preservação de evidências, análise crítica continuada, formação metodológica da equipe e compromisso permanente com transparência científica e qualidade interpretativa em todas as etapas documentadas do estudo.

12.18 Considerações finais

Os mapas temáticos e a análise da estrutura conceitual da ciência ocupam um ponto de convergência entre método quantitativo, interpretação qualitativa e visualização. Eles permitem observar como temas se agrupam, como frentes emergem, como tradições se estabilizam e como comunidades cognitivas se relacionam. Ao mesmo tempo, exigem cautela porque transformam dados incompletos, escolhas metodológicas e relações probabilísticas em representações visualmente persuasivas. A incorporação de IA intensifica tanto o potencial quanto o risco dessa operação.

O principal argumento deste capítulo é que a IA deve ser integrada à análise temática como instrumento de apoio, não como substituto do julgamento científico. *Prompts* podem ajudar a nomear *clusters*, comparar rótulos, identificar subtemas, construir narrativas e preparar validação por especialistas. No entanto, cada uma dessas tarefas deve ser acompanhada de evidências, incertezas, alternativas e critérios explícitos. A fluência textual da IA não deve ser confundida com validade epistemológica.

A estrutura conceitual da ciência não se revela integralmente em um único mapa. Ela precisa ser reconstruída por meio de múltiplas camadas: citações, termos, documentos, autores, instituições, temporalidade, contextos e leitura especializada. Mapas temáticos oferecem uma entrada poderosa nessa reconstrução, especialmente quando combinam diagramas estratégicos, análise de frentes, mapas conceituais e validação dos rótulos. A IA pode ampliar a capacidade do pesquisador de explorar essas camadas, mas somente a documentação metodológica permite transformar a exploração em conhecimento confiável.

Em termos práticos, recomenda-se que todo estudo de mapas temáticos assistido por IA registre a base de dados utilizada, a estratégia de busca, os procedimentos de limpeza, os parâmetros de rede, o algoritmo de agrupamento, os dados fornecidos ao modelo, os *prompts* empregados, as respostas obtidas, as decisões humanas, as validações e as limitações. Esse registro não é burocracia; é a condição para que outros pesquisadores compreendam como uma figura se tornou argumento.

O capítulo também procurou mostrar que nomear é interpretar. Atribuir um rótulo a um *cluster* não é uma operação estética, mas uma decisão analítica. Um rótulo pode ampliar a compreensão de uma área ou induzir leituras equivocadas. Por isso, *prompts* de rotulagem devem exigir justificativa, comparação, revisão crítica e validação. A qualidade do mapa depende tanto da técnica de visualização quanto da responsabilidade com que se transforma agrupamento em conceito.

Na trilha deste livro, dedicada à cientometria com inteligência artificial e *prompts*, os mapas temáticos representam um campo privilegiado de aplicação. Eles demandam exatamente aquilo que a IA pode oferecer com valor: síntese, comparação, organização textual e geração de hipóteses. Mas também exigem aquilo que a IA não pode garantir sozinha: conhecimento substantivo, julgamento crítico, responsabilidade interpretativa e compromisso com a reprodutibilidade. O futuro dos mapas temáticos assistidos por IA dependerá menos da automação completa da análise e mais da construção de protocolos em que máquinas ampliem a capacidade humana de perguntar, verificar e interpretar.

Capítulo 13

Análise temporal e detecção de tendências

13.1 Introdução

A ciência é uma prática histórica. Essa afirmação, aparentemente simples, tem consequências metodológicas profundas para a cientometria. Nenhuma área de conhecimento se apresenta ao pesquisador como uma fotografia imóvel; ao contrário, toda área é resultado de acúmulos, rupturas, esquecimentos, reativações, disputas terminológicas, ciclos de financiamento, transformações institucionais, mudanças tecnológicas e variações na forma de publicar. Por essa razão, uma análise cientométrica que ignore o tempo tende a converter processos dinâmicos em retratos estáticos. Ela pode descrever o volume total de publicações, listar autores ou indicar palavras-chave frequentes, mas dificilmente explicará como uma frente de pesquisa se consolidou, por que determinado vocabulário substituiu outro, quando uma agenda se tornou visível ou em que momento uma rede de colaboração passou de periférica a central (PRICE, 1965; FORTUNATO *et al.*, 2018).

No campo do mapeamento científico, a dimensão temporal já aparece como eixo indispensável para compreender trajetórias, linhas do tempo, animações, transições entre *clusters* e evolução de temas. O sumário-base da obra que orienta este livro atribui uma seção específica a representação do tempo, incluindo mapeamento longitudinal, trajetórias e abordagens baseadas em linha do tempo. Esta escolha é coerente com a literatura de ciência da ciência: mapas, redes e indicadores tornam-se mais explicativos quando deixam de ser apenas distribuições agregadas e passam a revelar sequências, ritmos e mudanças. A temporalidade permite distinguir um tema historicamente consolidado de uma moda passageira, um crescimento estrutural de um aumento causado por mudança de indexação, uma especialidade emergente de um simples reposicionamento terminológico (CHEN, 2004; COBO *et al.*, 2011; COBO *et al.*, 2012).

A incorporação de inteligência artificial a esse processo modifica a escala e a forma de trabalho. Modelos de linguagem podem resumir tendências, comparar períodos, explicar mudanças lexicais, sugerir hipóteses para deslocamentos temáticos e organizar narrativas interpretativas a partir de tabelas, redes e séries. Também podem ajudar a formular perguntas que antes exigiriam longa exploração manual: quais termos aparecem apenas nos últimos anos? Quais *clusters* cresceram em centralidade? Que autores migraram entre comunidades? Que países passaram a ocupar posições de articulação internacional? Quais tópicos aumentaram rapidamente, mas ainda apresentam baixa densidade conceitual? Entretanto, essas capacidades só produzem conhecimento confiável quando integradas a protocolos que preservam os dados, explicitam a fonte, documentam a estratégia de busca, delimitam períodos e distinguem observação, inferência e especulação.

A análise temporal em cientometria não deve ser confundida com profecia. Detectar uma tendência não significa predizer o futuro com certeza; formular um cenário prospectivo não significa afirmar que ele ocorrerá; identificar crescimento de publicações não significa demonstrar relevância científica; observar aumento de citações não significa provar superioridade epistêmica. O tempo torna os dados mais ricos, mas também torna a interpretação mais arriscada. Uma série ascendente pode resultar de maior interesse substantivo, de ampliação de cobertura da base, de mudança nos termos de busca, de crescimento geral da literatura, de políticas de financiamento, de uma crise social ou sanitária, de aumento de periódicos

predatórios ou de simples atraso de indexação em anos recentes. Por isso, a pergunta central deste capítulo não é apenas como detectar tendências, mas como interpretá-las com rigor.

Quadro 13.1 - Perguntas orientadoras para análise temporal em cientometria

Dimensão	Pergunta orientadora	Risco interpretativo
Crescimento	A produção aumenta, diminui ou se estabiliza ao longo do tempo?	Confundir expansão geral da base com interesse substantivo no tema.
Emergência	Há sinais convergentes de novo tema, novo *cluster* ou nova comunidade?	Classificar como emergente qualquer termo recente ou midiático.
Terminologia	Os termos usados para designar o tema mudaram?	Perder literatura precursora ou agrupar conceitos não equivalentes.
Trajetória	Autores, instituições ou países mudaram de posição no campo?	Inferir liderança ou estratégia sem evidência contextual.
Prospectiva	Quais cenários são plausíveis sob determinadas condições?	Transformar extrapolação em previsão determinística.

Fonte: Elaborado pelo autor (2026).

Neste capítulo, a engenharia de *prompts* é apresentada como prática de interrogação temporal. Um *prompt* bem desenhado pode forçar a IA a separar dados observados, padrões estatísticos, possíveis explicações, hipóteses alternativas e evidências necessárias para confirmação. Pode também solicitar que a resposta indique nível de incerteza, risco de viés, possível artefato de busca e limite de inferência. A IA, nesse sentido, funciona como interlocutor metodológico: organiza possibilidades, explicita cautelas e ajuda a transformar resultados numéricos em narrativa científica prudente. Contudo, o pesquisador permanece responsável pela definição dos recortes, pela validação dos dados, pela escolha dos indicadores, pela revisão das interpretações e pela decisão sobre o que pode ou não ser afirmado (WEI *et al.*, 2022; WHITE *et al.*, 2023; SAHOO *et al.*, 2024).

13.2 Crescimento da literatura

O crescimento da literatura científica é uma das questões mais antigas e persistentes da cientometria. Desde os estudos clássicos sobre a expansão da ciência moderna, pesquisadores procuram compreender em que ritmo a produção científica se multiplica, como esse crescimento varia entre campos e quais fatores históricos, institucionais e tecnológicos explicam acelerações ou desacelerações. A contagem de publicações por ano parece, à primeira vista, um indicador simples. No entanto, ela concentra problemas relevantes: a definição do *corpus*, a cobertura da base de dados, a inclusão de diferentes tipos documentais, a mudança de nome de periódicos, a indexação retroativa, a expansão de revistas de acesso aberto, a multiplicação de *preprints* e a internacionalização desigual da publicação científica (BORNMANN; MUTZ, 2015; BORNMANN; HAUNSCHILD; MUTZ, 2021).

Em estudos cientométricos, a série anual de publicações costuma ser utilizada para caracterizar maturidade, difusão e dinâmica de uma área. Um tema com poucos registros durante décadas e crescimento acentuado em período recente pode ser interpretado como emergente, desde que a interpretação seja sustentada por outros sinais. Um tema com grande volume histórico e crescimento estável pode ser considerado consolidado. Um tema com crescimento rápido seguido de queda pode indicar moda, substituição terminológica,

saturação ou deslocamento para outro vocabulário. O problema é que essas leituras dependem da comparação com contextos. Uma área pequena pode crescer 200 por cento e continuar pouco expressiva em termos absolutos, uma área grande pode crescer modestamente e ainda assim mobilizar milhares de documentos e comunidades robustas.

A análise do crescimento da literatura deve combinar medidas absolutas e relativas. O número absoluto de documentos revela escala, enquanto taxas de crescimento indicam ritmo. O crescimento percentual anual composto permite comparar campos de tamanhos diferentes, mas pode exagerar mudanças quando a base inicial é pequena. A participação relativa de um tema dentro de uma área mais ampla ajuda a identificar se o crescimento é específico ou apenas reflexo da expansão geral da literatura. Por exemplo, se publicações sobre um tópico duplicam em dez anos, mas a literatura do campo também duplica, a tendência específica pode ser menos intensa do que parece. Por outro lado, se o tópico cresce enquanto o campo permanece estável, há sinal mais forte de deslocamento de atenção.

A IA pode apoiar esse processo ao transformar tabelas de frequência em narrativas interpretativas, sugerir comparações, apontar anomalias e lembrar possíveis artefatos. Uma tabela com anos e quantidades pode ser enviada a um modelo para que ele identifique fases: período inicial, expansão, aceleração, estabilização ou declínio. Contudo, a IA não deve inventar causas para as mudanças. O *prompt* precisa instruir o modelo a distinguir aquilo que é visível na série daquilo que constitui explicação hipotética. Uma boa resposta deveria afirmar, por exemplo, que há crescimento acelerado entre 2018 e 2022, mas que a explicação desse crescimento exige verificar eventos externos, mudanças na estratégia de busca, cobertura da base e entrada de novos periódicos.

A primeira decisão metodológica na análise de crescimento é a delimitação temporal. Períodos muito curtos tendem a confundir flutuação com tendência. Períodos muito longos podem esconder mudanças internas, especialmente quando uma área passa por transformações terminológicas. Em campos recentes, a série pode começar artificialmente no ano em que a base passou a indexar determinado tipo de documento. Em campos antigos, a cobertura retrospectiva pode ser incompleta. Assim, o pesquisador deve registrar ano inicial, ano final, justificativa do recorte e eventuais exclusões de anos incompletos. Em bases como Web of Science, Scopus, Dimensions, PubMed, Lens ou OpenAlex, o ano mais recente pode estar subindexado, o que exige cautela antes de interpretar uma queda aparente.

Outra decisão importante diz respeito ao tipo documental. Artigos de pesquisa, revisões, *proceedings papers*, editoriais, cartas e *preprints* apresentam dinâmicas diferentes. Em algumas áreas, congressos são veículos centrais de comunicação; em outras, a publicação em periódicos predomina. Uma série que mistura todos os tipos documentais pode sugerir crescimento, quando na verdade o aumento decorre apenas de revisões ou de anais de eventos. A IA pode ajudar a construir quadros comparativos por tipo documental e a redigir justificativas metodológicas, mas a classificação deve depender dos metadados da base e de regras explícitas. Se o objetivo é analisar a evolução de uma frente de pesquisa, talvez seja recomendável separar artigos originais de revisões, pois revisões frequentemente aparecem depois que uma literatura já se acumulou.

Também é necessário considerar a diferença entre crescimento de publicações e crescimento de conhecimento. A cientometria mede registros da comunicação científica, não a totalidade do conhecimento produzido. Uma área pode publicar mais porque há mais pesquisadores, mais financiamento, mais periódicos, maior pressão por produtividade ou maior fragmentação de resultados em múltiplos artigos. O aumento da literatura pode indicar vitalidade, mas também pode gerar redundância, sobrecarga informacional e dificuldade de síntese. A IA, ao resumir tendências, tende a apresentar crescimento como sinal positivo,

a menos que seja instruída a considerar alternativas. Por isso, *prompts* para crescimento devem solicitar interpretações equilibradas e explicitar que aumento quantitativo não equivale automaticamente a relevância, qualidade ou inovação.

Um protocolo robusto de análise de crescimento deve incluir: fonte de dados, estratégia de busca, filtros, data da coleta, campos utilizados, critérios de limpeza, contagem total e por ano, taxa de crescimento, comparação com *corpus* de referência, verificação de anos incompletos, verificação de tipos documentais e interpretação com incerteza. A IA pode gerar esse protocolo a partir de informações fornecidas pelo pesquisador, mas não deve preencher lacunas inexistentes. Quando faltarem dados sobre cobertura ou critérios de inclusão, o modelo deve indicar ausência de informação. Essa exigência evita que a escrita cientométrica produza uma narrativa fluente, mas metodologicamente frágil.

Quadro 13.2 - Indicadores para analisar crescimento da literatura

Indicador	Uso	Cautela
Número anual de documentos	Descrever escala e ritmo de publicação.	Sensível a cobertura da base e ao tipo documental.
Taxa de crescimento anual	Comparar períodos e identificar acelerações.	Pode exagerar campos pequenos.
Participação relativa	Verificar se o tema cresce dentro de um *corpus* maior.	Depende de denominador bem definido.
Janelas móveis	Suavizar flutuações anuais.	Pode ocultar rupturas curtas.
Comparação entre bases	Testar robustez do padrão.	Bases diferem em cobertura e metadados.

Fonte: Elaborado pelo autor (2026).

13.3 Ciclo de vida de temas científicos

Temas científicos raramente surgem plenamente formados. Em geral, passam por fases de formulação inicial, dispersão terminológica, acumulação de evidências, institucionalização, diversificação, estabilização e, em alguns casos, declínio ou reconfiguração. A noção de ciclo de vida de temas científicos permite analisar esses movimentos sem reduzir a ciência a uma sucessão linear. Um tema pode emergir em uma área, migrar para outra, perder visibilidade em um vocabulário e reaparecer com outro nome. Pode ainda permanecer como tema básico, sustentando pesquisas aplicadas, mesmo quando deixa de ocupar a fronteira mais visível da literatura.

A literatura de mapeamento científico frequentemente distingue temas motores, básicos, emergentes e declinantes por meio de indicadores de centralidade e densidade em mapas temáticos. Um tema motor tende a ser simultaneamente desenvolvido internamente e conectado ao restante do campo. Um tema básico pode ser central, mas menos especializado. Um tema emergente ou declinante pode apresentar baixa centralidade e baixa densidade, exigindo análise temporal para distinguir ascensão de desaparecimento. Essa distinção é metodologicamente importante: um mapa estático pode colocar emergentes e declinantes no mesmo quadrante, mas apenas a comparação entre períodos permite compreender a direção do movimento (COBO *et al.*, 2011).

O ciclo de vida de um tema pode ser examinado por diferentes sinais: crescimento de publicações, crescimento de citações, aumento de palavras-chave, surgimento de conferências especializadas, criação de periódicos, formação de grupos de pesquisa, entrada de financiadores,

diversidade geográfica, coocorrência com conceitos vizinhos e consolidação de revisões de literatura. Nenhum desses sinais, isoladamente, é suficiente. Uma frente pode receber muitas publicações e poucas citações; outra pode ter poucas publicações e alto impacto; uma terceira pode ser altamente visível na mídia, mas ainda fraca na literatura revisada por pares. A IA pode auxiliar na integração desses sinais, mas deve ser orientada a indicar evidências convergentes e divergentes.

A fase embrionária de um tema costuma ser marcada por baixa frequência, vocabulário instável e dispersão entre áreas. Nesse momento, estratégias de busca baseadas apenas em termos atuais podem perder documentos precursores. A IA pode sugerir termos históricos, analogias conceituais e sinônimos, mas tais sugestões devem ser verificadas em tesauros, bases de dados e leitura especializada. Em uma área como inteligência artificial aplicada a cientometria, por exemplo, documentos anteriores podem usar expressões como *text mining*, *machine learning*, *automatic classification, topic modeling* ou *knowledge discovery* antes que a expressão *generative AI* se torne comum. O ciclo de vida terminológico nem sempre coincide com o ciclo de vida substantivo do problema.

Na fase de expansão, o tema passa a atrair mais autores, instituições e subtópicos. A coocorrência de palavras-chave torna-se mais densa, revisões aparecem e o campo começa a delimitar seus problemas. A análise temporal pode mostrar que termos antes periféricos passam a ocupar posição central. Nessa etapa, a IA pode ajudar a nomear subfases e construir narrativas, mas o pesquisador deve evitar teleologia. O fato de um tema crescer não significa que seu desenvolvimento era inevitável. Frequentemente, o crescimento depende de eventos contingentes: descoberta técnica, crise social, disponibilidade de dados, mudança regulatória, financiamento estratégico ou popularização de ferramentas.

A fase de institucionalização ocorre quando um tema passa a ter comunidades relativamente estáveis, autores recorrentes, redes de colaboração, periódicos ou conferências de referência e vocabulário reconhecível. Os indicadores de centralidade e densidade tendem a aumentar, mas isso varia por área. Em campos fortemente interdisciplinares, um tema pode ser institucionalizado sem formar comunidade fechada; em campos especializados, pode haver alta densidade interna e baixa articulação externa. A IA pode comparar descrições de diferentes períodos e sugerir se há sinais de institucionalização, desde que a resposta seja apresentada como inferência condicionada aos dados, não como classificação definitiva.

O declínio de um tema é igualmente complexo. Uma queda no número de publicações pode indicar esgotamento, mas também substituição terminológica, migração para aplicações, incorporação do tema a práticas rotineiras ou deslocamento para literatura cinzenta, patentes e documentos técnicos. Muitos conceitos deixam de aparecer como palavras-chave porque se tornam pressupostos da área. Nesse caso, a diminuição de frequência não significa irrelevância. A IA deve ser instruída a verificar se o tema desapareceu ou se foi absorvido por outros termos. Uma análise de declínio responsável inclui busca por sinônimos recentes, exame de citações e comparação com áreas adjacentes.

A perspectiva de ciclo de vida também ajuda a interpretar temas recorrentes, revividos ou reconfigurados. Conceitos podem ressurgir após décadas quando novas tecnologias permitem sua operacionalização. Ideias antigas podem tornar-se relevantes sob novas condições empíricas. A análise temporal assistida por IA pode identificar ressurgimentos lexicais e rupturas semânticas, mas deve considerar que palavras iguais podem significar coisas diferentes em épocas diferentes. Por isso, a leitura de amostras de documentos de cada período continua indispensável. O ciclo de vida não é apenas gráfico: ele é também semântico, institucional e epistêmico.

Quadro 13.3 - Fases possíveis do ciclo de vida temático

Fase	Sinais cientométricos	Interpretação cautelosa
Embrionária	Baixa frequência, termos dispersos, poucos grupos.	Pode haver literatura precursora sob outros nomes.
Emergente	Crescimento relativo, novos autores, *clusters* iniciais.	Exige distinguir novidade substantiva de hype.
Expansão	Aumento de publicações, citações e subtemas.	Crescimento não equivale a qualidade.
Institucionalização	Periódicos, eventos, redes e vocabulário estável.	Pode variar por disciplina e infraestrutura.
Declínio ou reconfiguração	Queda de termos ou migração semântica.	Pode ser substituição terminológica, não desaparecimento.

Fonte: Elaborado pelo autor (2026).

13.4 Detecção de temas emergentes

A detecção de temas emergentes é uma das aplicações mais valorizadas da cientometria temporal. Gestores de pesquisa, agências de fomento, universidades, empresas e grupos científicos desejam identificar precocemente áreas em expansão, oportunidades de colaboração, lacunas de conhecimento e possíveis linhas estratégicas. Contudo, o termo emergente deve ser usado com rigor. Um tema emergente não é simplesmente um tema recente; é um tema que apresenta sinais de crescimento, diferenciação, articulação e potencial de influência em uma comunidade ou em várias comunidades científicas. A emergência envolve tanto aumento de atenção quanto formação de estrutura (KLEINBERG, 2003; CHEN, 2006; MANE; BÖRNER, 2004; SMALL, 2006).

Existem diferentes estratégias para detectar emergência. A primeira é a análise de crescimento abrupto de frequência, seja de publicações, palavras-chave, termos em resumos ou citações. A segunda é a detecção de *bursts*, isto é, períodos em que uma referência, autor, termo ou palavra-chave recebe aumento intenso de atenção em relação ao seu comportamento anterior. A terceira é a análise de redes temporais, observando o surgimento de novos *clusters* ou a mudança de centralidade de *clusters* existentes. A quarta é a modelagem de tópicos, especialmente quando se deseja detectar mudanças sem depender apenas de palavras-chave fornecidas por autores. A quinta é a análise de sinais externos, como financiamento, patentes, *preprints*, ensaios clínicos, normas ou políticas públicas (KLEINBERG, 2003; CHEN, 2006; MANE; BÖRNER, 2004).

A IA generativa pode apoiar essas estratégias em pelo menos quatro momentos. Primeiro, pode ajudar a formular hipóteses sobre possíveis temas emergentes a partir de listas de termos, *clusters* ou títulos recentes. Segundo, pode comparar períodos e descrever mudanças no vocabulário. Terceiro, pode revisar interpretações, perguntando se a evidência apresentada é suficiente para afirmar emergência. Quarto, pode produzir narrativas de resultado com linguagem cautelosa. Entretanto, a IA não deve identificar emergência apenas por plausibilidade semântica. Um modelo pode considerar um tema importante porque ele soa atual, não porque os dados do *corpus* demonstram crescimento ou reorganização. O *prompt* deve, portanto, restringir a interpretação aos dados fornecidos.

Um erro frequente é confundir tendência cientométrica com tendência social ou midiática. Um assunto pode estar amplamente presente na imprensa e ainda não ter se consolidado como

literatura científica. O inverso também ocorre: um tema pode ser altamente ativo na pesquisa, mas pouco visível publicamente. A análise cientométrica mede sinais na comunicação científica indexada, não necessariamente na sociedade. Quando o pesquisador deseja conectar esses planos, precisa integrar fontes adicionais, como altmetria, políticas públicas, notícias, patentes ou bases de financiamento. A IA pode sugerir triangulações, mas a evidência deve ser mantida separada: crescimento em artigos, crescimento em mídia e crescimento em financiamento são fenômenos relacionados, mas distintos.

Outro risco é a inflação de emergência. Em relatórios estratégicos, há tendência a classificar qualquer tópico recente como emergente, especialmente quando a finalidade é justificar investimentos. A cientometria responsável exige critérios predefinidos: por exemplo, crescimento mínimo em determinado período, aumento de participação relativa, surgimento de *cluster* próprio, diversidade de autores, diversidade institucional e citações acima de um limiar. Esses critérios podem variar conforme o campo, mas devem ser explicitados. A IA pode ajudar a construir uma matriz de critérios, mas não deve substituir a decisão teórica e política sobre o que conta como emergência relevante.

A detecção de temas emergentes também exige cuidado com o ano final da série. Em muitas bases, os anos mais recentes estão incompletos. Um tema pode parecer em queda simplesmente porque a coleta ocorreu antes do fechamento do ano ou porque a indexação ainda não se completou. A interpretação deve mencionar essa limitação. Uma prática recomendável e excluir o ano corrente de cálculos de crescimento ou marcá-lo como parcial. Outra prática e comparar a proporção de documentos indexados no ano corrente com históricos de atrasos da base. A IA, ao interpretar séries, deve ser instruída a perguntar se o último ano e completo antes de concluir declínio (KLEINBERG, 2003; CHEN, 2006; MANE; BÖRNER, 2004; SMALL, 2006).

A combinação de métodos aumenta a robustez. Um tema emergente é mais convincente quando apresenta crescimento em publicações, aumento em termos de resumo, formação de *cluster*, entrada de novos autores, citações recentes e conexão com problemas de pesquisa reconhecíveis. Quando apenas um sinal aparece, a interpretação deve ser moderada. Por exemplo, um termo pode crescer porque se tornou moda retórica em títulos, mas ainda não corresponde a comunidade ou método distinto. A IA pode ser utilizada como ferramenta de auditoria: a partir de uma afirmação como "o tema X e emergente", o *prompt* pode solicitar que o modelo liste evidências necessárias, evidências ausentes e explicações alternativas.

A emergência deve ser situada no *corpus*. Um tema pode ser emergente em uma área nacional e consolidado internacionalmente; emergente em uma disciplina e antigo em outra; emergente em periódicos e maduro em patentes; emergente em países periféricos e estabilizado em centros tradicionais. Por isso, a detecção de tendências exige explicitar a escala: tema emergente em que base, em que período, em que idioma, em que comunidade e em relação a qual universo de comparação. A IA pode ajudar a redigir essa contextualização, evitando frases universalizantes como "o tema está em ascensão na ciência" quando os dados sustentam apenas "o tema cresceu no *corpus* analisado".

Quadro 13.4 - Evidências para sustentar a emergência de um tema

Evidência	Força do sinal	Teste recomendado
Crescimento de documentos	Moderada quando isolada.	Comparar com crescimento do campo.
Burst de termos ou referências	Forte quando acompanhado de leitura substantiva.	Verificar contexto e duração do *burst*.

Evidência	Força do sinal	Teste recomendado
Novo *cluster* em rede	Forte quando persistente entre períodos.	Testar parâmetros de normalização e *cluster*ing.
Diversificação institucional	Moderada a forte.	Verificar desambiguação e contagem fracionada.
Revisões recentes	Sinal de acumulação.	Diferenciar revisão de consolidação empírica.

Fonte: Elaborado pelo autor (2026).

13.5 Mudanças terminológicas ao longo do tempo

A ciência muda também por meio das palavras que utiliza. Termos aparecem, desaparecem, se tornam genéricos, são substituídos, apropriados ou ressignificados. Em estudos cientométricos, a mudança terminológica é uma das principais fontes de erro na análise temporal. Quando um pesquisador usa apenas a denominação atual de um tema, pode ignorar documentos anteriores que empregavam outro vocabulário. Quando não controla sinônimos, pode fragmentar uma tendência em múltiplas séries artificiais. Quando agrupa termos sem critério, pode fundir problemas distintos sob uma categoria excessivamente ampla. A IA pode ajudar a mapear variantes terminológicas, mas também pode inventar sinônimos inadequados, por isso, sua contribuição precisa ser validada.

Mudanças terminológicas podem ocorrer por diferentes mecanismos. Há substituições lexicais, quando um termo é progressivamente trocado por outro. Há expansões semânticas, quando um conceito passa a cobrir um conjunto maior de fenômenos. Há especializações, quando um termo antes amplo se torna associado a técnica ou subcampo particular. Há rebranding, quando comunidades adotam nova denominação para reposicionar uma agenda. Há tradução interdisciplinar, quando a mesma ideia circula por áreas com vocabulários distintos. Há ainda mudanças institucionais, quando descritores controlados em bases como MeSH ou thesauri são atualizados, afetando a recuperação bibliográfica.

A análise temporal de termos pode ser feita por frequência anual de palavras-chave, coocorrência de termos, *embedding* semântico, modelos de tópicos ou leitura comparativa de títulos e resumos. A escolha depende da pergunta. Se o objetivo é verificar quando um termo entrou no vocabulário da área, uma série de frequência pode bastar. Se o objetivo é compreender mudança conceitual, é necessário analisar contexto de uso. Termos iguais podem se deslocar semanticamente: "rede neural" nos anos 1990 não ocupa exatamente o mesmo campo técnico e institucional que "*deep learning*" nos anos 2010. A IA pode resumir contextos de uso de amostras temporais, mas a interpretação deve ser conferida por especialistas (BLEI; NG; JORDAN, 2003; BLEI; LAFFERTY, 2006).

O uso de palavras-chave de autores apresenta vantagens e limites. Elas revelam como os próprios autores rotulam seus trabalhos, mas são heterogêneas, sujeitas a moda e variam em granularidade. Palavras-chave de indexadores podem ser mais controladas, mas nem sempre capturam novidade rapidamente. Títulos são concisos e expressam foco, mas podem simplificar. Resumos oferecem maior contexto, mas exigem processamento textual mais complexo. Um protocolo temporal robusto deve combinar campos quando possível e explicar qual deles foi utilizado. A IA pode ajudar a comparar resultados obtidos por diferentes campos e identificar inconsistências.

A mudança terminológica é particularmente importante na detecção de declínio. Um termo em queda pode ter sido substituído por outro mais atual. Por exemplo, em diferentes domínios,

expressões relacionadas a automação, inteligência computacional, aprendizagem de máquina e inteligência artificial podem se sobrepor parcialmente, mas não são equivalentes. Em cientometria, termos como *bibliometric mapping, science mapping, knowledge domain visualization* e *research front analysis* podem designar práticas relacionadas com ênfases distintas. Agrupar tudo pode ampliar indevidamente o *corpus*; separar tudo pode perder continuidade histórica. A decisão deve ser fundamentada por leitura conceitual e testes de busca.

Prompts para mudança terminológica devem solicitar agrupamentos com justificativa e grau de confiança. Não basta pedir "liste sinônimos". O modelo deve diferenciar sinônimo, termo relacionado, termo mais amplo, termo mais específico, denominação histórica e tradução. Deve marcar casos ambíguos para revisão humana. Também deve ser instruído a não criar termos inexistentes. Uma boa prática é fornecer ao modelo apenas termos extraídos do *corpus*, pedindo que os organize, em vez de solicitar que imagine vocabulário. Quando termos sugeridos pela IA forem adicionados a estratégia de busca, devem ser testados separadamente para verificar ruído e recuperação.

A comparação entre períodos pode revelar mudanças interessantes. Um termo pode aparecer inicialmente associado a métodos, depois a aplicações e finalmente a políticas públicas. Outro pode migrar de uma área técnica para discussões éticas. Uma coocorrência antes rara pode se tornar central, indicando convergência entre agendas. A IA pode construir narrativas dessas transições a partir de matrizes temporais de coocorrência, desde que os dados sejam apresentados com clareza. O pesquisador deve exigir que a resposta indique quais termos sustentam cada afirmação e quais alternativas interpretativas permanecem abertas.

Finalmente, a mudança terminológica tem implicações para reprodutibilidade. Uma estratégia de busca formulada hoje pode não recuperar a mesma paisagem daqui a alguns anos, se bases atualizarem descritores, ampliarem indexação ou normalizarem metadados. Portanto, o registro da estratégia deve incluir *strings* exatas, campos pesquisados, data da coleta, thesauri consultados, decisões de tradução, termos excluídos e justificativas para agrupamentos. A IA pode transformar esses elementos em protocolo, mas o pesquisador deve preservar versões originais. Em estudos temporais, o vocabulário não é apenas instrumento de busca, é parte do objeto analisado.

Quadro 13.5 - Tipos de relação terminológica em estudos temporais

Relação	Descrição	Ação metodológica
Sinônimo	Termos intercambiáveis no *corpus*.	Agrupar com justificativa.
Termo mais amplo	Categoria que inclui subtemas.	Não agrupar automaticamente.
Termo mais específico	Subconjunto de um conceito maior.	Analisar como subsérie.
Termo histórico	Denominação usada em período anterior.	Incluir em busca longitudinal.
Termo relacionado	Conceito vizinho, mas não equivalente.	Manter separado e analisar coocorrência.

Fonte: Elaborado pelo autor (2026).

13.6 Trajetórias de autores, instituições e países

A análise temporal não se limita a temas. Autores, instituições e países também apresentam trajetórias. Um autor pode iniciar como colaborador periférico, tornar-se líder de um *cluster* e depois atuar como ponte entre comunidades. Uma instituição pode entrar tardiamente em

uma área e crescer rapidamente por meio de políticas de contratação ou financiamento. Um país pode ampliar participação internacional, diversificar colaborações ou concentrar-se em determinados subtemas. Essas trajetórias revelam a dimensão social e institucional da ciência, complementando a análise cognitiva de conceitos e frentes de pesquisa.

A trajetória de autores deve ser interpretada com cuidado. Produtividade anual, citações acumuladas, participação em redes de coautoria e mudança de temas podem indicar percurso científico, mas também são afetadas por carreira, área, orientação de estudantes, mudança institucional, colaborações de grande escala e práticas autorais. Em campos com hiperautoria, a contagem de publicações por autor pode ser pouco informativa sem fracionamento. Em áreas de humanidades, livros e capítulos podem ser mais relevantes que artigos indexados. Assim, a IA deve ser instruída a analisar autores apenas dentro do contexto disciplinar e dos metadados disponíveis (NEWMAN, 2001).

Trajetórias institucionais exigem desambiguação rigorosa. Nomes de universidades, departamentos, hospitais, centros de pesquisa e agências podem aparecer em múltiplas variantes. Fusões, mudanças de nome e transliterações complicam séries temporais. Uma instituição pode parecer crescer simplesmente porque variantes antes separadas foram unificadas na limpeza de dados. O contrário também ocorre: uma instituição pode parecer fragmentada quando a base registra unidades internas de forma inconsistente. A IA pode auxiliar na normalização e no resumo de trajetórias, mas a desambiguação deve ser documentada e, quando possível, apoiada por identificadores institucionais.

A análise por países apresenta problemas próprios. A atribuição de países geralmente depende das afiliações dos autores, não necessariamente do local onde a pesquisa foi concebida, financiada ou realizada. Trabalhos com múltiplos países podem ser contados de forma inteira ou fracionada. A contagem inteira valoriza presença, a fracionada distribui peso. Em séries temporais, a escolha altera conclusões sobre crescimento e participação. Além disso, países diferem em tamanho, investimento em pesquisa, idioma, políticas de publicação e cobertura em bases internacionais. Interpretar crescimento nacional sem normalizar por contexto pode produzir rankings simplistas (LEYDESDORFF; RAFOLS, 2009).

A trajetória relacional é tão importante quanto a trajetória quantitativa. Uma instituição pode publicar pouco, mas ocupar posição de intermediação em uma rede internacional. Um país pode ter alta produção, mas baixa diversidade de colaborações. Um autor pode ter poucas publicações recentes, mas alta influência histórica. Indicadores temporais de centralidade, modularidade e participação em *clusters* ajudam a compreender essas dinâmicas. A IA pode descrever mudanças de posição em rede, mas não deve inferir estratégias institucionais sem evidência. Uma frase como "a instituição assumiu papel de liderança" precisa ser sustentada por dados relacionais e, idealmente, por informações contextuais externas.

A análise de trajetórias também deve diferenciar coorte, período e idade acadêmica. Autores que entram em uma área em 2020 não podem ser comparados diretamente com autores ativos desde 2000 sem considerar tempo de exposição. Instituições novas podem apresentar altas taxas de crescimento por partir de bases pequenas. Países que ampliaram indexação em periódicos internacionais podem parecer mais produtivos sem necessariamente representar crescimento proporcional da pesquisa nacional. A IA pode ajudar a identificar esses efeitos e sugerir controles, como normalização por primeiro ano de aparição, janelas móveis ou comparação por coortes.

A dimensão ética é relevante. Análises temporais de autores podem ser usadas para avaliação individual, contratação, promoção ou distribuição de recursos. O uso inadequado de séries cientométricas pode reforçar desigualdades, penalizar interrupções de carreira, ignorar

maternidade, deslocamentos institucionais, trabalho coletivo invisível e diferenças disciplinares. A IA não deve ser usada para produzir julgamentos automáticos sobre trajetórias individuais. Em vez disso, pode apoiar análises descritivas e contextualizadas, sempre acompanhadas de advertências metodológicas e de critérios de uso responsável.

No caso de instituições e países, a análise temporal pode apoiar política científica, identificando áreas de crescimento, lacunas de colaboração, dependência de poucos parceiros, concentração regional ou oportunidades de internacionalização. Contudo, também pode induzir comportamento estratégico voltado a indicadores. Se gestores utilizam apenas crescimento de publicações como meta, podem estimular produção quantitativa em detrimento de qualidade, relevância ou integridade. Por isso, *prompts* para trajetórias institucionais devem solicitar que a IA apresente indicadores, limitações, possíveis efeitos perversos e alternativas de avaliação qualitativa.

Quadro 13.6 - Cuidados na interpretação de trajetórias científicas

Unidade	Indicadores úteis	Principais cautelas
Autor	Produção anual, citações, coautoria, mudança de temas.	Tempo de carreira, hiperautoria, interrupções e área.
Instituição	Produção, colaboração, centralidade, especialização.	Desambiguação, fusões, variantes e unidades internas.
País	Participação, colaboração internacional, liderança em *clusters*.	Tamanho do sistema científico, idioma, investimento e cobertura.

Fonte: Elaborado pelo autor (2026).

13.7 Cenários prospectivos com cautela metodológica

A análise temporal frequentemente leva a perguntas prospectivas. Se um tema cresce rapidamente, continuará crescendo? Se um país aumenta sua participação, tornar-se-á líder? Se um termo emergiu nos últimos anos, consolidar-se-á como campo? A cientometria pode oferecer insumos para cenários, mas não deve prometer previsões determinísticas. A ciência e influenciada por fatores imprevisíveis: descobertas, crises, políticas, guerras, pandemias, tecnologias disruptivas, mudanças regulatórias, movimentos sociais e transformações na comunicação científica. O futuro da ciência não é uma simples extrapolação de curvas passadas.

Cenário prospectivo é diferente de predição pontual. Uma predição afirma valor esperado: determinado número de publicações, citações ou colaborações em ano futuro. Um cenário descreve possibilidades plausíveis sob determinadas condições. Em estudos cientométricos responsáveis, a IA pode ajudar a formular cenários condicionais, como continuidade, aceleração, saturação, diversificação ou substituição terminológica. Cada cenário deve explicitar premissas, sinais que o sustentam, incertezas e indicadores de monitoramento. Essa abordagem é mais prudente do que afirmar que uma área "será dominante" ou que determinado tema "está destinado" a crescer.

A formulação de cenários deve partir de evidências históricas, mas também reconhecer limites. Uma série temporal curta não permite inferências robustas sobre tendência de longo prazo. Uma série longa pode incluir regimes diferentes que não deveriam ser projetados como se fossem homogêneos. Crescimento exponencial raramente se mantém indefinidamente; áreas podem saturar, dividir-se em subcampos ou perder financiamento. Temas muito recentes podem apresentar crescimento artificial por *hype* inicial. A IA pode ser instruída a formular

cenários apenas após listar o tamanho da série, a estabilidade dos dados, os possíveis pontos de ruptura e a qualidade da evidência.

O uso de modelos estatísticos para previsão, como regressão, modelos de séries temporais ou aprendizado de máquina, requer dados consistentes e perguntas bem delimitadas. A IA generativa, por si só, não calcula previsões confiáveis a menos que receba dados e instruções adequadas ou seja integrada a ferramentas estatísticas. Mesmo quando calcula, a interpretação deve considerar incerteza, intervalo de confiança, escolha do modelo e validação fora da amostra. Em capítulos anteriores, discutiu-se que *prompts* podem transformar tarefas analíticas em procedimentos documentados. No caso prospectivo, essa documentação é ainda mais importante, porque a linguagem de futuro tende a amplificar autoridade indevida.

A prospectiva cientométrica deve evitar o viés de novidade. Nem tudo que cresce rapidamente é transformador; nem tudo que é antigo é irrelevante. Estudos sobre ciência da ciência mostram que impacto, colaboração, novidade e consolidação possuem dinâmicas complexas. Em algumas áreas, temas básicos sustentam avanços importantes sem aparecer como tendência emergente. Em outras, um tópico muito visível pode gerar grande volume de literatura de baixa originalidade. A IA deve ser solicitada a apresentar tanto oportunidades quanto riscos: saturação, dependência de poucos grupos, fragilidade metodológica, excesso de publicação, baixa reprodutibilidade ou ausência de aplicações (SMALL, 2006; PARK; LEAHEY; FUNK, 2023).

Cenários prospectivos também devem considerar governança. Ao identificar tendências, relatórios cientométricos podem influenciar decisões de financiamento, contratação e estratégia institucional. Se as análises privilegiam apenas áreas em crescimento, podem reforçar desigualdades e marginalizar temas essenciais, mas menos populares. Se usam IA sem transparência, podem dar aparência técnica a escolhas políticas. Por isso, cenários devem ser apresentados como insumos para deliberação, não como comandos. O *prompt* adequado deve solicitar que o modelo indique implicações, mas também alternativas, *trade-offs* e questões éticas.

Quadro 13.7 - Diferença entre previsão, projeção e cenário

Termo	Característica	Uso responsável
Previsão	Estimativa de valor futuro por modelo.	Exige dados, modelo, validação e incerteza.
Projeção	Extrapolação condicional de tendência.	Depende de premissas explícitas.
Cenário	Narrativa plausível sob condições.	Serve para deliberação, não para certeza.
Hipótese prospectiva	Proposição verificável sobre evolução futura.	Deve conter sinais de confirmação e refutação.

Fonte: Elaborado pelo autor (2026).

Uma boa prática é formular hipóteses prospectivas testáveis. Em vez de escrever "o tema X dominara a área", o relatório pode afirmar: "se o crescimento relativo de X se mantiver por três anos, se novos *clusters* continuarem a surgir e se houver diversificação institucional, então X poderá consolidar-se como subcampo reconhecível". Essa fórmula transforma previsão vaga em condição verificável. A IA pode apoiar essa tradução, convertendo tendências observadas em hipóteses condicionais e indicadores de acompanhamento. O pesquisador, por sua vez, deve revisar se as condições são pertinentes e mensuráveis.

Por fim, a cautela metodológica deve aparecer na própria redação. Verbos como sugerir, indicar, apontar, ser compatível com e levantar a hipótese são mais adequados que provar, demonstrar definitivamente ou garantir. A linguagem cientométrica responsável reconhece que tendências são sinais probabilísticos, dependentes de *corpus*, método e contexto. A IA pode ajudar a reescrever conclusões em tom prudente, mas deve receber essa instrução explicitamente. Sem ela, modelos generativos tendem a produzir narrativas excessivamente assertivas. A prudência, neste capítulo, não é timidez interpretativa, é condição de validade científica.

13.8 O papel da inteligência artificial na análise temporal

A inteligência artificial aplicada à análise temporal deve ser compreendida como camada de apoio interpretativo, organizacional e redacional. Ela não substitui técnicas estatísticas, algoritmos de mapeamento ou conhecimento especializado. Sua contribuição principal está em transformar resultados dispersos em perguntas, checklist*s*, narrativas, hipóteses e protocolos. Quando o pesquisador possui tabelas anuais, *clusters* temporais, listas de termos e indicadores de crescimento, a IA pode ajudar a sintetizar padrões e a indicar pontos que merecem verificação. Essa função é especialmente valiosa em estudos com muitos subtemas e múltiplos períodos.

Um uso produtivo consiste em solicitar que a IA compare períodos de forma estruturada. Em vez de perguntar genericamente "quais são as tendências?", o pesquisador pode fornecer uma matriz com frequências por período e pedir uma resposta separada em crescimento absoluto, crescimento relativo, estabilidade, declínio, mudança terminológica e possíveis artefatos. Essa estrutura reduz a chance de conclusões impressionistas. Também permite que a saída seja comparada com cálculos independentes. A IA deve ser tratada como assistente de leitura dos resultados, não como fonte dos resultados.

Outro uso relevante é a construção de narrativas longitudinalmente consistentes. Estudos cientométricos frequentemente produzem muitos gráficos, tabelas e mapas, mas a dificuldade está em integrar esses produtos em uma história analítica. A IA pode ajudar a organizar a narrativa em fases: emergência, expansão, especialização, diversificação e estabilização. Pode sugerir transições entre seções, formular sumários executivos e adaptar a comunicação para gestores. Ainda assim, a narrativa deve permanecer ancorada em dados. O *prompt* deve exigir que cada afirmação substantiva seja associada a um indicador ou evidência fornecida.

A IA também pode funcionar como auditor de extrapolações. O pesquisador pode submeter uma conclusão preliminar ao modelo e pedir que ele identifique afirmações fortes demais, lacunas de evidência, possíveis vieses de base e alternativas explicativas. Esse uso crítico é frequentemente mais seguro do que pedir ao modelo que produza a interpretação inicial. Quando a IA é usada como revisora metodológica, ela ajuda a tornar explícitas as fragilidades que a fluidez da escrita pode esconder. A auditoria por *prompt*, entretanto, não substitui revisão por pares nem leitura substantiva da literatura.

A integração entre IA generativa e ferramentas de análise quantitativa deve ser documentada. Se os dados foram processados em R, Python, VOSviewer, CiteSpace, SciMAT ou bibliometrix, e a IA foi usada apenas para interpretar tabelas e redigir textos, isso deve ficar claro. Se a IA foi usada para classificar documentos, agrupar termos ou rotular *clusters*, a metodologia deve indicar modelo, data, versão quando disponível, *prompt*, parâmetros, amostra revisada e taxa de concordância humana. A transparência é particularmente importante porque modelos generativos podem mudar ao longo do tempo, gerando respostas diferentes para *prompts* semelhantes (WEI *et al.*, 2022; KOJIMA *et al.*, 2022; WHITE *et al.*, 2023).

A análise temporal assistida por IA deve preservar a possibilidade de replicação parcial. Ainda que a resposta de um modelo não seja perfeitamente reprodutível, o protocolo pode registrar entradas, *prompts*, critérios e saídas. A reprodutibilidade, nesse caso, não significa obter exatamente a mesma frase, mas permitir que outro pesquisador compreenda como a interpretação foi produzida, quais dados a sustentaram e quais decisões humanas foram tomadas. Essa perspectiva aproxima a engenharia de *prompts* de um diário metodológico: um registro das perguntas feitas ao sistema, das respostas recebidas e das validações aplicadas.

13.9 Diferenciando tendência real, flutuação e artefato de busca

Um dos desafios centrais da análise temporal e diferenciar tendência real de flutuação aleatória ou artefato metodológico. Tendência real, no contexto cientométrico, não significa necessariamente realidade objetiva independente dos dados, mas padrão consistente no *corpus* analisado, sustentado por múltiplos sinais e robusto a variações razoáveis de método. Flutuação e variação instável, geralmente sensível a poucos documentos ou anos. Artefato de busca e padrão produzido por decisão técnica: *string* mal formulada, mudança de descritor, filtro documental, cobertura da base, duplicatas, erros de metadados ou exclusão de sinônimos (CHEN, 2006; KLEINBERG, 2003).

A primeira estratégia para diferenciar esses fenômenos é testar robustez. Uma tendência que aparece apenas em uma *string* estreita, mas desaparece quando sinônimos são incluídos, deve ser interpretada com cautela. Um crescimento que ocorre apenas em uma base pode refletir cobertura, e não necessariamente dinâmica do campo. Uma queda concentrada no ano mais recente pode ser atraso de indexação. Um pico isolado pode decorrer de chamada especial de periódico, conferência, relatório temático ou crise conjuntural. A IA pode listar essas possibilidades, mas precisa receber dados suficientes: *string*, base, período, tipos documentais e frequências.

A segunda estratégia é triangulação. Tendências mais confiáveis costumam aparecer em diferentes indicadores: crescimento de documentos, aumento de termos nos resumos, citações recentes, novos autores, diversificação institucional e coocorrências estáveis. Quando apenas um indicador aponta mudança, a interpretação deve ser apresentada como sinal preliminar. A triangulação não elimina incerteza, mas reduz dependência de um único artefato. *Prompts* de revisão devem solicitar explicitamente a separação entre evidências convergentes é evidências ausentes.

A terceira estratégia é leitura de amostras. Análise quantitativa sem leitura qualitativa pode confundir termo com tema. Ao detectar crescimento de uma expressão, é necessário verificar se os documentos realmente tratam do fenômeno esperado. Uma palavra pode ser usada de modo periférico, metafórico ou diferente entre campos. A IA pode resumir amostras, mas o pesquisador deve revisar casos duvidosos. A amostragem deve incluir anos de crescimento, anos anteriores e documentos de diferentes *clusters*. Essa prática ajuda a distinguir mudança substantiva de ruído terminológico.

A quarta estratégia é normalização por campo ou *corpus* de referência. Se toda a literatura de uma área cresce por causa de aumento geral de publicações, um tema pode parecer ascendente sem ganhar participação relativa. A comparação com um denominador adequado é essencial. Esse denominador pode ser a produção total da área, a produção da base, a produção de um conjunto de periódicos ou o total de documentos recuperados por uma estratégia ampla. A IA pode explicar a diferença entre crescimento absoluto e relativo para leitores não especialistas, evitando interpretações exageradas.

Por fim, deve-se registrar incerteza. Em vez de ocultar limites, o relatório temporal deve explicitar quais conclusões são fortes, moderadas ou preliminares. A IA pode ser solicitada a classificar afirmações por nível de confiança, mas essa classificação deve derivar de critérios fornecidos pelo pesquisador. Uma conclusão com alta confiança deve ter base temporal suficiente, sinais convergentes e baixa sensibilidade a mudanças metodológicas. Uma conclusão preliminar deve ser marcada como tal. Essa prática torna a análise mais honesta é mais útil para tomada de decisão.

13.10 Prompts do capítulo

Os *prompts* apresentados a seguir foram desenhados para transformar a análise temporal em um procedimento auditável. Eles podem ser usados com tabelas, descrições de gráficos, matrizes de termos, resultados de softwares de mapeamento ou resumos de *clusters*. O princípio geral é que a IA deve interpretar apenas os dados fornecidos, explicitar incertezas e separar descrição de explicação. *R*ecomenda-se registrar o *prompt*, a entrada, a saída, a data de uso, a versão do sistema quando disponível e as decisões humanas tomadas após a resposta.

O primeiro *prompt* interpreta séries temporais. Ele deve ser usado quando o pesquisador possui dados anuais de produção, citação, ocorrência de termos ou participação relativa. O segundo *prompt* detecta mudança temática, comparando períodos ou *clusters*. O terceiro *prompt* diferencia tendência real de artefato de busca. O quarto *prompt* formula hipóteses prospectivas, sempre em linguagem condicional. Em todos os casos, o pesquisador deve adaptar variáveis, períodos, *corpus* e indicadores ao seu estudo.

Prompt 13.10.1 Interpretar séries temporais

Atue como especialista em Cientometria. Analise a série temporal abaixo, referente a [indicador] no *corpus* [descrição do *corpus*], no período [ano inicial-ano final]. Separe a resposta em:

(1) descrição objetiva dos padrões observados;

(2) fases temporais possíveis;

(3) crescimento absoluto e relativo, quando os dados permitirem;

(4) anomalias, picos ou quedas;

(5) interpretações plausíveis;

(6) explicações alternativas;

(7) limites metodológicos;

(8) conclusão prudente em até 150 palavras.

Não invente causas externas.

Quando uma explicação depender de evidência adicional, marque-a como hipótese.

Dados: [colar tabela].

Uso recomendado: Use quando houver tabela anual de publicações, citações, termos, documentos por país ou indicadores de colaboração.

Prompt 13.10.2 Detectar mudança temática

Atue como analista de mapeamento científico. Compare os períodos [período A] e [período B] com base nos termos, *clusters* ou palavras-chave abaixo. Identifique:

(1) temas que cresceram;

(2) temas que diminuíram;

(3) temas novos;

(4) temas possivelmente substituídos por nova terminologia;

(5) temas estáveis;

(6) mudanças de coocorrência;

(7) possíveis artefatos de busca;

(8) amostras de documentos que deveriam ser lidas para validação.

Classifique cada achado como forte, moderado ou preliminar e justifique o grau de confiança.

Dados: [colar matriz ou lista por período].

Uso recomendado: Use em estudos de evolução de palavras-chave, mapas temáticos, *clusters* temporais e frentes de pesquisa.

Prompt 13.10.3 Diferenciar tendência real de artefato de busca

Atue como auditor metodológico de um estudo cientométrico. A afirmação preliminar é: [colar afirmação].

Os dados utilizados são: [colar dados, estratégia de busca, base, período e filtros].

Avalie se a afirmação pode decorrer de uma tendência real, de flutuação ou de artefato de busca.

Liste evidências favoráveis, evidências contrárias, informações ausentes, testes de robustez recomendados e uma versão reescrita da afirmação em linguagem metodologicamente prudente.

Não aceite a conclusão sem verificar alternativas.

Uso recomendado: Use antes de afirmar emergência, declínio, liderança, consolidação ou ruptura temática.

Prompt 13.10.4 Formular hipóteses prospectivas

Atue como especialista em prospectiva cientométrica. Com base nos resultados temporais abaixo, formule até quatro hipóteses prospectivas condicionais para os próximos [n] anos. Para cada hipótese, indique:

(1) premissa;

(2) evidências atuais;

(3) indicadores de acompanhamento;

(4) sinais que refutariam a hipótese;

(5) riscos de extrapolação;

(6) grau de incerteza.

Escreva em linguagem cautelosa, sem afirmar previsões determinísticas.

Dados: [colar resultados temporais].

Uso recomendado: Use para relatórios estratégicos, políticas científicas, planejamento institucional e monitoramento de temas emergentes.

Prompt 13.10.5 Reescrever conclusões temporais sem extrapolação

Revise o parágrafo abaixo como editor científico. Identifique frases que extrapolam os dados, transformam associação em causalidade, confundem crescimento com qualidade ou apresentam previsões sem evidência.

Em seguida, reescreva o texto em tom formal científico, mantendo as conclusões apenas no nível sustentado pelos dados.

Parágrafo: [colar texto].

Uso recomendado: Use na etapa final de redação de resultados e discussão.

13.11 Estudo de caso ilustrativo: a evolução de um tema cientométrico hipotético

Para ilustrar a aplicação dos conceitos deste capítulo, considere um estudo hipotético sobre "inteligência artificial generativa em estudos cientométricos" no período de 2018 a 2026. O *corpus* foi construído em bases bibliográficas internacionais com termos relacionados a *generative artificial intelligence*, *large language models*, ChatGPT, *bibliometrics*, *scientometrics*, *science mapping* e *research evaluation*. O objetivo não é apresentar resultados empíricos reais, mas demonstrar como a análise temporal poderia ser conduzida e interpretada.

Na fase inicial, entre 2018 e 2020, documentos recuperados provavelmente seriam poucos e poderiam usar termos mais amplos, como *machine learning*, *natural language processing*, *text mining* e *automated bibliometric analysis*. A expressão *generative AI* ainda não ocuparia papel central. Uma estratégia de busca baseada apenas em ChatGPT perderia quase toda a literatura precursora. A primeira decisão metodológica seria construir famílias terminológicas separadas: uma para métodos de IA, outra para modelos de linguagem e outra para objetos cientométricos. A IA poderia ajudar a organizar essas famílias, mas os termos deveriam ser testados nas bases.

Em um segundo período, de 2021 a 2023, surgiriam sinais de transição. Modelos de linguagem de grande escala se tornariam mais visíveis, e publicações passariam a discutir usos de IA generativa na escrita acadêmica, revisão de literatura, recuperação de informação e análise textual. Um pico em 2023 poderia refletir o impacto público de sistemas conversacionais. Entretanto, esse pico não deveria ser interpretado automaticamente como consolidação cientométrica. Seria preciso verificar se os documentos tratam efetivamente de cientometria ou apenas mencionam bibliometria em discussões gerais sobre pesquisa acadêmica.

Em um terceiro período, de 2024 a 2026, a hipótese de emergência poderia ser avaliada por sinais adicionais: crescimento de artigos específicos, formação de *clusters*, aparição de protocolos, estudos comparativos, discussões sobre alucinação em referências, ferramentas de automação bibliométrica e debates éticos. Se esses sinais convergissem, o relatório poderia afirmar que o tema apresenta indícios de consolidação inicial no *corpus* analisado. Se o crescimento fosse concentrado em comentários, editoriais e ensaios, a conclusão deveria ser mais cautelosa: há aumento de atenção, mas ainda não necessariamente uma área metodologicamente estabilizada.

A IA poderia apoiar a interpretação recebendo uma tabela com anos, contagens, tipos documentais e termos mais frequentes. Um *prompt* adequado solicitaria: identificar fases, apontar possíveis artefatos, distinguir literatura precursora de literatura focal e sugerir evidências necessárias para confirmar emergência. A resposta deveria evitar frases como "a IA generativa revolucionou a cientometria" se os dados indicarem apenas crescimento recente de discussões. A formulação prudente seria: "os dados sugerem intensificação da atenção ao uso de modelos generativos em tarefas cientométricas, mas a consolidação metodológica depende de evidências adicionais, como validação empírica, reprodutibilidade e adoção por comunidades especializadas".

Esse caso hipotético mostra que a análise temporal é inseparável da estratégia de busca, da interpretação terminológica e da validação humana. A IA pode acelerar a leitura dos sinais, mas também pode amplificar narrativas de novidade. Em temas associados a tecnologias populares, a cautela é ainda mais importante. A linguagem prospectiva deve ser condicional, e a detecção de tendências deve ser apoiada por critérios explícitos.

13.12 Checklist metodológico para análises temporais assistidas por IA

Antes de interpretar qualquer série temporal, o pesquisador deve verificar se a pergunta exige escala anual, períodos agregados ou janelas móveis. Séries anuais são intuitivas, mas podem ser ruidosas quando o *corpus* é pequeno. Períodos de três ou cinco anos reduzem flutuações, mas podem esconder rupturas. Janelas móveis permitem observar tendências suavizadas, mas introduzem dependência entre pontos. A escolha deve ser justificada pela natureza do campo e pelo volume de dados.

O checklist metodológico deve incluir também verificação da completude dos anos recentes, da consistência dos filtros, da presença de duplicatas, da normalização de termos, da comparabilidade entre bases e da existência de eventos externos relevantes. Em estudos com múltiplas bases, é necessário verificar se o crescimento aparece em todas ou apenas em uma. Em estudos com termos emergentes, é necessário comparar o termo atual com variantes históricas. Em estudos prospectivos, é necessário separar resultados observados de hipóteses.

Quadro 13.8 - Checklist de qualidade para análise temporal com IA

Etapa	Pergunta de controle	Evidência esperada
Corpus	A estratégia de busca e o período estão documentados?	String, base, data da coleta e filtros.
Série temporal	O ano recente está completo?	Nota metodológica sobre indexação.
Termos	Sinônimos e variantes históricas foram avaliados?	Tabela de normalização terminológica.

Etapa	Pergunta de controle	Evidência esperada
Robustez	A tendência foi testada por indicador ou base alternativa?	Comparação de resultados.
IA	*Prompt*, entrada e saída foram arquivados?	Registro de auditoria.
Interpretação	Conclusões separam dado, inferência e hipótese?	Texto revisado em linguagem prudente.

Fonte: Elaborado pelo autor (2026).

Na etapa de uso de IA, o checklist deve perguntar: quais dados foram fornecidos ao modelo? o modelo recebeu instruções para não inventar causas? a resposta separou descrição, explicação e hipótese? o nível de incerteza foi explicitado? houve revisão humana? a saída foi comparada com gráficos e tabelas originais? as conclusões foram reescritas em linguagem prudente? o *prompt* e a resposta foram arquivados? Essas perguntas transformam a IA de instrumento opaco em componente documentado do método.

Um checklist não garante qualidade, mas reduz omissões. Em análises temporais, muitas falhas decorrem de detalhes simples: ano incompleto, mudança de *string*, ausência de sinônimos, interpretação de pico isolado, confusão entre crescimento absoluto e relativo ou falta de normalização por campo. A IA pode ajudar a aplicar o checklist, mas a responsabilidade permanece com o pesquisador. O objetivo é produzir análises que possam ser discutidas, corrigidas e reproduzidas parcialmente por outros pesquisadores.

13.13 Protocolo resumido de reprodutibilidade para o capítulo

Em uma pesquisa empírica, a análise temporal deve ser acompanhada de um pacote mínimo de reprodutibilidade. Esse pacote não precisa expor dados restritos, mas deve permitir que outro pesquisador compreenda o caminho analítico. O primeiro componente é o registro da coleta: bases consultadas, estratégia de busca, filtros, campos exportados, data de coleta e versão da base quando disponível. O segundo componente é a preservação dos dados brutos, separados dos dados limpos. O terceiro é a documentação das transformações: deduplicação, normalização de autores, agrupamento de termos, exclusão de registros e tratamento de anos incompletos.

O quarto componente é o registro dos scripts ou ferramentas utilizadas. Quando forem empregados pacotes como bibliometrix, VOSviewer, CiteSpace ou SciMAT, recomenda-se registrar versão, parâmetros de normalização, limiares de inclusão, método de *cluster*ing e critérios de visualização. O quinto componente é o registro dos *prompts*. Em estudos assistidos por IA, *prompt*, entrada, saída e revisão humana devem ser tratados como parte do caderno de pesquisa. Isso não significa que toda resposta de IA tenha valor científico, mas que sua influência sobre a interpretação precisa ser rastreável.

Quadro 13.9 - Matriz de conclusões temporais

Conclusão	Evidência de suporte	Nível de confiança	Limite principal
O tema apresenta crescimento recente.	Série anual e participação relativa.	Moderado a forte, se robusto.	Dependência da cobertura da base.
Um *cluster* novo surgiu no último período.	Rede temporal e termos característicos.	Moderado.	Sensibilidade a parâmetros de *cluster*ing.
Há mudança terminológica.	Substituição de termos em títulos e resumos.	Forte quando confirmada por leitura.	Possível polissemia.
O tema pode consolidar-se futuramente.	Crescimento, diversificação e citações recentes.	Preliminar.	Cenário condicional, não previsão.

Fonte: Elaborado pelo autor (2026).

O sexto componente é a matriz de conclusões. Para cada conclusão temporal, o pesquisador deve indicar indicador de suporte, período, unidade de análise, nível de confiança e limitações. Essa matriz é útil para evitar que a discussão final inclua afirmações não sustentadas pelos resultados. A IA pode ajudar a preencher uma versão preliminar da matriz, mas o pesquisador deve revisá-la. O produto final deve permitir ao leitor distinguir claramente entre dados observados, interpretações plausíveis e hipóteses prospectivas.

13.14 Técnicas quantitativas e híbridas para análise temporal

A análise temporal em cientometria pode empregar técnicas descritivas, inferenciais e relacionais. A técnica mais simples é a contagem anual de documentos, mas ela raramente é suficiente para explicar mudança. Quando o objetivo é detectar fases, a regressão segmentada pode indicar pontos de inflexão, isto é, períodos nos quais a inclinação da curva se altera. Essa abordagem é útil para distinguir crescimento gradual de aceleração abrupta, mas exige cuidado com séries curtas e com valores extremos. Em campos com poucos documentos, um único dossiê temático ou congresso especial pode criar uma aparente ruptura. Assim, a escolha do método deve considerar volume, dispersão e estabilidade do *corpus*.

Outra técnica amplamente utilizada é a detecção de *burst*. A ideia central é identificar termos, referências ou autores que apresentam aumento subitamente elevado de frequência ou citação durante um intervalo. Em estudos de ciência, *bursts* podem sinalizar frentes de pesquisa, descobertas relevantes, debates emergentes ou reativação de literatura clássica. Contudo, a interpretação de *burst* exige contexto. Um *burst* curto pode indicar atenção episódica; um *burst* prolongado pode indicar consolidação; um *burst* tardio em referência antiga pode sugerir redescoberta ou uso retrospectivo. A IA pode ajudar a descrever esses padrões, mas deve ser instruída a não tratar todo *burst* como inovação.

Modelos de tópicos dinâmicos oferecem outra possibilidade, especialmente quando o *corpus* contém títulos e resumos em larga escala. Diferentemente de modelos estáticos, modelos dinâmicos procuram acompanhar como a distribuição de palavras associadas a um tópico se altera ao longo do tempo. Essa abordagem pode revelar transformações semânticas que não aparecem em contagens de palavras-chave. Entretanto, tópicos modelados estatisticamente não são categorias naturais; eles dependem de parametrização, pré-processamento, número de tópicos, qualidade dos textos e interpretação humana. A IA generativa pode auxiliar na

rotulagem preliminar, mas a interpretação deve ser validada por leitura de documentos representativos (BLEI; LAFFERTY, 2006; WANG; BLEI; HECKERMAN, 2012).

Janelas móveis constituem recurso frequente em mapeamento longitudinal. Em vez de comparar anos isolados, o pesquisador agrega períodos sobrepostos, como 2010-2014, 2011-2015 e 2012-2016. Essa estratégia reduz ruído e permite observar continuidade, mas também cria dependência entre janelas. Mudanças podem parecer suaves porque parte dos documentos se repete entre períodos. Janelas não sobrepostas, por sua vez, oferecem comparações mais independentes, mas podem perder detalhe. A decisão deve ser informada pelo tamanho do *corpus* e pela pergunta de pesquisa. A IA pode sugerir vantagens e limites, mas não deve escolher janelas sem conhecer o volume anual de documentos.

A análise de envelhecimento de citações também contribui para compreender tendências. Um tema pode ter alta produção recente, mas baixa citação por ainda não ter tido tempo de acumular impacto. Outro pode ter poucas publicações recentes, mas citações persistentes, indicando papel fundacional. Janelas de citação e normalização por idade do documento são essenciais para evitar comparações injustas. A IA, ao interpretar impacto temporal, deve ser instruída a considerar atraso de citação e diferenças disciplinares. Em campos rápidos, citações podem surgir em poucos anos; em humanidades e algumas ciências sociais, a circulação é mais lenta.

A análise relacional temporal observa como redes mudam. Redes de cocitação podem revelar transformações na estrutura intelectual; redes de coautoria mostram reorganização social; redes de coocorrência de termos indicam mudanças conceituais. O desafio e comparar redes de tamanhos diferentes. Mudanças em centralidade podem decorrer de crescimento do *corpus* e não de mudança substantiva. Técnicas de normalização e limiares estáveis ajudam, mas sempre envolvem escolhas. A IA pode descrever diferenças entre mapas, desde que receba parâmetros e seja orientada a mencionar que layout visual não é evidência suficiente de mudança (ROSVALL; BERGSTROM, 2010; WALTMAN; VAN ECK; NOYONS, 2010).

Métodos híbridos combinam estatística, rede e texto. Um protocolo pode começar por crescimento anual, seguir para termos emergentes, construir redes por período, identificar *clusters*, rotular temas com apoio de IA e validar com leitura de amostras. Essa combinação oferece visão mais robusta do que qualquer técnica isolada. Entretanto, quanto mais complexo o protocolo, maior a necessidade de documentação. Cada transformação cria oportunidades de viés: filtros, lematização, tradução, exclusão de termos, limiar de ocorrência e algoritmo de comunidade. A IA deve ser usada para explicitar essas escolhas, não para ocultar sua complexidade.

Quadro 13.10 - Técnicas para análise temporal e seus usos

Técnica	Pergunta que responde	Limite principal
Regressão segmentada	Há pontos de inflexão no crescimento?	Requer série suficiente e controle de *outliers*.
Burst detection	Quais termos ou referências tiveram aumento abrupto?	Pode confundir atenção episódica com consolidação.
Modelos de tópicos dinâmicos	Como temas latentes mudam semanticamente?	Depende de parâmetros e interpretação humana.
Janelas móveis	Como suavizar variações anuais?	Pode ocultar rupturas curtas.
Redes temporais	Como comunidades e relações mudam?	Sensível a normalização e limiares.

Fonte: Elaborado pelo autor (2026).

Em síntese, técnicas quantitativas e híbridas devem ser escolhidas a partir da pergunta, e não da disponibilidade de software. A pergunta "o tema cresce?" exige indicadores diferentes da pergunta "o significado do tema mudou?" ou "a comunidade se reorganizou?". *Prompts* cientométricos podem ajudar a traduzir pergunta em método, mas a validade depende da coerência entre unidade de análise, fonte, período e indicador. O pesquisador deve resistir à tentação de aplicar todas as técnicas possíveis. Em análises temporais, parcimônia metodológica frequentemente é sinal de rigor.

13.15 Visualização longitudinal e comunicação de tendências

A visualização longitudinal desempenha papel central na interpretação de tendências cientométricas. Gráficos de linhas, áreas empilhadas, mapas sobrepostos, diagramas aluviais, mapas de calor, linhas do tempo e redes por período tornam visíveis padrões que seriam difíceis de perceber em tabelas. Todavia, visualizações também podem induzir interpretações indevidas. Escalas truncadas, cores sugestivas, agregações arbitrárias e layouts instáveis podem exagerar mudanças. Por isso, a visualização deve ser tratada como argumento metodológico, não como ornamento. A IA pode sugerir tipos de gráfico, mas o pesquisador deve avaliar se o gráfico responde à pergunta é preserva proporcionalidade (CHEN, 2004; VAN ECK; WALTMAN, 2014).

Gráficos de linha são apropriados para séries anuais de produção, citação ou ocorrência de termos. Eles facilitam a identificação de tendência, picos e quedas. Entretanto, quando há muitos temas, tornam-se ilegíveis. Mapas de calor podem ser mais adequados para comparar termos ao longo de períodos, mostrando intensidades por cor. Áreas empilhadas ajudam a visualizar participação relativa de categorias, mas podem dificultar a leitura de categorias pequenas. Diagramas aluviais representam fluxos entre *clusters* ou categorias ao longo do tempo, mostrando fusões, divisões e continuidades. Cada forma visual enfatiza um aspecto e oculta outro.

Em mapeamento científico, visualizações de redes por período devem ser comparadas com cautela. Layouts gerados independentemente podem colocar *clusters* em posições diferentes sem que isso represente mudança substantiva. Uma rede aparentemente mais dispersa pode resultar de parâmetro de layout, não de fragmentação real. Para comparações, é recomendável manter critérios de inclusão, normalização e escala sempre que possível, ou explicar diferenças. A IA, ao interpretar mapas, deve receber dados estruturais além da imagem ou da descrição visual: número de nós, arestas, densidade, modularidade, principais *clusters* e termos característicos.

Linhas do tempo são particularmente úteis para combinar eventos bibliométricos e eventos contextuais. Um estudo pode marcar publicações seminais, lançamento de tecnologias, mudanças regulatórias, criação de programas de financiamento, crises sanitárias ou marcos institucionais. Essa combinação ajuda a formular hipóteses explicativas, mas não prova causalidade. Se uma curva cresce após determinado evento, pode haver relação, mas a evidência temporal é apenas um indício. A IA pode ajudar a montar uma linha do tempo narrativa, desde que o *prompt* exija separar coincidência temporal de explicação causal.

A comunicação de tendências para públicos não especialistas exige simplificação sem distorção. Gestores podem desejar saber "qual tema está crescendo", mas o analista precisa explicar se o crescimento e absoluto, relativo, robusto e contextualizado. A IA pode adaptar linguagem, criar sumários executivos e propor títulos de gráficos, mas deve evitar sensacionalismo. Expressões como "explosão", "revolução" e "dominância" devem ser usadas somente quando sustentadas por evidência forte e definidas tecnicamente. Em geral, termos

como expansão, intensificação, aumento relativo e consolidação inicial são mais adequados (HICKS *et al.*, 2015).

Legendas e notas metodológicas são parte essencial da visualização. Um gráfico temporal deve informar fonte dos dados, data de coleta, tipo documental, período, tratamento do ano corrente e critério de contagem. Quando uma visualização resulta de agrupamento automatizado ou rotulagem por IA, isso também deve ser indicado. A IA pode gerar legendas analíticas, mas o *prompt* deve exigir inclusão das limitações. Uma legenda ideal não apenas descreve o que se vê, ela orienta o leitor sobre como não interpretar indevidamente o gráfico.

Visualizações também podem apoiar validação. Ao apresentar tendências para especialistas, mapas e linhas do tempo facilitam a identificação de erros: termos mal agrupados, períodos estranhos, ausências de eventos conhecidos, *clusters* mal nomeados ou países sub-representados. A IA pode gerar perguntas para guiar essa validação, como "este termo deveria estar agrupado com aquele?", "a queda após determinado ano pode ser efeito de indexação?" ou "há evento externo que explique o pico?". Assim, a visualização se torna instrumento de diálogo entre dados, IA e conhecimento especializado.

Em suma, visualizar tendências e construir uma narrativa visual sobre a ciência. Essa narrativa deve ser precisa, proporcional e transparente. A IA pode apoiar escolhas gráficas e redação, mas a ética da visualização permanece humana. Um bom gráfico cientométrico temporal deve permitir ao leitor perceber mudanças, mas também compreender limites. Deve convidar a interpretação, não impor conclusões. Deve mostrar o tempo como processo, não como destino.

Quadro 13.11 - Visualizações temporais recomendadas

Visualização	Uso preferencial	Cautela
Linha temporal	Séries anuais de indicadores.	Evitar escala truncada.
Mapa de calor	Comparar muitos termos por período.	Explicar normalização de cores.
Diagrama aluvial	Mostrar fluxos entre *clusters*.	Não inferir causalidade pelo fluxo visual.
Mapa sobreposto	Comparar posições temáticas.	Controlar layout e escala.
Linha do tempo anotada	Relacionar dados e eventos.	Distinguir coincidência de explicação.

Fonte: Elaborado pelo autor (2026).

13.16 Qualidade dos dados, bases bibliográficas e vieses temporais

Nenhuma análise temporal é melhor que seus dados. Em cientometria, a qualidade temporal depende da consistência de metadados ao longo dos anos. Bases bibliográficas mudam cobertura, políticas de indexação, campos disponíveis, critérios de classificação e procedimentos de atualização. Uma área pode parecer crescer porque a base passou a indexar mais periódicos de determinado país ou idioma. Um tópico pode parecer recente porque documentos antigos não tinham resumos ou palavras-chave disponíveis. Uma queda pode resultar de atraso de indexação. Esses problemas não invalidam a análise, mas precisam ser reconhecidos.

A comparação entre bases é particularmente importante. Web of Science, Scopus, PubMed, Dimensions, Crossref, Lens, OpenAlex e outras fontes possuem perfis distintos. Algumas privilegiam periódicos; outras incluem *preprints*, livros, anais ou dados mais

amplos. Algumas possuem melhor cobertura histórica em determinadas áreas; outras são mais abertas, mas podem ter metadados heterogêneos. Em análises temporais, a escolha da base pode alterar o ano de início aparente de um tema, a intensidade de crescimento e a distribuição geográfica. A IA pode ajudar a redigir uma seção de limitações, mas não deve sugerir equivalência entre bases diferentes.

Os campos bibliográficos também variam no tempo. Palavras-chave podem estar ausentes em registros antigos; resumos podem ser incompletos; afiliações podem ser registradas de forma irregular; referências citadas podem não estar disponíveis para todos os documentos. Uma análise baseada em resumos pode favorecer períodos recentes, não porque os temas eram inexistentes antes, mas porque os resumos antigos não foram indexados. Um estudo de colaboração internacional pode ser prejudicado por afiliações incompletas em anos remotos. O pesquisador deve avaliar a completude de cada campo antes de comparar períodos.

O tratamento de *preprints* introduz desafios adicionais. Em algumas áreas, *preprints* antecipam tendências antes da publicação em periódicos. Em outras, eles são menos relevantes. Incluir *preprints* pode melhorar sensibilidade para temas emergentes, mas também pode gerar duplicatas quando o mesmo trabalho aparece depois como artigo. Excluir *preprints* pode atrasar a detecção de frentes recentes. A decisão deve ser alinhada ao objetivo: monitoramento antecipado pode incluir *preprints*; avaliação consolidada pode priorizar literatura revisada por pares. A IA pode ajudar a listar vantagens e riscos, mas a regra deve ser explicitada.

O idioma e outro viés temporal e geográfico. Temas podem emergir em literaturas nacionais antes de se tornarem visíveis em inglês. Bases internacionais podem sub-representar produção local, especialmente em ciências sociais, humanidades e áreas aplicadas. Análises que consideram apenas literatura em inglês podem produzir narrativas de emergência tardia ou centralidade de países anglófonos. A IA, ao interpretar trajetórias, deve ser instruída a reconhecer limites linguísticos. Quando o estudo tiver interesse global ou latino-americano, fontes regionais e repositórios institucionais podem ser necessários.

Duplicatas e variantes de metadados afetam séries temporais. O mesmo documento pode aparecer em mais de uma base, com anos diferentes ou tipos documentais distintos. Uma versão online *first* pode ter ano diferente da versão final. Um registro de conferência pode preceder o artigo. Sem regras claras, a contagem anual pode ser distorcida. O pré-processamento deve preservar o ano utilizado e justificar critérios de deduplicação. A IA pode auxiliar na identificação de casos suspeitos, mas decisões devem ser tomadas por regras documentadas.

A indexação retrospectiva é um problema sutil. Algumas bases adicionam documentos antigos em anos recentes. Se o pesquisador coleta em 2026, pode encontrar registros de 1980 que não estavam disponíveis em coletas anteriores. Isso é positivo para cobertura, mas complica comparações históricas entre estudos. Além disso, a data de publicação é diferente da data de entrada na base. Para crescimento da literatura, normalmente usa-se ano de publicação; para monitoramento de disponibilidade informacional, poderia interessar ano de indexação. A pergunta define qual temporalidade importa.

Por fim, a qualidade dos dados deve ser reportada. Um capítulo ou artigo cientométrico temporal deve incluir nota sobre bases, cobertura, campos, exclusões, ano corrente, duplicatas, normalização e limitações. A IA pode transformar esse conjunto em texto metodológico claro. Contudo, se o pesquisador não fornecer os elementos, o modelo pode produzir uma seção genérica e insuficiente. Em análise temporal, a transparência sobre dados não é detalhe técnico; é fundamento da interpretação.

13.17 Validação humana, amostragem e triangulação interpretativa

A validação humana é indispensável em análises temporais assistidas por IA. O motivo é simples: tendências cientométricas frequentemente envolvem significados disciplinares que não podem ser inferidos apenas por padrões numéricos. Um termo em crescimento pode representar técnica, teoria, aplicação, objeto empírico, moda discursiva ou exigência de financiamento. Um *cluster* pode reunir documentos próximos por vocabulário, mas distantes teoricamente. Um aumento de citações pode indicar influência, controvérsia, crítica ou método amplamente usado. A IA pode sugerir leituras, mas especialistas precisam avaliar pertinência.

A amostragem de documentos é uma prática eficiente de validação. Para cada tendência detectada, recomenda-se selecionar documentos de diferentes anos, tipos documentais, *clusters* e níveis de citação. A leitura deve verificar se os documentos realmente tratam do tema esperado, se o vocabulário mudou, se há documentos irrelevantes e se as categorias fazem sentido. Em amostras grandes, a IA pode resumir documentos e criar matrizes de codificação preliminar, mas a amostra deve incluir revisão humana. O objetivo não é ler todo o *corpus*, mas testar a validade semântica dos sinais.

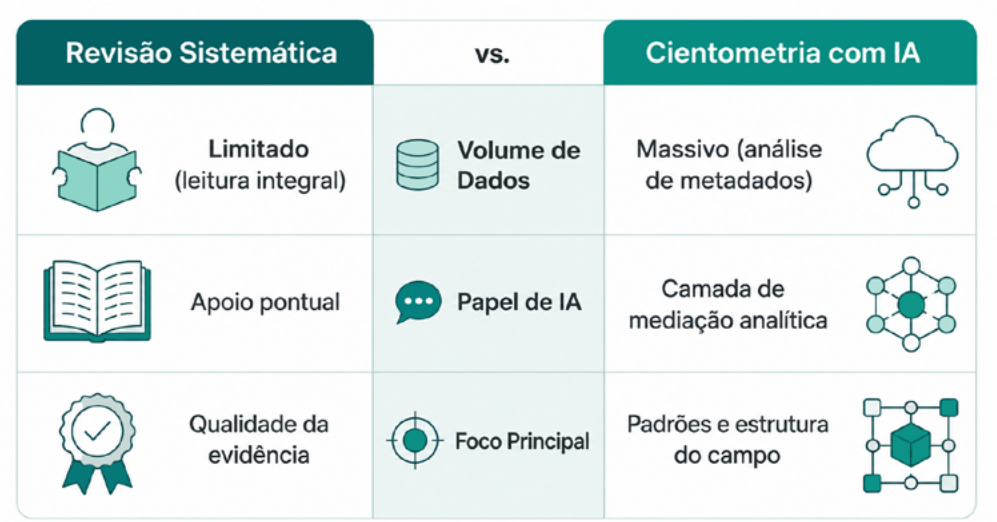
A triangulação interpretativa combina métodos, fontes e atores. Métodos diferentes podem confirmar ou questionar uma tendência: série de publicações, coocorrência de termos, rede de citações, leitura de resumos e indicadores de colaboração. Fontes diferentes podem revelar cobertura desigual: uma tendência em Scopus pode não aparecer em PubMed; um tema em OpenAlex pode refletir inclusão maior de *preprints*. Atores diferentes, como especialistas de área, bibliotecários e analistas de dados, podem identificar problemas distintos. A IA deve ser apenas um participante instrumental nesse processo.

A revisão por especialistas é particularmente importante na rotulagem de temas. Modelos de linguagem tendem a produzir rótulos plausíveis e elegantes, mas nem sempre disciplinarmente corretos. Um *cluster* com termos como *"network", "model", "learning"* e *"science"* pode receber rótulo genérico que oculta subcampo específico. Especialistas podem reconhecer que o *cluster* corresponde a uma tradição teórica, método ou debate particular. A validação deve registrar alterações feitas nos rótulos sugeridos pela IA. Esse registro aumenta transparência e evita falsa autoridade automática.

A validação também deve examinar conclusões negativas. Quando um tema parece declinar, é necessário verificar se ele não migrou para outro termo ou para outra fonte. Quando uma instituição parece desaparecer, pode ter mudado de nome. Quando um país parece pouco presente, pode publicar em bases regionais fora do *corpus*. A IA pode listar explicações alternativas, mas não conhece necessariamente a história institucional local. Por isso, a ausência em dados bibliográficos deve ser interpretada como ausência no *corpus*, não como ausência na ciência.

A concordância entre avaliadores pode ser útil quando a IA participa de classificações. Se o modelo agrupa termos ou classifica documentos por período, uma amostra pode ser revisada por dois avaliadores humanos. Divergências indicam categorias ambíguas, *prompts* insuficientes ou problema conceitual. O resultado pode levar a ajustes de *prompt*, revisão de categorias ou exclusão de casos duvidosos. Essa prática aproxima o uso de IA de protocolos de análise de conteúdo e revisão sistemática, nos quais confiabilidade e decisão humana são explicitadas.

Figura 13 - Comparação: Revisão sistemática vs Cientometria com IA



Fonte: Imagem gerada por IA, ChatGPT (2026).

A triangulação com eventos contextuais deve ser feita com cuidado. Ao observar crescimento de publicações após uma política pública ou lançamento tecnológico, o pesquisador pode formular hipótese de relação, mas não deve afirmar causalidade sem evidências adicionais. Entrevistas, documentos de política, dados de financiamento e análise institucional podem complementar a cientometria. A IA pode ajudar a organizar essas fontes, mas a inferência causal requer desenho específico. Em muitos casos, a formulação mais adequada e "o crescimento coincide com" ou "e compatível com", não "foi causado por".

Em resumo, a validação humana não é obstáculo ao uso de IA; é condição para que ela seja cientificamente útil. A IA acelera a exploração, mas a comunidade de pesquisa fornece critério. A amostragem permite testar significado; a triangulação permite avaliar robustez; a revisão especializada permite corrigir rótulos e narrativas. Análises temporais confiáveis nascem da combinação entre computação, interpretação e responsabilidade.

13.18 Redação de resultados temporais com apoio de prompts

A redação de resultados temporais é uma etapa crítica. Mesmo quando a análise quantitativa foi correta, o texto pode exagerar conclusões. Frases como "o tema explodiu", "a área foi revolucionada", "o país tornou-se líder" ou "a tecnologia determinou o crescimento" devem ser substituídas por formulações proporcionais aos dados. A IA pode ajudar tanto a produzir a primeira versão quanto a revisar excessos. O *prompt* deve solicitar linguagem formal, indicação de período, unidade de análise, indicador e limite. Uma frase bem escrita informa o que mudou, onde, quando, segundo qual dado e com qual cautela.

Uma estrutura útil para resultados temporais contém quatro movimentos. Primeiro, descrição: "o número de documentos aumentou de X para Y entre A e B". Segundo, contextualização: "esse aumento corresponde a crescimento relativo de Z em comparação ao *corpus* geral". Terceiro, interpretação prudente: "o padrão sugere intensificação da atenção ao tema". Quarto, limite: "entretanto, a interpretação deve considerar a cobertura da base e o ano recente incompleto". Essa estrutura reduz o risco de transformar gráfico em narrativa causal simplista. A IA pode ser instruída a seguir esses quatro movimentos.

Na discussão, o pesquisador pode relacionar tendências a literatura existente. Se estudos anteriores identificaram crescimento semelhante, isso fortalece a interpretação. Se os resultados divergem, é preciso comparar *corpus*, período, base e estratégia de busca. A IA pode ajudar a formular parágrafos comparativos, mas referências devem ser verificadas. Modelos generativos podem produzir citações inexistentes ou associar autores a afirmações incorretas. Portanto, a redação assistida por IA deve usar referências previamente selecionadas e conferidas pelo pesquisador.

A redação de cenários prospectivos requer ainda mais cautela. Em vez de "o tema continuará crescendo", recomenda-se "caso as condições observadas se mantenham, o tema poderá apresentar continuidade de crescimento". Em vez de "o *cluster* se consolidará", pode-se escrever "a persistência do *cluster* em períodos subsequentes será indício de consolidação". Em vez de "a instituição lidera a área", pode-se escrever "a instituição ocupa posição central no *corpus* analisado segundo os indicadores considerados". A IA pode reescrever conclusões usando esse tipo de modalização.

Textos temporais devem evitar ambiguidade entre tempo de publicação, tempo de citação e tempo de indexação. Um artigo publicado em 2022 pode receber citações em 2023 e ser indexado em 2024. Cada temporalidade responde à pergunta diferente. Ao escrever resultados, deve-se informar se o ano se refere a publicação, citação, coleta, indexação ou evento. A IA pode revisar o texto em busca dessa ambiguidade. Esse cuidado é particularmente relevante em estudos que combinam bases, *preprints* e documentos online *first*.

A escrita também deve lidar com incerteza de modo transparente. Incerteza não enfraquece a pesquisa; ela mostra maturidade metodológica. Frases como "os dados sugerem", "o padrão é compatível com", "a hipótese requer validação adicional" e "o resultado deve ser interpretado no limite do *corpus*" ajudam a calibrar conclusões. A IA pode gerar uma escala verbal de confiança para as principais afirmações, desde que o pesquisador defina critérios. Essa escala pode ser apresentada em quadro ou usada internamente na revisão.

Outro aspecto importante é a comunicação de resultados negativos ou nulos. Se não há crescimento, se uma tendência esperada não aparece ou se um tema é instável, isso também é achado. A pressão por novidade pode levar relatórios a ignorar estabilidade, mas estabilidade pode indicar consolidação, saturação ou continuidade de uma agenda. A IA pode ajudar a escrever resultados nulos de forma informativa, evitando interpretações forçadas. Nem toda análise temporal precisa descobrir emergência; muitas vezes, a contribuição está em mostrar permanência ou ausência de evidência.

Por fim, a redação assistida por *prompts* deve ser revisada como qualquer texto científico. O pesquisador deve verificar se termos técnicos foram usados corretamente, se a voz do texto é coerente, se as referências existem, se os limites metodológicos estão presentes e se a argumentação acompanha os resultados. A IA pode aumentar velocidade, mas não reduz a responsabilidade autoral. Em cientometria com inteligência artificial, escrever bem e também escrever de modo rastreável, verificável e proporcional aos dados.

13.19 Governança, ética e responsabilidade na detecção de tendências

A detecção de tendências científicas não é uma atividade neutra. Relatórios sobre temas emergentes podem orientar editais, prioridades institucionais, contratações, distribuição de bolsas, abertura de laboratórios e estratégias de internacionalização. Quando um tema e classificado como promissor, ele pode atrair recursos; quando e classificado como declinante, pode perder legitimidade. Por isso, a análise temporal assistida por IA deve ser acompanhada de governança ética. O pesquisador deve explicitar critérios, limites e incertezas, evitando que

resultados sejam usados como justificativa automática para decisões que exigem deliberação política e científica (HICKS *et al.*, 2015; FORTUNATO *et al.*, 2018).

Um risco recorrente é o reforço de desigualdades. Tendências detectadas em bases internacionais tendem a privilegiar literaturas em inglês, periódicos de maior visibilidade e países com sistemas científicos mais bem indexados. Se esses resultados forem usados para definir prioridades, áreas locais, temas sociais relevantes ou produções em idiomas nacionais podem ser subvalorizados. A IA pode agravar o problema ao produzir narrativas universalizantes com base em dados parciais. Assim, *prompts* devem exigir que a resposta especifique o escopo da evidência: tendência no *corpus* analisado, na base consultada, no período definido e sob os filtros aplicados.

A governança também envolve transparência no uso de IA. Quando modelos generativos ajudam a nomear *clusters*, resumir tendências ou formular cenários, o relatório deve informar essa participação. Não é necessário transformar cada *prompt* em anexo volumoso, mas os *prompts* principais, critérios de validação e decisões humanas devem ser preservados. Essa documentação protege a integridade do estudo e permite que leitores avaliem o grau de dependência da interpretação em relação ao sistema utilizado. Em contextos institucionais, esse registro pode ser decisivo para auditar recomendações estratégicas.

Outro ponto ético diz respeito à avaliação de pessoas. Séries temporais de autores podem ser interpretadas como trajetórias de produtividade ou impacto, mas nunca devem ser usadas isoladamente para julgar mérito individual. Interrupções de carreira, desigualdades de gênero, responsabilidades familiares, acesso desigual a financiamento, diferenças disciplinares e contribuições não publicadas afetam a produção científica. A IA não deve produzir rankings normativos de pesquisadores sem contexto. Quando a unidade de análise incluir autores, o texto deve enfatizar que os indicadores descrevem registros bibliográficos, não valor humano ou qualidade integral da contribuição.

A responsabilidade também inclui prevenir usos performativos das tendências. Se uma universidade passa a financiar apenas temas detectados como emergentes, pode estimular comportamento adaptativo: pesquisadores rebatizam projetos para aderir a termos em moda, periódicos criam dossiês oportunistas e grupos abandonam problemas importantes de longo prazo. A cientometria temporal deve apoiar entendimento da ciência, não reduzir a ciência a corrida por sinais de crescimento. Cenários prospectivos responsáveis devem apresentar oportunidades e riscos, indicando que decisões estratégicas devem combinar dados, avaliação qualitativa, missão institucional e compromisso social.

Finalmente, a detecção de tendências deve preservar pluralidade epistêmica. A ciência avança tanto por áreas quentes quanto por trabalhos lentos, acumulativos, teóricos, locais ou de manutenção de infraestrutura intelectual. Nem toda contribuição relevante aparece como *burst*, crescimento acelerado ou *cluster* novo. A IA, se orientada apenas por novidade, pode invisibilizar continuidade e profundidade. Um capítulo sobre análise temporal deve, portanto, concluir com uma advertência: tendências são sinais a serem interpretados, não ordens a serem seguidas. O papel da cientometria com IA é ampliar a compreensão crítica da dinâmica científica, e não automatizar a escolha do futuro.

13.20 Considerações finais

A análise temporal e a detecção de tendências ocupam lugar estratégico na cientometria contemporânea. Elas permitem compreender não apenas quanto se publica, mas como agendas se formam, como conceitos se transformam, como comunidades se reorganizam e como instituições e países se posicionam na paisagem científica. Quando associadas a

inteligência artificial e a *prompts* bem estruturados, essas análises ganham capacidade de síntese e interpretação. Contudo, também ganham novos riscos: narrativas excessivamente fluentes, previsões sem base, causalidades indevidas e ocultamento de incertezas.

O capítulo defendeu que a IA deve ser utilizada como instrumento metodológico auxiliar. Ela pode interpretar séries, comparar períodos, detectar possíveis mudanças temáticas, auditar artefatos de busca e formular hipóteses prospectivas. Mas não deve substituir cálculos, validação humana, leitura especializada, controle de dados e prudência epistêmica. Em estudos cientométricos temporais, a qualidade da interpretação depende menos da eloquência da narrativa é mais da rastreabilidade do processo.

Uma tendência cientométrica é sempre situada: depende de *corpus*, base, período, estratégia de busca, normalização, campo, idioma e tipo documental. Por isso, conclusões devem ser apresentadas com limites explícitos. O uso responsável de *prompts* contribui para essa tarefa quando exige que a IA separe dado, inferência, hipótese e especulação. Ao final, analisar o tempo na ciência e analisar mudança sob incerteza. A inteligência artificial pode ajudar a organizar essa incerteza, mas não deve apagá-la.

A análise temporal amplia a cientometria porque desloca a pergunta "o que existe no *corpus*?" para "como o *corpus* se transformou?". Essa mudança é decisiva. Ao incorporar o tempo, o pesquisador consegue identificar fases, transições, rupturas, permanências e deslocamentos. A IA generativa, por sua vez, amplia a capacidade de formular perguntas, sintetizar resultados e revisar interpretações, desde que submetida a protocolos que impeçam alucinação, causalidade indevida e excesso de confiança.

O capítulo mostrou que crescimento da literatura, ciclo de vida temático, emergência, mudança terminológica, trajetórias institucionais e cenários prospectivos são dimensões interligadas. Uma tendência robusta raramente aparece em apenas um indicador. Ela costuma envolver combinação de sinais: aumento relativo, articulação em rede, estabilidade terminológica ou mudança explicável, entrada de atores, citações recentes e coerência substantiva. A tarefa do pesquisador é avaliar essa convergência, e não apenas selecionar o gráfico mais convincente (BORNMANN; MUTZ, 2015; CHEN, 2006; COBO *et al.*, 2011).

A principal contribuição dos *prompts*, nesse contexto, é tornar a interpretação mais disciplinada. Um *prompt* pode exigir que a IA separe descrição, explicação, hipótese e incerteza. Pode pedir que ela identifique artefatos e reescreva conclusões em linguagem prudente. Pode transformar resultados temporais em matriz de monitoramento. Essas funções tornam a IA uma tecnologia de apoio metodológico, e não uma autoridade sobre o futuro da ciência. Em uma cientometria responsável, a inteligência artificial não substitui o juízo humano: ela deve tornar esse juízo mais explícito, documentado e criticável (WEI *et al.*, 2022; WHITE *et al.*, 2023).

Capítulo 14

IA para leitura, sumarização e classificação de literatura

14.1 Introdução

A expansão contínua da literatura científica produziu uma condição paradoxal para a pesquisa contemporânea. Nunca houve tantos registros bibliográficos, resumos indexados, artigos a*bert*os, bases de dados interoperáveis e ferramentas de busca; ao mesmo tempo, nunca foi tão difícil formar uma visão substantiva, crítica e atualizada de um campo de conhecimento em tempo hábil. Em áreas dinâmicas, o volume de publicações excede a capacidade de leitura integral de indivíduos e equipes pequenas. Em áreas interdisciplinares, a dispersão terminológica e institucional dificulta o reconhecimento de convergências. Em áreas emergentes, a velocidade de publicação cria incerteza sobre o que constitui moda passageira, tendência consolidada, frente de pesquisa ou lacuna real. Nesse cenário, a inteligência artificial aparece como instrumento promissor para ampliar a capacidade de leitura e organização da literatura, mas também como fonte de novos riscos metodológicos.

A cientometria sempre dependeu de operações de redução: contar publicações, agrupar autores, calcular citações, extrair palavras-chave, construir redes, nomear *clusters* e interpretar padrões. Tais operações não eliminam a complexidade da ciência; elas a tornam manejável por meio de representações. A IA generativa e os modelos de linguagem introduzem uma nova camada nessa tradição: podem condensar resumos, propor categorias, identificar métodos, distinguir populações estudadas, sugerir lacunas, comparar descritores e redigir narrativas. Todavia, essa camada é discursiva, probabilística e sensível ao *prompt*. Por isso, seu uso exige critérios de governança metodológica. A pergunta fundamental não é se a IA pode ler textos científicos, mas sob quais condições suas leituras podem ser úteis, controláveis, auditáveis e epistemicamente responsáveis (SMALL, 1973; WHITE; GRIFFITH, 1981; VAN ECK; WALTMAN, 2010; ARIA; CUCCURULLO, 2017).

Neste capítulo, a leitura assistida por IA é definida como o conjunto de procedimentos nos quais modelos computacionais são utilizados para examinar, resumir, classificar ou extrair informações de documentos científicos, com o objetivo de apoiar decisões analíticas em estudos cientométricos. A expressão "assistida" é central: indica que a IA participa do processo como instrumento, e não como autoridade final. A máquina pode sugerir agrupamentos temáticos; o pesquisador avalia sua validade. A máquina pode resumir mil resumos; o pesquisador verifica se a síntese preserva diversidade e incerteza. A máquina pode identificar lacunas; o pesquisador decide se elas são lacunas empíricas, lacunas de indexação, lacunas de busca ou lacunas produzidas por um recorte inadequado.

A relevância do tema decorre da convergência entre três tradições. A primeira é a tradição da recuperação da informação e da revisão de literatura, que busca localizar, selecionar e sintetizar estudos de modo transparente. A segunda é a tradição cientométrica, que analisa estruturas, dinâmicas e indicadores da produção científica. A terceira é a tradição do processamento de linguagem natural, que desenvolve métodos para representação, classificação, sumarização e extração de informação em textos. A IA contemporânea, especialmente os modelos de linguagem baseados em arquiteturas *transformer*, não substitui nenhuma dessas tradições; ela as tensiona, amplia e exige que sejam reinterpretadas à luz de novas capacidades de geração textual (VASWANI *et al.*, 2017; DEVLIN *et al.*, 2019; BROWN *et al.*, 2020).

O uso de IA para leitura de literatura é particularmente relevante em estudos baseados em títulos, resumos e palavras-chave. Esses campos textuais não equivalem ao artigo completo, mas constituem uma fonte concentrada de informação sobre objeto, método, teoria, população, evidência e vocabulário disciplinar. Em análises cientométricas, eles são frequentemente empregados para mapas conceituais, coocorrência de termos, detecção de temas emergentes, rotulagem de *clusters* e caracterização de frentes de pesquisa. Com IA, torna-se possível complementar esses procedimentos com classificações semânticas, justificativas textuais e identificação de casos ambíguos. Porém, quanto maior a sofisticação linguística da saída, maior deve ser a exigência de validação.

Este capítulo defende uma tese metodológica: a IA deve ser incorporada à leitura, à sumarização e à classificação da literatura como componente de um protocolo reprodutível. Isso implica registrar dados de entrada, versão dos modelos quando possível, parâmetros disponíveis, *prompts*, instruções de classificação, exemplos usados, critérios de exclusão, respostas obtidas, alterações humanas e medidas de concordância. Sem esse registro, a IA pode produzir resultados aparentemente úteis, mas cientificamente frágeis. Com esse registro, os modelos de linguagem podem se tornar instrumentos de triagem, organização e interpretação preliminar, aumentando a escala da análise sem abandonar o rigor.

A discussão a seguir está organizada em seis partes. A primeira trata da sumarização de grandes conjuntos de resumos. A segunda discute a classificação temática assistida. A terceira aborda a extração de métodos, objetos e populações estudadas. A quarta analisa a identificação de lacunas de pesquisa. A quinta compara revisões sistemáticas e estudos cientométricos. A sexta apresenta critérios de controle de qualidade em classificações feitas por IA. Ao final, são apresentados *prompts* aplicáveis a diferentes etapas do processo, bem como orientações para documentação e validação.

14.2 Sumarização de grandes conjuntos de resumos

A sumarização de grandes conjuntos de resumos é uma das aplicações mais imediatas da inteligência artificial em estudos cientométricos. Em vez de ler isoladamente centenas ou milhares de registros, o pesquisador pode solicitar ao modelo que produza descrições agregadas sobre temas, métodos, populações, países, períodos, tipos de evidência e vocabulário dominante. Essa operação não elimina a necessidade de leitura humana, mas pode reorganizar a entrada cognitiva do pesquisador. Em um estudo exploratório, a IA pode ajudar a identificar eixos temáticos preliminares; em uma etapa intermediária, pode verificar se determinados subtemas aparecem em segmentos específicos do *corpus*; em uma etapa final, pode apoiar a redação de um panorama descritivo, desde que fundamentado em dados verificáveis.

A primeira distinção necessária é entre resumo de documento e resumo de *corpus*. O resumo de documento condensa o conteúdo de um artigo ou registro específico. O resumo de *corpus* procura sintetizar padrões recorrentes em um conjunto de registros. Em cientometria, o segundo tipo é mais importante, pois o objetivo não é substituir a leitura de um artigo, mas compreender a estrutura coletiva da literatura. Quando se pede a um modelo que "resuma os resumos", é preciso indicar se a saída deve apresentar frequências aproximadas, categorias temáticas, exemplos de termos, mudanças temporais, lacunas, contradições ou apenas uma visão narrativa. A ausência dessa especificação favorece respostas genéricas, pouco úteis e muitas vezes excessivamente seguras.

Historicamente, a sumarização automática possui raízes anteriores à IA generativa. O trabalho de Luhn, publicado em periódico científico no final da década de 1950, já explorava métodos estatísticos para identificar palavras significativas em textos técnicos. Posteriormente,

técnicas baseadas em grafos, como *TextRank*, propuseram ordenar sentenças ou termos por relações de coocorrência e centralidade. Métodos de avaliação como ROUGE procuraram comparar resumos automáticos com resumos de referência. A diferença trazida pelos modelos de linguagem contemporâneos é que eles produzem resumos abstrativos fluentes, isto é, textos que não apenas extraem sentenças, mas reescrevem e combinam informações. Essa fluência aumenta a utilidade comunicacional, mas também eleva o risco de introdução de elementos não presentes nos dados (LUHN, 1958; MIHALCEA; TARAU, 2004; LIN, 2004).

Em estudos cientométricos, a sumarização deve ser orientada por perguntas analíticas. Uma pergunta descritiva pode solicitar: "quais são os temas mais recorrentes nos resumos?". Uma pergunta metodológica pode solicitar: "quais desenhos de pesquisa aparecem com maior frequência?". Uma pergunta sociogeográfica pode solicitar: "quais países, regiões ou populações são mencionados?". Uma pergunta temporal pode solicitar: "quais termos aparecem nos anos iniciais e quais aparecem nos anos recentes?". Uma pergunta crítica pode solicitar: "quais aspectos são pouco estudados ou aparecem apenas marginalmente?". A qualidade da sumarização depende da clareza da pergunta e da adequação do campo textual utilizado.

A sumarização de grandes conjuntos de resumos deve ser precedida por uma preparação do *corpus*. É preciso remover duplicatas, garantir que os metadados foram exportados corretamente, separar campos textuais, corrigir problemas de codificação, preservar identificadores únicos e registrar a origem dos dados. Em muitos casos, é recomendável trabalhar por lotes, pois modelos de linguagem têm limites de contexto. A divisão por lotes, no entanto, introduz outro problema: cada lote pode gerar uma síntese local diferente. Por isso, a metodologia deve prever uma segunda etapa de metassumarização, na qual sínteses parciais são comparadas, agrupadas e auditadas. Essa etapa deve exigir que a IA indique recorrências, divergências e incertezas, não apenas produza uma narrativa harmonizada (NOWAKOWSKA, 2025).

Um procedimento robusto consiste em organizar a sumarização em três níveis. No primeiro nível, cada lote de resumos recebe uma síntese descritiva padronizada. No segundo, as sínteses dos lotes são agregadas em categorias temporárias. No terceiro, uma amostra de documentos é lida por humanos para verificar se as categorias fazem sentido. Essa arquitetura reduz a dependência de uma única resposta e favorece a rastreabilidade. A IA pode ser solicitada a fornecer justificativas ancoradas em trechos ou palavras recorrentes, mas o pesquisador deve evitar que essas justificativas sejam tratadas como evidência sem conferência. A saída deve ser considerada mapa preliminar, não conclusão final.

Ao sumarizar resumos, é útil distinguir cobertura, saliência e interpretação. Cobertura diz respeito a quais tópicos aparecem no *corpus*. Saliência diz respeito à importância relativa desses tópicos em termos de recorrência, centralidade ou presença em documentos influentes. Interpretação diz respeito ao significado desses tópicos para a dinâmica da área. A IA pode apoiar as três dimensões, mas com graus diferentes de confiabilidade. A cobertura pode ser verificada por contagens de termos e amostragem. A saliência exige dados quantitativos complementares. A interpretação depende de conhecimento de domínio. Quando um modelo afirma que um tema é "central", deve-se perguntar: central em frequência, central em citações, central em rede, central em debate teórico ou central apenas na narrativa gerada?

A sumarização também deve respeitar a heterogeneidade disciplinar. Em algumas áreas, resumos são estruturados e informativos; em outras, são breves, retóricos ou pouco metodológicos. Em ciências biomédicas, pode haver indicação clara de participantes, intervenção, comparação e resultado. Em humanidades, o resumo pode enfatizar debate teórico, objeto cultural ou argumento interpretativo. Em engenharia, pode privilegiar método, protótipo, desempenho e aplicação. Assim, um *prompt* genérico tende a favorecer disciplinas com resumos padronizados e pode sub-representar campos com estilos argumentativos

diferentes. O *prompt* cientométrico deve solicitar explicitamente que o modelo considere variações de linguagem e registre quando a informação não estiver disponível.

A sumarização de grandes conjuntos de resumos possui valor especial na fase exploratória de um estudo. Ela pode indicar se a estratégia de busca recuperou documentos coerentes com o escopo, se há ruído temático relevante, se alguns termos estão dominando indevidamente o *corpus* ou se determinados subcampos foram omitidos. Também pode apoiar a comunicação inicial com especialistas, apresentando um panorama para discussão. Contudo, não deve ser usada isoladamente para excluir documentos, definir categorias finais ou elaborar conclusões substantivas. A função principal é orientar a atenção analítica, não automatizar o juízo científico.

Em termos práticos, recomenda-se que a saída da IA em tarefas de sumarização contenha sempre pelo menos cinco campos: síntese geral, temas principais, temas secundários, evidências textuais ou termos indicativos, e incertezas. Quando a tarefa envolve grandes conjuntos, é recomendável incluir ainda número de documentos do lote, período temporal, fonte de dados e eventuais limitações. A ausência de tais campos dificulta a comparação entre lotes e impede que a síntese seja auditada posteriormente. Uma síntese cientometricamente útil é aquela que pode ser confrontada com os dados, e não apenas aquela que soa plausível.

A avaliação da sumarização deve combinar critérios qualitativos e quantitativos. Entre os critérios qualitativos estão fidelidade ao *corpus*, ausência de inferências não autorizadas, clareza, preservação da diversidade e identificação de casos ambíguos. Entre os critérios quantitativos estão cobertura de categorias esperadas, concordância com amostras humanas, estabilidade entre execuções e relação com frequências de termos. A literatura sobre avaliação de sumarização automática, incluindo métricas como ROUGE, mostra que a comparação lexical com resumos de referência pode ser útil, mas insuficiente para avaliar sínteses científicas complexas. Em cientometria, a fidelidade metodológica e a utilidade analítica são tão importantes quanto a semelhança textual (LIN, 2004).

Por fim, a sumarização com IA deve ser documentada como operação analítica. O relatório deve informar quais campos foram usados, quantos registros foram processados, como os lotes foram formados, qual *prompt* foi aplicado, quais instruções limitavam inferências, como as respostas foram agregadas e que validação humana foi realizada. Sem essa documentação, o uso de IA se torna invisível e impede a reprodutibilidade. Com documentação, a IA deixa de ser uma "caixa de texto" e passa a ser componente rastreável do fluxo cientométrico.

Prompt 14.2.1 Sumarização de um lote de resumos

Atue como analista de literatura científica em um estudo cientométrico. Receba o conjunto de resumos abaixo e produza uma síntese descritiva, sem acrescentar informações externas. Organize a resposta em:

(1) temas principais;

(2) temas secundários;

(3) métodos mencionados;

(4) objetos ou populações estudadas;

(5) termos recorrentes; (

6) lacunas ou ausências percebidas apenas com base nos resumos;

(7) pontos incertos que exigem revisão humana.

Para cada tema, indique evidências textuais breves e não afirme causalidade nem importância científica sem dados de citação, rede ou validação especializada.

14.3 Classificação temática assistida

A classificação temática assistida por IA é uma operação central para transformar um *corpus* textual em categorias analíticas. Diferentemente da simples contagem de palavras, a classificação temática procura reconhecer sentidos, relações, objetos de estudo, abordagens e enquadramentos teóricos. Em estudos cientométricos, essa tarefa pode apoiar a construção de mapas conceituais, a identificação de subcampos, a comparação entre áreas, a análise temporal de temas e a redação de diagnósticos. A IA pode ser usada para sugerir categorias, aplicar um esquema predefinido, classificar documentos em múltiplos rótulos ou indicar casos ambíguos. Entretanto, cada uma dessas modalidades exige regras próprias de validação.

Há duas estratégias gerais de classificação: indutiva e dedutiva. Na estratégia indutiva, o pesquisador fornece títulos, resumos e palavras-chave para que a IA proponha categorias emergentes. Essa modalidade é útil quando a área é pouco conhecida, interdisciplinar ou recente. Na estratégia dedutiva, o pesquisador define previamente categorias e critérios de inclusão; a IA então classifica documentos de acordo com esse esquema. Essa modalidade é mais adequada quando existe taxonomia consolidada, protocolo de revisão, classificação institucional ou interesse comparativo. Em ambos os casos, a IA deve apresentar justificativa textual e grau de confiança, permitindo revisão humana.

A classificação temática não deve ser confundida com modelagem de tópicos. Modelos estatísticos como LDA identificam padrões de coocorrência de palavras e estimam tópicos latentes; modelos baseados em *embeddings* agrupam documentos por proximidade semântica; modelos generativos podem atribuir rótulos interpretáveis a conjuntos de textos. Cada abordagem tem vantagens e limitações. A modelagem de tópicos é útil para exploração sistemática e replicável, mas seus tópicos podem ser difíceis de interpretar. A classificação por IA generativa é comunicativamente clara, mas pode criar rótulos sedutores sem base suficiente. Uma abordagem robusta combina métodos: usa modelagem para estrutura inicial, IA para rotulagem preliminar e especialistas para validação (BLEI; NG; JORDAN, 2003; GRIFFITHS; STEYVERS, 2004).

A literatura em processamento de linguagem natural mostra que modelos treinados em *corpora* científicos podem melhorar o desempenho em tarefas específicas. SciBERT, por exemplo, foi desenvolvido para textos científicos e avaliado em tarefas de classificação e extração. BioBERT mostrou ganhos em mineração de textos biomédicos. Esses modelos indicam que o vocabulário científico possui propriedades diferentes do texto geral. No contexto dos grandes modelos de linguagem, essa constatação permanece relevante: *prompts* aplicados a literatura científica devem reconhecer especificidade terminológica, abreviações, nomes de métodos, estruturas de resumo e padrões disciplinares. A IA pode parecer fluente, mas fluência não equivale a domínio disciplinar (BELTAGY; LO; COHAN, 2019; LEE *et al.*, 2020).

O primeiro passo para uma classificação assistida por IA é a definição da unidade classificatória. Pode-se classificar o documento inteiro, o resumo, a frase, o parágrafo, a palavra-chave, a referência citada ou o *cluster*. Classificar o documento inteiro tende a produzir categorias mais gerais. Classificar frases pode melhorar a identificação de métodos, populações e resultados, mas aumenta o custo de processamento. Classificar *clusters* exige que a rede ou o agrupamento já tenham sido gerados. A escolha da unidade deve estar alinhada à pergunta cientométrica. Uma pesquisa sobre evolução temática pode usar documentos por ano; uma pesquisa sobre métodos pode usar frases de metodologia; uma pesquisa sobre interdisciplinaridade pode classificar referências ou periódicos.

O segundo passo é estabelecer a granularidade das categorias. Categorias muito amplas, como "tecnologia", "educação" ou "saúde", podem ser inúteis para análise fina. Categorias excessivamente específicas podem gerar fragmentação e baixa consistência. A IA

tende a obedecer ao nível de granularidade sugerido pelo *prompt*; por isso, a instrução deve indicar número máximo de categorias, possibilidade de categoria secundária, tratamento de documentos híbridos e critérios para casos duvidosos. A classificação *multilabel*, na qual um documento pode pertencer a mais de uma categoria, é frequentemente mais adequada para literatura científica interdisciplinar do que a classificação exclusiva.

O terceiro passo é construir um *codebook*. O *codebook* é um documento que define categorias, critérios de inclusão, critérios de exclusão, exemplos positivos, exemplos negativos e regras de decisão. Quando a IA é usada sem *codebook*, ela pode alterar o significado das categorias entre lotes, especialmente em *corpus* longo. Quando o *codebook* é fornecido no *prompt*, aumenta-se a estabilidade, embora não se elimine a necessidade de conferência. O *codebook* também zfacilita a comparação entre classificadores humanos e IA. Em estudos reprodutíveis, ele deve ser anexado ao protocolo, assim como as versões dos *prompts* e as amostras de calibração.

A classificação assistida pode seguir um desenho em três rodadas. Na primeira rodada, a IA propõe categorias emergentes com base em uma amostra diversificada de documentos. Na segunda, o pesquisador revisa, consolida e define o *codebook*. Na terceira, a IA aplica o *codebook* ao *corpus* completo ou a lotes sucessivos, marcando incertezas. Após essa aplicação, uma amostra estratificada deve ser revisada por especialistas. Essa sequência reduz o risco de impor categorias prematuras e também evita aceitar categorias automáticas sem controle. Em termos metodológicos, a IA funciona como instrumento de exploração e aplicação, enquanto a comunidade de especialistas fornece a validação semântica.

Um problema recorrente é a estabilidade da classificação. Um mesmo modelo pode classificar o mesmo documento de modo diferente em execuções distintas, especialmente quando o *prompt* é aberto, a temperatura é alta ou as categorias são ambíguas. Para controlar esse problema, recomenda-se usar instruções determinísticas, limitar o número de categorias, exigir justificativas curtas e repetir a classificação de amostras críticas. Quando houver divergências entre execuções, o documento deve ser marcado para revisão humana. A instabilidade não é apenas defeito técnico; ela sinaliza que o documento pode ser conceitualmente híbrido, que as categorias são mal definidas ou que o *prompt* permite interpretações amplas demais.

A classificação temática também deve preservar a possibilidade de "não classificar". Em muitos protocolos, a IA é forçada a escolher uma categoria mesmo quando a evidência é insuficiente. Isso produz falsa precisão. Um *prompt* cientometricamente responsável deve incluir opções como "informação insuficiente", "fora do escopo", "ambíguo" e "revisão humana necessária". Essa prática evita que a classificação se transforme em atribuição arbitrária. Em contextos de avaliação institucional, política científica ou decisão estratégica, a categoria "ambíguo" pode ser tão importante quanto as categorias substantivas, pois revela limites do *corpus* ou do esquema classificatório.

A classificação assistida por IA deve ser interpretada como mediação entre texto e estrutura. Ao classificar documentos, o pesquisador cria uma matriz que poderá ser cruzada com anos, autores, instituições, países, citações, periódicos e redes. Essa matriz pode revelar crescimento de temas, concentração institucional, especialização regional ou mudança de métodos. Porém, tais análises só serão confiáveis se a classificação for suficientemente válida. Uma taxa elevada de erro em categorias raras pode distorcer a interpretação de temas emergentes. Da mesma forma, uma categoria com muitos documentos heterogêneos pode parecer dominante sem representar um programa de pesquisa coerente.

Em suma, a classificação temática assistida por IA é poderosa quando tratada como processo iterativo, documentado e validado. Seu valor não está apenas em economizar tempo, mas em explicitar critérios, testar categorias e tornar visíveis zonas de ambiguidade. O produto ideal não é uma tabela final aparentemente perfeita, mas uma classificação acompanhada de justificativas, graus de confiança, amostras revisadas e descrição transparente das decisões tomadas.

Prompt 14.3.1 Classificação temática assistida com justificativa

Classifique os resumos abaixo em até seis categorias temáticas. Para cada documento, apresente:

- identificador do documento;

- categoria principal; categoria secundária, se houver;

- justificativa textual baseada no título, resumo ou palavras-chave;

- grau de confiança em escala de 1 a 5;

- indicação de casos duvidosos.

Não invente informações ausentes. Use a categoria "informação insuficiente" quando o resumo não permitir decisão segura.

Ao final, liste categorias que parecem sobrepostas e recomende ajustes no *codebook*.

14.4 Extração de métodos, objetos e populações estudadas

A extração de métodos, objetos e populações estudadas amplia a classificação temática ao deslocar a atenção do "assunto" para a arquitetura da pesquisa. Dois artigos podem tratar do mesmo tema, mas usar métodos diferentes; podem investigar objetos distintos; podem estudar populações ou contextos geográficos diversos; podem mobilizar escalas temporais incompatíveis. Em cientometria, essa extração permite construir diagnósticos mais ricos do que simples contagens de publicações por tema. É possível perguntar quais métodos dominam determinada área, quais populações são negligenciadas, quais objetos empíricos concentram evidências, quais países são estudados e quais combinações metodológicas emergem ao longo do tempo.

A IA pode apoiar a extração de informações porque modelos de linguagem são capazes de reconhecer padrões linguísticos em resumos: expressões como "este estudo analisa", "foi conduzida uma revisão", "os participantes foram", "utilizou-se um *survey*", "a amostra incluiu", "foram coletados dados" ou "os resultados indicam" sinalizam componentes da pesquisa. Entretanto, a confiabilidade da extração varia muito conforme o estilo do resumo. Em resumos estruturados, a tarefa é mais simples; em resumos narrativos, o modelo pode inferir elementos não explicitados. Por isso, *prompts* de extração devem exigir distinção entre informação explícita e inferência provável.

Um campo consolidado para pensar essa tarefa é a extração PICO em literatura biomédica: população, intervenção, comparação e desfecho. *Corpora* anotados, como o EBM-NLP, foram criados para apoiar o desenvolvimento de métodos de processamento de linguagem natural para identificar esses elementos em resumos de ensaios clínicos. Embora a cientometria não se limite à saúde, a lógica é útil: definir previamente quais elementos interessam, indicar exemplos, separar evidência textual de interpretação e validar contra anotações humanas. Em

ciências sociais, por exemplo, os elementos podem ser população, contexto, método, teoria e objeto. Em engenharia, podem ser problema, solução, técnica, base de dados e métrica de desempenho (NYE *et al.*, 2018).

A extração de métodos exige atenção terminológica. Termos como "análise de conteúdo", "estudo de caso", "experimento", "simulação", "*survey*", "revisão sistemática", "análise bibliométrica", "entrevista" ou "métodos mistos" possuem significados disciplinares específicos. A IA pode reconhecê-los, mas também pode normalizar indevidamente métodos diferentes sob rótulos amplos. Por exemplo, "modelagem de equações estruturais" não deve ser simplesmente convertida em "método quantitativo" se a pergunta do estudo exige granularidade. O *prompt* deve especificar o nível de detalhe desejado e solicitar que o modelo preserve o termo original ao lado do termo normalizado.

A extração de objetos de estudo é igualmente delicada. Em algumas áreas, o objeto é material ou empírico: pacientes, escolas, cidades, empresas, sensores, documentos, políticas, algoritmos. Em outras, é conceitual: justiça, inovação, memória, governança, identidade, risco. A IA pode confundir tema e objeto. Um artigo sobre "governança da inteligência artificial em universidades" pode ter como tema a governança da IA, como objeto as universidades e como método análise documental. Se o *prompt* não diferenciar essas dimensões, a classificação tende a misturá-las. Essa distinção é essencial para análises cientométricas que buscam mapear não apenas sobre o que se escreve, mas o que efetivamente se investiga.

A extração de populações estudadas permite revelar assimetrias e invisibilidades. Em estudos biomédicos, pode indicar idade, sexo, condição clínica ou grupo de pacientes. Em educação, pode identificar estudantes, docentes, gestores, escolas públicas ou ensino superior. Em políticas públicas, pode indicar municípios, comunidades, beneficiários ou órgãos governamentais. Em ciência da informação, pode identificar pesquisadores, usuários, bibliotecas, periódicos ou repositórios. A IA pode acelerar essa identificação, mas deve registrar quando a população não é informada. A ausência de população explícita é um dado relevante; não deve ser preenchida por suposição.

Além de métodos, objetos e populações, a IA pode extrair teorias, instrumentos, bases de dados, países, setores econômicos, tipos de documento, fontes de financiamento e softwares. Essa extração transforma resumos em uma base analítica rica. Contudo, cada campo adicionado aumenta a chance de erro. Uma prática recomendável é começar com poucos campos de alta relevância, validar a qualidade da extração e expandir gradualmente. Em vez de pedir à IA que extraia vinte dimensões de uma só vez, o pesquisador pode usar cadeias de *prompts*: primeiro, métodos, depois objetos, depois populações, depois contextos, sempre com amostras de validação.

A extração deve produzir saídas tabulares. Cada linha corresponde a um documento e cada coluna corresponde a uma variável extraída. Além do valor extraído, recomenda-se incluir evidência textual, tipo de evidência, grau de confiança e observação. A coluna "tipo de evidência" pode indicar "explícita", "inferida" ou "ausente". Essa estrutura permite auditoria posterior e integração com ferramentas estatísticas. Sem evidência textual, a tabela se torna opaca; com evidência textual, o pesquisador pode revisar rapidamente os casos problemáticos.

A validação de extrações feitas por IA deve considerar sensibilidade e precisão. Se o objetivo é identificar todos os estudos que usam determinado método, pode-se preferir alta sensibilidade, aceitando mais falsos positivos para revisão humana. Se o objetivo é produzir estatísticas finais, a precisão torna-se mais importante. O *prompt* deve refletir essa escolha. Uma instrução como "se houver dúvida, marque para revisão humana" tende a aumentar

a confiabilidade. Uma instrução como "classifique todos os documentos" pode aumentar a completude aparente, mas comprometer a validade.

A extração também se beneficia de vocabulários controlados, ontologias e listas de sinônimos. Quando se fornece ao modelo uma lista de métodos esperados ou de populações de interesse, a saída tende a ser mais consistente. Entretanto, listas fechadas podem invisibilizar métodos emergentes ou terminologia local. Uma solução é permitir uma coluna de "termo não previsto", na qual a IA registra expressões que não se encaixam no vocabulário inicial. Esses termos podem alimentar a revisão do protocolo e revelar mudanças terminológicas na área.

Em estudos cientométricos assistidos por IA, a extração de métodos, objetos e populações tem valor estratégico porque aproxima indicadores quantitativos de conteúdo substantivo. Ela permite responder não apenas "quanto se publicou", mas "que tipo de conhecimento foi produzido", "sobre quem", "com quais métodos" e "em quais contextos". Essa aproximação é indispensável para evitar diagnósticos puramente volumétricos. Uma área pode crescer em número de publicações e, ainda assim, manter baixa diversidade metodológica ou concentrar-se em poucas populações. A IA pode ajudar a revelar essas configurações, desde que a extração seja controlada e verificável.

O uso responsável da IA nessa etapa exige que o pesquisador trate cada saída como anotação preliminar. O modelo deve ser orientado a reconhecer incerteza, indicar evidências e não preencher lacunas por plausibilidade. Quando o documento não informa método, população ou objeto, a resposta correta é registrar ausência, não inferir. Em pesquisa científica, a ausência de informação também é informação sobre a qualidade dos resumos, a transparência metodológica da área e os limites da base de dados.

Prompt 14.4.1 Extração de métodos, objetos e populações

Atue como codificador metodológico em um estudo cientométrico. Para cada resumo fornecido, extraia os seguintes campos:

- método de pesquisa;
- objeto de estudo
- população ou *corpus* analisado;
- contexto geográfico ou institucional;
- tipo de evidência textual usada; grau de confiança;
- observação sobre ambiguidade.

Use apenas informações presentes no título, resumo e palavras-chave. Quando um campo não estiver explícito, escreva "não informado".

Diferencie informação explícita de inferência provável e marque todas as inferências para revisão humana.

14.5 Identificação de lacunas de pesquisa

A identificação de lacunas de pesquisa é uma das tarefas mais solicitadas a modelos de linguagem, mas também uma das mais arriscadas. Uma lacuna não é simplesmente aquilo que o modelo não mencionou ou aquilo que parece pouco frequente em um conjunto de resumos. Lacuna de pesquisa é uma relação entre estado do conhecimento, relevância teórica ou prática,

evidência disponível, limitações metodológicas e necessidades futuras. Em cientometria, a IA pode apoiar a detecção preliminar de ausências e sub-representações, mas a validação de lacunas exige triangulação com dados, especialistas e literatura substantiva.

Existem diferentes tipos de lacuna. A lacuna temática ocorre quando determinado assunto é pouco estudado em comparação com sua relevância. A lacuna metodológica aparece quando uma área é dominada por poucos desenhos de pesquisa. A lacuna populacional ocorre quando grupos, territórios ou contextos são sub-representados. A lacuna temporal surge quando não há atualização recente ou quando períodos críticos são negligenciados. A lacuna teórica aparece quando conceitos são aplicados sem diálogo com determinadas tradições. A lacuna de evidência ocorre quando há muitos estudos descritivos, mas poucos estudos avaliativos, longitudinais ou comparativos. A IA deve ser instruída a diferenciar esses tipos, pois cada um exige evidência diferente.

Um erro comum é interpretar baixa frequência como lacuna. Em um *corpus* delimitado por termos de busca, a baixa frequência pode resultar de estratégia inadequada, variação terminológica, cobertura limitada da base, idioma, exclusão de literatura cinzenta ou uso de palavras-chave diferentes. A IA, se não for orientada, pode transformar artefatos de busca em oportunidades de pesquisa. Por isso, todo *prompt* de lacuna deve incluir uma pergunta de auditoria: "esta ausência pode ser produto do recorte?". A resposta deve considerar fonte de dados, período, idioma, campos pesquisados, descritores e critérios de inclusão.

A IA pode ser útil para identificar lacunas quando trabalha sobre resultados já estruturados. Por exemplo, uma matriz que cruza temas e anos pode revelar temas emergentes com baixa consolidação. Uma matriz que cruza métodos e temas pode revelar subcampos com pouca diversidade metodológica. Uma matriz que cruza países e objetos pode indicar concentração geográfica. Uma tabela de extração de populações pode mostrar grupos ausentes. Nesses casos, a IA não precisa inventar lacunas; ela interpreta padrões observáveis. O *prompt* deve fornecer os dados e solicitar que o modelo separe evidência observada, hipótese explicativa e recomendação cautelosa.

A identificação de lacunas também exige distinguir lacuna científica de lacuna editorial. Alguns temas podem estar ausentes de periódicos indexados em bases internacionais, mas presentes em relatórios técnicos, dissertações, livros, anais locais ou literatura em outros idiomas. Em áreas aplicadas, a literatura cinzenta pode ser fundamental. Em humanidades, livros e capítulos podem ter centralidade maior que artigos. Em políticas públicas, documentos institucionais podem conter evidências relevantes. Assim, um estudo cientométrico baseado apenas em artigos indexados deve evitar afirmar que um tema "não é estudado"; deve afirmar, quando apropriado, que "não aparece com frequência no *corpus* analisado".

A IA pode apoiar a elaboração de uma tipologia de lacunas. A partir de resumos, categorias e indicadores, o modelo pode propor lacunas temáticas, metodológicas, populacionais e geográficas. Mas cada lacuna deve ser acompanhada por evidência, nível de confiança e verificação necessária. Por exemplo: "há aparente baixa presença de estudos longitudinais sobre o tema X; evidência: apenas três documentos classificados como longitudinais; verificação necessária: confirmar classificação metodológica e revisar termos relacionados a acompanhamento temporal". Esse formato transforma a lacuna em hipótese auditável.

Outro cuidado importante é evitar que a IA confunda lacuna com agenda normativa. Modelos de linguagem tendem a sugerir recomendações amplas, como "são necessários mais estudos interdisciplinares", "pesquisas futuras devem considerar diversidade" ou "é preciso desenvolver abordagens longitudinais". Essas afirmações podem ser razoáveis, mas frequentemente são genéricas. Um capítulo ou relatório cientométrico deve produzir lacunas

específicas, associadas a padrões do *corpus*. A recomendação deve explicar por que a lacuna importa, que evidência a sustenta, que limites impedem afirmação forte e que tipo de estudo poderia preenchê-la.

A detecção de lacunas pode ser integrada à análise temporal. Um tema que cresce rapidamente, mas tem baixa diversidade metodológica, pode indicar oportunidade de amadurecimento. Um tema que desaparece pode representar declínio real, substituição terminológica ou mudança de indexação. Um país com crescimento de publicações, mas baixa participação em redes internacionais, pode indicar lacuna de colaboração. Um conceito muito citado, mas pouco operacionalizado nos resumos, pode indicar lacuna entre discurso teórico e investigação empírica. A IA pode ajudar a formular essas hipóteses, mas não deve apresentá-las como diagnóstico definitivo.

Os *prompts* para lacunas devem exigir linguagem modalizada. Expressões como "os dados sugerem", "há indício", "no *corpus* analisado", "aparentemente", "exige validação" e "pode refletir" são mais adequadas do que "comprova", "demonstra" ou "revela definitivamente". Essa prudência não enfraquece a análise; ao contrário, aumenta sua credibilidade. A cientometria lida com representações parciais da ciência publicada, e a IA trabalha sobre essas representações. A dupla mediação recomenda cautela interpretativa.

Em síntese, a IA pode ser poderosa para gerar hipóteses de lacunas, especialmente quando o *corpus* é grande e multidimensional. Seu valor está em organizar padrões dispersos e sugerir perguntas para investigação posterior. Entretanto, lacunas só se tornam achados quando passam por validação metodológica, comparação com o desenho da busca, revisão por especialistas e, quando necessário, leitura integral de documentos-chave. A lacuna cientométrica responsável é sempre acompanhada de escopo, evidência, incerteza e condição de validade.

Prompt 4 - Identificação cautelosa de lacunas de pesquisa

Com base na matriz de temas, métodos, populações, países e anos fornecida, identifique possíveis lacunas de pesquisa. Para cada lacuna, informe:

- tipo de lacuna;
- evidência observada no *corpus*;
- possível explicação alternativa;
- risco de ser artefato de busca ou indexação;
- nível de confiança;
- verificação necessária.

Não use linguagem conclusiva. Diferencie "ausência no *corpus* analisado" de "ausência na literatura".

14.6 Comparação entre revisões sistemáticas e estudos cientométricos

Revisões sistemáticas e estudos cientométricos compartilham a necessidade de busca transparente, critérios explícitos e documentação rigorosa, mas possuem objetivos distintos. A revisão sistemática busca responder a uma pergunta substantiva sobre evidência, frequentemente avaliando qualidade, resultados e conclusões de estudos primários. O estudo cientométrico busca mapear padrões da produção científica, tais como crescimento, impacto, colaboração, estrutura intelectual, temas e circulação. A IA pode apoiar ambos, mas o tipo de

tarefa e o padrão de validação diferem. Confundir essas tradições pode levar a uso inadequado de resumos, indicadores e classificações.

Na revisão sistemática, a triagem de títulos e resumos é orientada por critérios de inclusão e exclusão. O objetivo é decidir se cada estudo deve ou não integrar a síntese. Na cientometria, o *corpus* pode ser mantido mais amplo para mapear a estrutura da literatura, mesmo quando alguns documentos não seriam elegíveis para uma revisão substantiva estrita. Por exemplo, uma revisão sistemática sobre eficácia de uma intervenção pode excluir estudos teóricos; um estudo cientométrico sobre a evolução do tema pode incluí-los para compreender a formação do campo. Assim, a IA deve receber instruções diferentes conforme o desenho de pesquisa.

As diretrizes PRISMA e *PRISMA-S* enfatizam transparência no relato de revisões sistemáticas, especialmente busca, seleção e fluxo de estudos. Embora não sejam protocolos cientométricos em sentido estrito, oferecem princípios úteis: registrar bases, datas, estratégias, número de registros, critérios e motivos de exclusão. Em estudos cientométricos assistidos por IA, princípios análogos devem ser aplicados à classificação e à sumarização. O pesquisador deve relatar quantos registros foram processados por IA, quantos foram revisados por humanos, quais critérios orientaram a classificação, como casos ambíguos foram tratados e quais limitações decorrem da automação (PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021).

A literatura sobre automação de revisões sistemáticas oferece lições importantes para a cientometria. Estudos sobre triagem assistida por aprendizado de máquina mostram que algoritmos podem reduzir carga de trabalho, mas dependem de treinamento, validação e métricas de desempenho. Trabalhos recentes sobre modelos de linguagem indicam potencial para triagem e extração, mas também variabilidade, instabilidade e necessidade de avaliação por domínio. A conclusão metodológica é clara: a IA pode ser auxiliar, segundo avaliador, filtro preliminar ou geradora de hipóteses; não deve ser adotada como substituto não auditado da decisão humana (COHEN *et al.*, 2006; WALLACE *et al.*, 2010; THOMAS; MCNAUGHT; ANANIADOU, 2011; TSAFNAT *et al.*, 2014; O'MARA-EVES *et al.*, 2015; MARSHALL; WALLACE, 2019; BOLAÑOS *et al.*, 2024).

Uma diferença crucial entre revisão sistemática e cientometria está na relação com a qualidade dos estudos. Revisões sistemáticas frequentemente avaliam risco de viés, desenho metodológico e força da evidência. Estudos cientométricos, por sua vez, podem analisar todos os documentos recuperados independentemente de qualidade metodológica, pois seu objeto é a produção científica publicada. Isso não significa que qualidade seja irrelevante; significa que a pergunta é diferente. Ao usar IA para classificar literatura, o pesquisador deve indicar se a tarefa é descrever o *corpus*, selecionar evidência ou avaliar qualidade. Cada finalidade requer *prompt*, critérios e validação específicos.

A sumarização também difere. Em uma revisão sistemática, a síntese deve responder à pergunta de evidência e, muitas vezes, integrar resultados. Em um estudo cientométrico, a síntese pode descrever temas, redes, métodos, instituições e evolução temporal. Uma sumarização gerada por IA que mistura resultados substantivos e padrões bibliométricos pode produzir confusão. Portanto, *prompts* devem separar síntese cientométrica de síntese de evidência. A primeira responde "como a literatura se organiza"; a segunda responde "o que a literatura conclui". Ambas são úteis, mas não são intercambiáveis.

Os estudos cientométricos podem preceder revisões sistemáticas. Um mapa temático assistido por IA pode indicar subcampos, *clusters*, períodos de crescimento e terminologias relevantes, ajudando a formular perguntas de revisão mais precisas. Da mesma forma, uma revisão sistemática pode fornecer *corpus* validado para análise cientométrica posterior. Essa complementaridade é especialmente valiosa em áreas emergentes, nas quais ainda não há

vocabulário estabilizado. A IA pode apoiar a transição entre mapeamento e revisão, mas a passagem de uma lógica para outra deve ser explicitada no protocolo.

Em termos de classificação, revisões sistemáticas costumam exigir decisão binária ou multicategórica relacionada à elegibilidade. Estudos cientométricos frequentemente exigem classificação temática, metodológica, institucional ou temporal. Um mesmo resumo pode ser "não elegível" para revisão, mas ainda relevante para mapear a discussão. Por isso, recomenda-se evitar *prompts* únicos que prometem "classificar e revisar" simultaneamente. A separação das tarefas melhora a qualidade: primeiro delimita-se *corpus*, depois classifica-se conteúdo, depois, se necessário, avalia-se elegibilidade para revisão substantiva.

A IA também pode auxiliar na comparação entre resultados de revisões e mapas cientométricos. Por exemplo, um estudo pode mostrar que determinado tema é muito publicado, mas as revisões sistemáticas indicam baixa qualidade da evidência. Outro caso pode mostrar tema pouco publicado, mas com alta concentração em estudos robustos. A integração entre análise quantitativa da produção e avaliação qualitativa da evidência é promissora, desde que os métodos sejam relatados separadamente. A IA pode ajudar a construir narrativas integradas, mas deve respeitar os limites de cada tradição.

Finalmente, a comparação entre revisões sistemáticas e estudos cientométricos reforça a importância da pergunta de pesquisa. O mesmo conjunto de documentos pode ser lido de maneiras diferentes. Um especialista em bibliometria pergunta por padrões de produção e estrutura; um revisor sistemático pergunta por evidência e qualidade; um gestor de pesquisa pergunta por capacidades e oportunidades; um especialista pergunta por debates teóricos. A IA pode adaptar a leitura a cada finalidade, mas somente se o *prompt* declarar a finalidade e limitar a inferência. Sem essa clareza, a saída tende a ser híbrida, genérica e metodologicamente imprecisa.

Prompt 14.6.1 Comparar revisão sistemática e análise cientométrica

> A partir do tema [TEMA] e do *corpus* descrito, diferencie o que poderia ser respondido por uma revisão sistemática e o que poderia ser respondido por um estudo cientométrico.
>
> Para cada abordagem, indique objetivo, unidade de análise, critérios de inclusão, tipo de classificação, tipo de síntese, papel da IA, forma de validação e principais limitações.
>
> Não misture avaliação de qualidade da evidência com mapeamento de produção científica.

14.7 Controle de qualidade em classificações feitas por IA

O controle de qualidade é a condição que transforma uma classificação feita por IA em procedimento cientificamente defensável. Sem controle de qualidade, a classificação é apenas uma saída textual. Com controle, ela se torna dado analítico passível de auditoria. O controle deve abranger o *prompt*, o *corpus*, as categorias, as respostas, a estabilidade, a revisão humana e a documentação. Em estudos cientométricos, a preocupação é ainda maior porque classificações incorretas podem afetar mapas, tendências, indicadores e recomendações institucionais.

O primeiro componente do controle de qualidade é a amostragem. Antes de aplicar a IA ao *corpus* completo, recomenda-se selecionar uma amostra representativa e outra amostra

de casos difíceis. A amostra representativa deve cobrir anos, áreas, fontes, idiomas e tipos de documento. A amostra difícil deve incluir resumos curtos, documentos interdisciplinares, termos ambíguos, categorias sobrepostas e registros com metadados incompletos. A IA deve ser testada nessas amostras, e seus erros devem orientar revisão do *prompt* e do *codebook*. Essa etapa reduz a chance de falhas sistemáticas em larga escala.

O segundo componente é a construção de padrão de referência. Uma equipe humana pode classificar uma amostra e produzir uma matriz de referência. A classificação da IA é então comparada com essa referência por medidas como acurácia, precisão, revocação e F1, ou por análise qualitativa de divergências. Em classificações *multilabel*, as métricas devem ser adaptadas. Em tarefas interpretativas complexas, a concordância não deve ser reduzida a número; deve-se analisar por que a IA divergiu. Divergências podem revelar erro do modelo, ambiguidade do texto, categoria mal definida ou discordância legítima entre especialistas.

O terceiro componente é a exigência de justificativa textual. Quando a IA classifica um documento, deve informar que trecho ou expressão motivou a decisão. Essa exigência não garante verdade, pois o modelo pode selecionar justificativas plausíveis depois da decisão; ainda assim, torna a revisão mais eficiente. O pesquisador pode verificar rapidamente se a justificativa está no resumo, se sustenta o rótulo e se há inferência indevida. Em tarefas de alta relevância, a justificativa deve ser acompanhada por citação literal curta do resumo ou por indicação de campos usados.

O quarto componente é o controle de consistência. A mesma tarefa pode ser executada mais de uma vez para uma amostra crítica. Se os resultados variarem muito, há instabilidade. O pesquisador pode reduzir temperatura, tornar o *prompt* mais específico, fornecer exemplos, dividir categorias ou exigir resposta em formato fixo. A consistência também deve ser avaliada entre lotes. Uma categoria não deve mudar de significado no meio do *corpus*. Para isso, o *codebook* deve ser fornecido em todas as execuções, e as saídas devem ser inspecionadas periodicamente.

O quinto componente é a marcação de incerteza. A IA deve ser autorizada a dizer que não sabe. Paradoxalmente, muitos *prompts* induzem falsa certeza ao exigir sempre uma categoria. Um protocolo responsável inclui campos como grau de confiança, motivo da incerteza e revisão humana necessária. Casos de baixa confiança podem ser encaminhados a especialistas. Em análises finais, é possível relatar a proporção de documentos classificados automaticamente, revisados manualmente e considerados ambíguos. Essa transparência fortalece a credibilidade do estudo.

O sexto componente é a auditoria de viés. Modelos de linguagem podem favorecer categorias mais frequentes, terminologia em inglês, áreas com maior presença no treinamento, estilos de resumo mais padronizados ou conceitos dominantes. Em *corpus* multilíngues, podem interpretar termos locais de modo inadequado. Em campos interdisciplinares, podem forçar documentos híbridos em categorias convencionais. A auditoria deve comparar desempenho por idioma, período, área, tipo de periódico e densidade do resumo. Erros assimétricos são tão importantes quanto erros gerais.

O sétimo componente é a preservação dos dados originais. O pesquisador deve manter registros brutos, dados pré-processados, saídas da IA e versões revisadas. Quando a IA normaliza termos ou classifica documentos, o valor original não deve ser apagado. Essa preservação permite refazer a análise, testar novos *prompts*, corrigir erros e responder a questionamentos. Em cientometria, onde decisões de limpeza e classificação podem alterar resultados, a preservação do histórico é parte do rigor.

O oitavo componente é a rastreabilidade das transformações. Cada documento deve manter identificador único ao longo de todo o fluxo: busca, exportação, limpeza, sumarização, classificação, revisão e análise. Sem identificadores, torna-se impossível verificar como uma decisão afetou um resultado. A IA deve receber e devolver identificadores. *Prompts* que solicitam respostas narrativas sem IDs são inadequados para análise reprodutível. O formato tabular com identificador, rótulo, justificativa e confiança é preferível à resposta em prosa livre quando a classificação alimentará estatísticas ou mapas.

O nono componente é a revisão humana proporcional ao risco. Nem toda tarefa exige revisão integral, mas tarefas com consequências avaliativas, políticas ou estratégicas exigem maior controle. Uma classificação exploratória para formular hipóteses pode usar amostragem menor. Uma classificação usada para relatório institucional, distribuição de recursos ou diagnóstico de capacidades deve ter revisão mais robusta. A intensidade da validação deve ser compatível com o uso pretendido dos resultados. Essa é uma aplicação do princípio de responsabilidade proporcional.

O décimo componente é a documentação no relatório final. O capítulo de metodologia deve informar que IA foi usada, para quais tarefas, com quais dados, sob quais *prompts*, como os resultados foram validados, quais limitações foram identificadas e como casos ambíguos foram tratados. Também deve informar se a IA gerou texto interpretativo ou apenas classificação. Essa transparência é necessária para que leitores avaliem a confiabilidade do estudo. O uso não declarado de IA em etapas analíticas pode comprometer a integridade metodológica.

O controle de qualidade não deve ser visto como obstáculo à inovação. Ao contrário, ele permite usar IA com maior segurança. Sem validação, o pesquisador pode ser enganado pela fluência. Com validação, pode explorar a escala e a velocidade dos modelos sem abandonar o julgamento científico. A combinação entre IA, métricas de desempenho, revisão humana e documentação cria um regime de confiança graduada: algumas saídas são aceitas, outras são corrigidas, outras são descartadas e outras se tornam objeto de investigação adicional.

Em síntese, classificações feitas por IA devem ser tratadas como hipóteses codificadas. Elas podem ser úteis, rápidas e informativas, mas exigem teste. O objetivo não é alcançar automação total, e sim produzir um fluxo de trabalho no qual a IA realiza tarefas repetitivas ou exploratórias enquanto humanos supervisionam decisões conceituais, corrigem ambiguidades e interpretam resultados. Essa divisão de trabalho é coerente com a proposta geral deste livro: usar *prompts* para ampliar a análise da ciência sem renunciar ao rigor científico.

Prompt 14.7.1 Auditoria de classificações feitas por IA

Revise a tabela de classificações produzida por IA.

Para cada linha, verifique se a categoria atribuída é sustentada pela justificativa textual. Marque: “válida”, “duvidosa”, “provável erro” ou “exige leitura integral”.

Identifique padrões de erro, categorias sobrepostas, documentos com baixa confiança, possíveis vieses por idioma ou área e recomendações para melhorar o *prompt* e o *codebook*.

Não corrija silenciosamente; registre a razão de cada alteração sugerida.

O *prompt* central a seguir sintetiza a abordagem do capítulo. Ele foi concebido para classificação temática de resumos em estudos cientométricos, mas pode ser adaptado a

extração metodológica, identificação de lacunas e análise temporal. Sua estrutura explicita papel, tarefa, limite de categorias, justificativa textual, grau de confiança e tratamento de casos duvidosos. Esses elementos são essenciais para transformar uma interação com IA em procedimento metodológico auditável.

Prompt 14.7.2 Prompt central para classificação de resumos com evidência e incerteza

> Classifique os resumos abaixo em até seis categorias temáticas.
>
> Para cada documento, apresente: categoria principal, categoria secundária se houver, justificativa textual baseada no resumo, grau de confiança e indicação de casos duvidosos.
>
> Use somente as informações fornecidas. Não invente método, população, país, teoria ou resultado.
>
> Quando a evidência textual for insuficiente, use a categoria "não classificável com segurança". Ao final, apresente:
>
> (1) descrição breve de cada categoria;
>
> (2) documentos que exigem revisão humana;
>
> (3) possíveis sobreposições entre categorias;
>
> (4) riscos metodológicos da classificação;
>
> (5) recomendações para melhorar o *codebook*.

14.8 Cadeia de prompts para leitura, sumarização e classificação

Um estudo cientométrico assistido por IA raramente deve depender de um único *prompt*. A abordagem mais robusta é construir uma cadeia de *prompts*, na qual cada etapa tem entrada, tarefa, saída e validação específicas. Essa cadeia pode iniciar com leitura exploratória de amostra, seguir para geração de categorias, depois para construção de *codebook*, aplicação da classificação, auditoria de casos incertos, consolidação de matriz e redação de resultados. A cadeia reduz a complexidade de cada tarefa e permite detectar erros em pontos intermediários. Além disso, facilita a documentação, pois cada *prompt* se torna parte do protocolo.

A primeira etapa da cadeia é a familiarização com o *corpus*. O objetivo é identificar temas recorrentes, variações terminológicas e possíveis ruídos. A saída não deve ser usada como classificação final. A segunda etapa é a proposição de categorias preliminares, com descrição, exemplos e critérios. A terceira etapa é a revisão humana dessas categorias. A quarta etapa é a aplicação do *codebook* consolidado. A quinta etapa é a auditoria das classificações. A sexta etapa é a análise agregada e a redação. Cada etapa deve produzir artefatos: síntese, *codebook*, tabela, lista de incertezas e registro de alterações.

Essa cadeia pode ser integrada a ferramentas estatísticas e bibliométricas. A IA pode classificar resumos; o pesquisador exporta a tabela; a tabela é cruzada com ano, país, instituição, citações e *clusters*; os resultados são visualizados; a IA pode auxiliar na redação interpretativa; e a interpretação final é revisada por humanos. O valor da IA está na conexão entre texto e matriz analítica. Contudo, essa conexão só é confiável quando o caminho entre entrada textual e indicador final é rastreável.

Em projetos de grande escala, a cadeia de *prompts* deve prever controle de versões. Alterações no *prompt* podem alterar resultados. Se o pesquisador modifica a definição de uma

categoria após classificar metade do *corpus*, precisa reclassificar os documentos anteriores ou registrar a mudança. Da mesma forma, se usa modelos diferentes em momentos diferentes, deve informar essa heterogeneidade. A reprodutibilidade não exige que todo estudo produza resultados idênticos em qualquer modelo, mas exige que as condições de produção dos dados analíticos sejam conhecidas.

A cadeia de *prompts* também deve incluir uma etapa de contraditório. Em vez de apenas pedir que a IA classifique, pode-se pedir que ela critique a própria classificação, identifique documentos de fronteira, aponte categorias redundantes e sugira evidências que faltam. Esse uso crítico é mais valioso do que o uso meramente produtivo. A IA deve ser tratada como instrumento para ampliar perguntas e detectar fragilidades, não apenas como máquina de gerar tabelas.

Prompt 14.8.1 Familiarização com o corpus

Leia a amostra de títulos, resumos e palavras-chave.

Identifique temas recorrentes, variações terminológicas, possíveis ruídos de busca, termos ambíguos e subáreas aparentes.

Não classifique documento por documento.

Produza apenas uma leitura exploratória e indique o que precisa ser confirmado por análise quantitativa ou revisão humana.

Prompt 14.8.2 Construção preliminar de codebook

Com base na amostra e nos objetivos do estudo, proponha um codebook com até oito categorias.

Para cada categoria, apresente definição, critérios de inclusão, critérios de exclusão, exemplos de termos, casos de fronteira e observações para revisão humana.

Evite categorias muito amplas e indique possíveis sobreposições.

Prompt 14.8.3 Aplicação do codebook

Aplique o *codebook* abaixo aos documentos fornecidos.

Para cada documento, informe ID, categoria principal, categoria secundária opcional, justificativa textual, confiança, evidência explícita e necessidade de revisão humana.

Não altere as categorias sem registrar sugestão em coluna separada.

Prompt 14.8.4 Consolidação interpretativa

> A partir da tabela validada de classificações e dos indicadores agregados, redija uma interpretação cientométrica.
>
> Separe descrição objetiva, interpretação provável, hipóteses, limitações e cuidados de não extrapolação.
>
> Não afirme qualidade científica, impacto social ou causalidade sem evidência apropriada.

14.9 Estudo de caso ilustrativo

Considere um estudo cientométrico sobre "inteligência artificial na educação superior". A busca recupera 4.800 registros entre 2015 e 2026, contendo título, resumo, palavras-chave, autores, afiliações, ano, periódico e citações. O objetivo inicial é compreender a evolução temática da área, identificar métodos predominantes e mapear lacunas. Uma análise puramente bibliométrica poderia indicar crescimento anual, países mais produtivos, periódicos centrais e redes de coautoria. Entretanto, para responder que tipos de IA são estudados, em quais contextos educacionais e com quais métodos, é necessário examinar o conteúdo textual dos documentos. É nesse ponto que a IA pode auxiliar.

Na primeira etapa, uma amostra estratificada por ano é enviada ao modelo para leitura exploratória. A IA identifica subtemas como aprendizagem personalizada, avaliação automática, tutores inteligentes, análise de aprendizagem, ética, integridade acadêmica, formação docente e uso de modelos generativos. O pesquisador percebe que o termo "AI" aparece em documentos muito diferentes, desde sistemas especialistas até ChatGPT. Essa constatação orienta a criação de categorias históricas: IA educacional clássica, *learning analytics*, IA generativa, ética e governança, avaliação automatizada e formação docente. A categoria "outros" é evitada inicialmente, mas casos não classificáveis são marcados para revisão.

Na segunda etapa, a IA ajuda a construir um *codebook*. Para "IA generativa", o critério de inclusão exige menção a modelos de linguagem, ChatGPT, geração de texto, geração de feedback ou ferramentas generativas. Para "*learning analytics*", exige menção a dados de aprendizagem, predição, *dashboards* ou mineração educacional. Para "ética e governança", inclui privacidade, viés, transparência, políticas e regulação. O *codebook* distingue tema principal e tema secundário, pois muitos artigos sobre IA generativa também tratam de ética. Essa decisão evita perda de informação.

Na terceira etapa, a IA classifica os 4.800 registros em lotes. Cada linha da saída contém ID, categoria principal, categoria secundária, método, objeto, população, confiança e justificativa. Os documentos com confiança baixa são separados. Uma amostra de 500 registros é revisada por dois pesquisadores. A taxa de concordância mostra bom desempenho em categorias amplas, mas baixa precisão em "formação docente", porque a IA confunde artigos sobre professores como usuários com artigos sobre programas de formação. O *codebook* é ajustado e a categoria é reclassificada. Esse exemplo mostra que erro da IA pode revelar fragilidade do esquema analítico.

Na quarta etapa, os resultados são cruzados com anos. Observa-se crescimento acentuado da categoria IA generativa a partir de 2023. A IA é então usada para redigir uma interpretação preliminar, mas o *prompt* exige diferenciar tendência real de artefato de busca. O modelo indica que o crescimento pode refletir tanto expansão substantiva quanto inclusão de termos como "ChatGPT" na estratégia de busca. O pesquisador verifica a *string* e cria

uma análise de sensibilidade sem o termo "ChatGPT". A comparação mostra que parte do crescimento decorre do termo, mas a tendência permanece quando se incluem termos mais gerais. A IA apoiou a interpretação, mas a validação dependeu de teste metodológico.

Na quinta etapa, busca-se identificar lacunas. A IA sugere baixa presença de estudos longitudinais, pouca atenção a instituições de países de baixa renda e concentração em percepção de estudantes. O pesquisador verifica os dados extraídos: poucos documentos foram classificados como longitudinal, a maioria dos países pertence a regiões de alta renda e muitos estudos usam *surveys* de percepção. A lacuna é então formulada com cautela: "no *corpus* analisado, há indícios de predominância de estudos transversais baseados em percepção, com menor presença de desenhos longitudinais e contextos institucionais periféricos". A formulação evita afirmar que a literatura mundial carece desses estudos sem verificar outras fontes.

Esse caso ilustra a tese do capítulo. A IA não substitui o estudo cientométrico; ela atua como componente de uma cadeia analítica. Sem *prompts* estruturados, produziria sínteses genéricas. Com *prompts*, *codebook*, amostragem, revisão humana e análise de sensibilidade, contribui para transformar grandes conjuntos de resumos em informações interpretáveis. O ganho principal não é apenas velocidade, mas capacidade de documentar o caminho entre *corpus* textual, categorias, indicadores e narrativa final.

14.10 Notas metodológicas

Quadro metodológico: boas práticas

- Preservar identificadores únicos em todas as etapas da leitura assistida por IA, evitando que resumos processados em lotes percam vínculo com metadados originais.
- Separar tarefas de sumarização, classificação, extração e interpretação, pois cada uma exige *prompt*, saída e validação próprios.
- Usar categorias provisórias na etapa exploratória e categorias documentadas no momento de aplicação ao *corpus* completo.
- Incluir sempre campos de justificativa textual, grau de confiança e indicação de revisão humana necessária.
- Evitar *prompts* que obriguem classificação quando o texto é insuficiente; permitir respostas como "não informado" e "ambíguo".
- Validar amostras com especialistas e relatar métricas ou análise qualitativa de divergências.
- Comparar resultados de IA com métodos bibliométricos tradicionais, como coocorrência de termos, *clusters*, anos e frequências.
- Documentar modelo, data de execução, *prompt*, parâmetros disponíveis, divisão de lotes e alterações humanas.
- Não usar a IA para fabricar referências, preencher campos ausentes ou inferir conclusões não sustentadas pelo *corpus*.
- Tratar a saída da IA como evidência intermediária e não como achado final sem validação.

Quadro metodológico: erros frequentes

- Erro 1: solicitar uma síntese geral de muitos resumos sem informar a pergunta cientométrica. Esse erro gera textos amplos, descritivos e pouco úteis para análise. A correção consiste em declarar o objetivo, a unidade de análise, o tipo de saída e as limitações desejadas.

- Erro 2: aceitar categorias propostas pela IA como taxonomia final. Categorias automáticas podem refletir viés do modelo, frequência lexical ou convenções genéricas. A correção é revisar categorias, criar *codebook* e testar em amostra.
- Erro 3: confundir classificação temática com avaliação de qualidade. Um documento bem classificado não é necessariamente metodologicamente robusto. A correção é separar mapeamento cientométrico e avaliação crítica da evidência.
- Erro 4: ignorar documentos ambíguos. Casos de fronteira podem revelar interdisciplinaridade, mudança terminológica ou fragilidade do protocolo. A correção é criar fila de revisão humana e relatar sua proporção.
- Erro 5: usar IA para identificar lacunas sem verificar se a lacuna é artefato de busca. A correção é testar variações da estratégia, revisar termos, comparar bases e consultar especialistas.
- Erro 6: apagar dados originais após normalização. A correção é manter campos originais e campos normalizados, preservando rastreabilidade.
- Erro 7: redigir resultados com linguagem forte demais. A correção é exigir que a IA separe dado observado, inferência e hipótese, evitando causalidade não sustentada.
- Erro 8: não relatar o uso de IA. A correção é incluir na metodologia a descrição das tarefas automatizadas, *prompts*, validação e limitações.

14.11 Roteiro para documentação reprodutível

Figura 14 - Roteiro para documentação reprodutível

Fonte: Imagem gerada por IA, ChatGPT (2026).

A documentação reprodutível deve começar antes da primeira execução do modelo. O pesquisador registra objetivo, *corpus*, campos textuais, critérios de inclusão, estratégia de busca, versão da base, data de exportação e formato dos dados. Em seguida, registra a preparação dos dados: remoção de duplicatas, normalização de codificação, exclusão de registros sem resumo e divisão em lotes. Cada lote deve ter identificador. A IA deve receber esses identificadores e devolvê-los na saída.

O protocolo deve armazenar todos os *prompts* utilizados. *Prompts* alterados devem receber número de versão. Quando um *prompt* é testado e descartado, sua exclusão também pode ser documentada, especialmente se influenciou decisões. As respostas da IA devem ser preservadas em formato bruto, antes de correções humanas. Isso permite auditoria e comparação. A versão corrigida deve registrar quem revisou, quando revisou e que alterações realizou. Em projetos colaborativos, recomenda-se manter planilha de decisões metodológicas

Para classificações, a documentação deve incluir *codebook*, exemplos, critérios de decisão e métricas de validação. Para sumarizações, deve incluir lote, número de documentos, campos usados e sínteses geradas. Para extrações, deve incluir campos extraídos, definição de evidência explícita e tratamento de ausência. Para lacunas, deve incluir dados que sustentam a hipótese, explicações alternativas e verificações realizadas. Essa documentação transforma o uso de IA em método examinável.

Além disso, o relatório final deve incluir seção de limitações específica sobre IA. Essa seção pode mencionar dependência da qualidade dos resumos, possibilidade de erro de classificação, instabilidade de respostas, limites de contexto, risco de viés linguístico e necessidade de validação especializada. A transparência não diminui o valor do estudo; ela permite que leitores compreendam seu alcance. Em um livro sobre cientometria com IA, essa postura é central: a inovação metodológica deve vir acompanhada de responsabilidade epistemológica.

14.12 Síntese conclusiva

A IA para leitura, sumarização e classificação de literatura científica representa uma ampliação relevante do repertório metodológico da cientometria. Ela permite lidar com grandes conjuntos de resumos, construir categorias exploratórias, extrair métodos e populações, formular hipóteses de lacunas e apoiar a redação de narrativas sobre campos científicos. No entanto, a mesma capacidade de gerar textos coerentes pode ocultar erros, inferências indevidas e simplificações. O desafio, portanto, não é apenas usar IA, mas desenhar protocolos que transformem sua produção em elemento verificável do processo de pesquisa.

O capítulo mostrou que sumarizar não é sintetizar evidência; classificar não é validar categorias; extrair informações não é compreender integralmente um estudo; identificar lacunas não é demonstrar ausência na literatura; e redigir narrativas não é interpretar cientificamente sem dados. Esses limites não reduzem a utilidade da IA. Ao contrário, definem o espaço legítimo de sua aplicação. A IA é mais útil quando orienta a atenção, estrutura dados textuais, explicita incertezas e acelera tarefas repetitivas. É menos confiável quando substitui julgamento, inventa referências, preenche ausências ou transforma padrões parciais em conclusões fortes.

O papel dos *prompts* é decisivo. Um *prompt* adequado não apenas solicita uma tarefa; ele define unidade de análise, contexto, critérios, formato, limites, grau de confiança, justificativa e revisão humana. Em estudos cientométricos, *prompts* são parte do método. Devem ser escritos, testados, versionados e relatados. Ao tratar *prompts* como instrumentos de pesquisa, o analista desloca o uso da IA do improviso para a metodologia. Esse deslocamento é a condição para que a inteligência artificial contribua para uma cientometria mais escalável, transparente e crítica.

Em última instância, a leitura assistida por IA não elimina a dimensão interpretativa da ciência. Ela torna mais visíveis certas regularidades textuais, mas não decide seu significado. A classificação automática pode organizar um campo, mas não substitui o debate teórico sobre suas fronteiras. A sumarização pode condensar milhares de resumos, mas não substitui a leitura atenta de obras centrais. A identificação de lacunas pode orientar agendas, mas não substitui a deliberação coletiva sobre relevância científica e social. A cientometria com IA será metodologicamente robusta na medida em que combinar escala computacional, validação humana e responsabilidade interpretativa.

Apêndice A.14 - Notas complementares

A classificação assistida por IA deve ser pensada como uma prática sociotécnica. Ela depende da qualidade dos dados, da clareza dos objetivos, da estabilidade do vocabulário, da competência dos especialistas e da forma como os resultados serão usados. Em contextos de avaliação de programas, a tolerância ao erro é menor do que em estudos exploratórios. Em contextos de inovação, categorias abertas podem ser mais úteis do que taxonomias fechadas. Em contextos de política científica, a transparência do procedimento é tão importante quanto a sofisticação técnica.

A sumarização de literatura em larga escala é mais confiável quando vinculada a campos específicos do registro bibliográfico. O título indica enquadramento; o resumo oferece narrativa condensada; as palavras-chave expressam escolhas terminológicas; as referências revelam linhagens intelectuais; as citações recebidas sugerem circulação. Um resumo gerado por IA que ignora essas diferenças pode parecer coerente, mas empobrece a análise. A boa prática consiste em informar ao modelo que campo está sendo usado e que tipo de inferência é permitido a partir dele.

A identificação de lacunas deve sempre reconhecer o recorte do *corpus*. Um estudo que utiliza apenas Scopus não pode afirmar ausência total de produção; um estudo em inglês não pode concluir sobre literatura multilíngue; um estudo baseado em resumos não pode decidir sobre detalhes metodológicos que aparecem apenas no texto integral. A IA deve ser instruída a repetir essas condições de validade na própria saída, para impedir que a redação final converta limitações em conclusões.

A extração de entidades pode ser combinada com análise de redes. Métodos, populações e objetos extraídos dos resumos podem formar redes bipartidas com autores, países, periódicos ou anos. Essa integração permite observar, por exemplo, quais instituições se concentram em determinados métodos, quais países investigam determinadas populações e quais temas migram de um subcampo para outro. A IA ajuda a produzir as variáveis textuais; a análise cientométrica examina suas relações estruturais.

O uso de modelos de linguagem deve respeitar confidencialidade e direitos autorais. Quando o *corpus* contém textos completos protegidos, dados sensíveis, informações pessoais ou documentos não públicos, o pesquisador deve avaliar condições de uso, políticas institucionais e alternativas locais. Mesmo quando são usados apenas resumos públicos, convém preservar metadados de forma responsável. A ética da IA na cientometria inclui não apenas precisão, mas também governança dos dados utilizados.

A redação assistida por IA deve ser separada da análise assistida por IA. Um modelo pode ajudar a transformar resultados em texto, mas não deve alterar números, inventar interpretações ou ocultar incertezas. O *prompt* de redação deve fornecer dados consolidados e proibir extrapolações. Essa distinção é especialmente importante porque a qualidade estilística

da escrita pode criar impressão de robustez metodológica. A clareza textual deve servir à transparência, não mascarar fragilidades.

A comparação entre classificações humanas e classificações por IA pode gerar aprendizagem metodológica. Quando a IA erra sistematicamente, o erro pode indicar termos ambíguos, resumos mal estruturados, categorias vagas ou assimetrias disciplinares. O pesquisador deve explorar esses erros como dados sobre o próprio processo de classificação. Em vez de tratar a IA apenas como ferramenta de produtividade, pode utilizá-la como dispositivo para explicitar decisões que antes ficavam tácitas.

A evolução dos modelos de linguagem não elimina a necessidade de princípios metodológicos. Modelos futuros podem ser mais precisos, ter janelas de contexto maiores e integrar busca em bases acadêmicas. Ainda assim, a cientometria continuará exigindo definição de *corpus*, unidade de análise, indicadores, validação e interpretação. A tecnologia pode mudar; as exigências de transparência, coerência e responsabilidade permanecem.

Capítulo 15

Visualização científica assistida por IA

15.1 Introdução

A visualização científica ocupa uma posição singular na cientometria contemporânea. Ela permite que padrões de produção, impacto, colaboração, especialização temática e circulação do conhecimento sejam percebidos de modo relacional e sintético, sobretudo quando o volume de registros excede a capacidade de leitura direta. Em um estudo cientométrico, uma tabela de frequência pode informar quantos artigos foram publicados por ano, mas um gráfico de série temporal pode evidenciar regimes de crescimento, inflexões, saturações ou efeitos de eventos externos. Uma matriz de coautoria pode conter todas as relações entre autores, mas um mapa de rede pode tornar visível a existência de comunidades, pontes, periferias e centros. Um conjunto de palavras-chave pode ser apenas uma lista; um mapa temático pode apresentar sua organização cognitiva em torno de tópicos, frentes de pesquisa e zonas de transição. Desse modo, visualizar não é apenas embelezar resultados. Visualizar é formular uma hipótese de leitura sobre a estrutura dos dados.

A introdução da inteligência artificial generativa nesse processo amplia as possibilidades e também os riscos. Por um lado, modelos de linguagem podem ajudar a escolher tipos de visualização coerentes com objetivos analíticos, sugerir comparações, revisar legendas, indicar possíveis vieses de escala, propor narrativas para relatórios executivos e traduzir achados visuais para públicos não especialistas. Por outro lado, modelos generativos podem recomendar gráficos inadequados, interpretar *clusters* como se fossem categorias ontológicas estáveis, confundir associação visual com causalidade, omitir a incerteza dos dados e produzir explicações persuasivas sem base empírica suficiente. O problema central, portanto, não é se a IA deve ou não participar da visualização científica, mas sob quais condições metodológicas essa participação pode ser considerada legítima, auditável e útil.

No campo da visualização de informação, há uma tradição consolidada que ressalta a relação entre percepção, codificação gráfica e tarefa analítica. Trabalhos clássicos sobre percepção gráfica demonstraram que diferentes marcas visuais não possuem a mesma precisão interpretativa. Comparações baseadas em posição em uma escala comum tendem a ser percebidas com maior acurácia do que comparações baseadas em área, ângulo ou tonalidade. Essa constatação é decisiva para a cientometria, pois muitos erros de interpretação derivam da escolha de codificações visuais incompatíveis com a pergunta. Um gráfico de setores pode parecer simples, mas pode ser inadequado para comparar muitos países, áreas ou periódicos. Um mapa de rede pode ser impressionante, mas pode ocultar valores absolutos, efeitos de normalização e decisões de limiar. A IA deve ser instruída a considerar tais princípios, e não apenas a sugerir o gráfico mais popular ou visualmente atraente (CLEVELAND; MCGILL, 1984, 1985; SHNEIDERMAN, 1996; HEER; BOSTOCK; OGIEVETSKY, 2010).

A visualização científica em cientometria também é marcada por uma tensão entre exploração e comunicação. Na fase exploratória, gráficos e mapas funcionam como instrumentos de descoberta: permitem localizar *outliers*, detectar comunidades, perceber termos emergentes, comparar bases e identificar inconsistências nos dados. Nessa fase, múltiplas visualizações provisórias são úteis, inclusive aquelas que serão descartadas posteriormente. Na fase comunicacional, entretanto, a visualização deve ser mais seletiva e robusta: precisa sustentar uma afirmação, explicitar o recorte de dados, ser legível para o

público-alvo e evitar ambiguidade. A IA pode contribuir em ambas as fases, mas os *prompts* devem distinguir claramente se a tarefa é explorar, diagnosticar, validar, explicar ou persuadir. A ausência dessa distinção tende a gerar visualizações excessivamente genéricas (BÖRNER; CHEN; BOYACK, 2003; KEIM *et al.*, 2008; SEGEL; HEER, 2010).

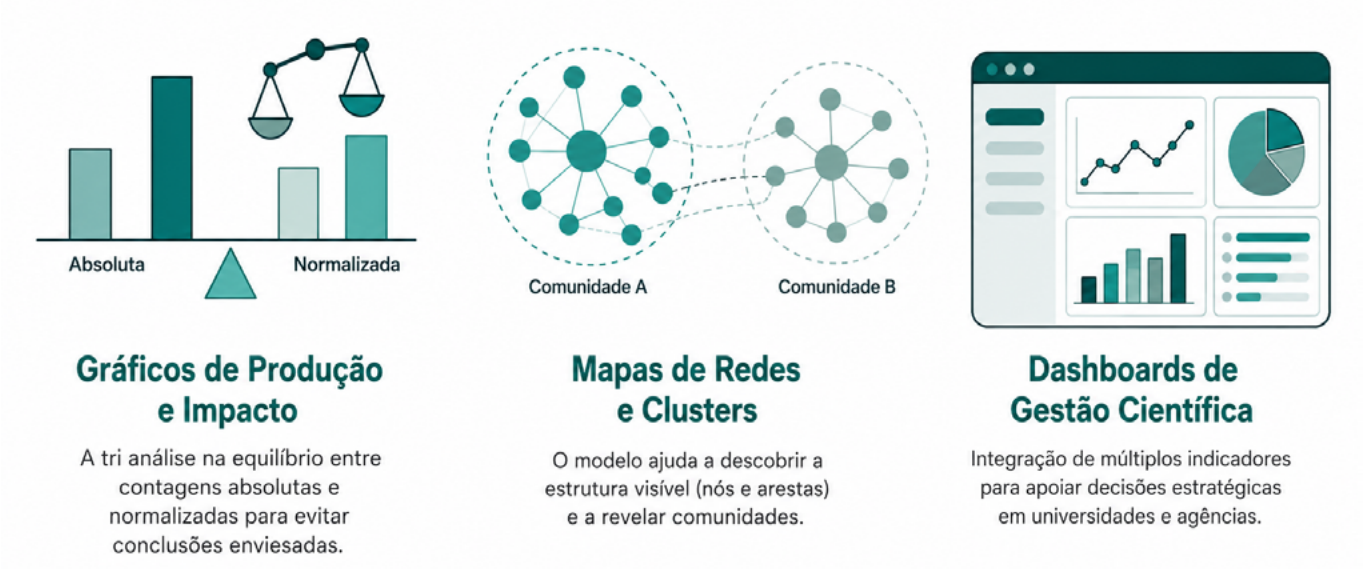
Na trilha deste livro, a visualização aparece após a formulação da pergunta, a busca, a coleta, o pré-processamento, a seleção de indicadores, a análise relacional, a construção de mapas temáticos e a detecção de tendências. Essa posição não é casual. Uma visualização confiável depende de decisões anteriores: estratégia de busca, cobertura da base, tratamento de duplicatas, padronização de nomes, normalização de citações, escolha de limiares, critérios de agrupamento, escalas de tempo e unidades de análise. Quando essas etapas não são documentadas, o gráfico final pode transmitir uma aparência de precisão que o processo não sustenta. A IA, ao operar sobre descrições parciais, tende a preencher lacunas com suposições. Por isso, um bom *prompt* de visualização deve fornecer contexto metodológico suficiente para que a recomendação visual seja situada (VAN ECK; WALTMAN, 2010; CHEN, 2006; COBO *et al.*, 2011; ARIA; CUCCURULLO, 2017).

A visualização assistida por IA deve ser compreendida como uma prática de mediação. Ela media o relacionamento entre dados bibliográficos e interpretação; entre métricas e decisão; entre pesquisadores e gestores; entre especialistas e públicos ampliados; entre o rigor metodológico e a necessidade de comunicação clara. Essa mediação exige responsabilidade semiótica, pois cada escolha visual enfatiza alguns aspectos e silencia outros. Em um gráfico de produção científica, escolher escala linear ou logarítmica altera a percepção do crescimento. Em um mapa geográfico, representar produção absoluta favorece países grandes, enquanto normalizar por população, investimento ou número de pesquisadores produz outra leitura. Em um diagrama de Sankey, a espessura dos fluxos sugere continuidade e transferência, mas nem toda mudança temática implica fluxo causal. O capítulo propõe critérios para lidar com essas decisões.

Um princípio transversal deste capítulo é que a boa visualização cientométrica deve ser orientada por pergunta. Não se começa escolhendo entre gráfico de barras, rede, mapa temático ou *dashboard*; começa-se identificando que tipo de relação precisa ser comunicada. Se a pergunta é sobre crescimento, uma série temporal pode ser adequada. Se é sobre distribuição de impacto por periódicos, um gráfico de barras ordenado ou *boxplot* pode ser preferível. Se é sobre comunidades de colaboração, uma rede pode ser útil. Se é sobre transição entre temas ao longo do tempo, um Sankey ou diagrama aluvial pode ser pertinente. Se é sobre governança institucional, um *dashboard* pode integrar diferentes indicadores, desde que suas métricas sejam explicadas e não convertidas em ranking simplista (SHNEIDERMAN, 1996; HEER; BOSTOCK; OGIEVETSKY, 2010).

Outro princípio é que a visualização deve preservar a distinção entre dado, transformação e interpretação. O dado é o registro bibliográfico coletado; a transformação é o cálculo, filtro, normalização ou agrupamento aplicado; a interpretação é a leitura que o pesquisador propõe. Em muitos relatórios, essas camadas se confundem: a rede parece revelar uma realidade objetiva, quando na verdade resulta de uma cadeia de decisões. A IA pode ajudar a explicitar essa cadeia se for provocada por *prompts* adequados. Por exemplo, antes de recomendar um mapa, pode ser solicitada a enumeração dos pressupostos que tornam o mapa defensável. Antes de escrever uma legenda, pode ser solicitado que o modelo diferencie o que a figura mostra diretamente, o que sugere indiretamente e o que não permite afirmar.

Figura 15.1 - Tipos de visualização



Fonte: Imagem gerada por IA, ChatGPT (2026).

A visualização científica assistida por IA também se vincula à literacia visual. Muitos usuários de relatórios cientométricos não possuem formação em análise de redes, estatística ou bibliometria. Gestores de pesquisa, coordenadores de programas, avaliadores externos e equipes de planejamento institucional podem interpretar gráficos de forma intuitiva, atribuindo significado excessivo a cores, proximidades, tamanhos de nós ou posições espaciais. Assim, parte da função da visualização é educar o leitor. Legendas, notas metodológicas, escalas, unidades, critérios de seleção e avisos de limitação não são acessórios: são elementos de governança interpretativa. A IA pode ajudar a produzir explicações em camadas, com versões técnica, executiva e pedagógica, desde que alimentada com os dados e restrições corretos.

Por fim, o capítulo defende que a visualização científica não deve ser avaliada apenas pela beleza ou pela inovação formal. A pergunta decisiva é se a figura permite uma interpretação mais correta, transparente e útil do fenômeno estudado. Uma visualização simples, como um gráfico de barras bem ordenado e anotado, pode ser mais científica do que uma rede densa e colorida que não permite inferir nada com segurança. Um *dashboard* pode ser útil para monitoramento contínuo, mas pode ser perigoso se transformar indicadores parciais em julgamento automático de qualidade. A IA pode acelerar a produção de alternativas visuais, mas a decisão final deve permanecer ancorada em critérios científicos, éticos e comunicacionais. Essa é a perspectiva adotada nas seções seguintes

15.2 Gráficos de produção e impacto

Os gráficos de produção e impacto constituem a porta de entrada visual para muitos estudos cientométricos. Eles apresentam a evolução do número de publicações, a distribuição de documentos por autores, instituições, países, periódicos, áreas do conhecimento, tipos documentais e períodos. Também podem representar citações, médias de citação, citações normalizadas, percentis de impacto, *h-index*, crescimento anual, concentração editorial e relações entre produtividade e visibilidade. Embora pareçam simples, esses gráficos exigem decisões metodológicas cuidadosas. A escolha entre contagens absolutas e medidas normalizadas, entre escala anual e acumulada, entre médias e medianas, entre citações totais e citações por documento, altera profundamente a leitura. A IA pode apoiar a escolha, mas

deve ser instruída a considerar a natureza da variável, a distribuição dos dados, o tamanho da amostra, a assimetria das citações e o público que receberá o resultado (HICKS *et al.*, 2015; CLEVELAND; MCGILL, 1984).

Em estudos de produção científica, gráficos de linha são adequados para representar séries temporais quando o objetivo é evidenciar crescimento, estabilidade, declínio ou inflexões. Todavia, o gráfico de linha pressupõe continuidade temporal e pode sugerir transições suaves entre anos que, na prática, são agregações discretas. Para séries muito curtas ou com poucos registros, barras anuais podem comunicar melhor a variação. Para comparar múltiplas áreas ou países, linhas sobrepostas podem gerar poluição visual; nesse caso, pequenos múltiplos ou gráficos facetados podem ser mais legíveis. Um *prompt* de apoio à escolha visual deve pedir que a IA avalie o número de categorias, o período, a magnitude das diferenças e o risco de sobrecarga cognitiva. O modelo não deve apenas indicar um tipo de gráfico, mas justificar a escolha em termos de tarefa analítica.

A análise de impacto apresenta desafios adicionais porque as distribuições de citação são altamente assimétricas. Poucos documentos concentram muitas citações, enquanto a maioria recebe poucas ou nenhuma. Por isso, médias simples podem ser enganosas, especialmente quando comparadas entre áreas com padrões de citação distintos. Visualizações baseadas em mediana, quartis, *boxplots*, histogramas ou distribuições acumuladas podem revelar melhor a estrutura dos dados. A IA pode ser solicitada a propor formas de representar assimetria e dispersão, evitando que um único número sintetize indevidamente o desempenho. Em vez de perguntar apenas 'qual gráfico usar para citações?', o pesquisador deve informar se deseja comparar grupos, mostrar desigualdade interna, identificar *outliers* ou comunicar desempenho agregado (HICKS *et al.*, 2015).

Gráficos de barras ordenadas são frequentemente úteis para rankings de autores, periódicos, instituições ou países. Contudo, todo ranking visual carrega uma dimensão normativa. Ao ordenar entidades por número de publicações ou citações, a figura pode sugerir superioridade, quando talvez apenas mostre volume, cobertura de base ou vantagem disciplinar. Recomenda-se que gráficos de ranking sejam acompanhados de notas de escopo e, quando possível, de indicadores complementares. Um gráfico de barras de produtividade institucional pode ser enriquecido por uma segunda dimensão que informe colaboração internacional, impacto normalizado ou proporção de documentos em acesso aberto. A IA pode auxiliar na redação de legendas que evitem linguagem avaliativa indevida, substituindo expressões como 'melhor desempenho' por 'maior número de registros no *corpus* analisado'.

Um erro comum é representar produção e impacto no mesmo eixo ou na mesma escala sem explicitar diferenças conceituais. Produção mede volume; impacto citacional mede reconhecimento ou uso posterior em publicações indexadas; atenção social mede circulação em ambientes digitais; colaboração mede relações de autoria. Misturar essas dimensões em visualizações compostas pode ser útil para análise estratégica, mas requer clareza. Gráficos de dispersão podem representar produção no eixo horizontal e impacto no eixo vertical, permitindo identificar entidades com alta produção e alto impacto, baixa produção e alto impacto, ou alta produção e baixo impacto. Entretanto, tal visualização não deve ser interpretada como medida completa de qualidade. A IA deve ser instruída a propor quadrantes analíticos com cautela e a indicar fatores de confusão.

Em gráficos de dispersão, a escala dos eixos, a presença de rótulos e a seleção de pontos destacados são decisões cruciais. Se muitas instituições ou autores forem representados, os rótulos podem se sobrepor e tornar o gráfico ilegível. Estratégias como destacar apenas os dez maiores valores, usar transparência, criar versões interativas ou separar grupos por facetas podem melhorar a leitura. A IA pode sugerir essas alternativas, mas deve receber informações

sobre o tamanho do conjunto e o objetivo da comunicação. Para relatórios executivos, pode ser mais apropriado apresentar uma versão simplificada, com pontos destacados e nota metodológica. Para anexos técnicos, pode-se incluir versão completa, tabela suplementar e critérios de exclusão de rótulos.

Gráficos de crescimento acumulado podem ser úteis para demonstrar a expansão de uma literatura, mas também podem gerar uma impressão inevitável de progresso, pois curvas acumuladas tendem naturalmente a crescer. Quando o objetivo é detectar mudanças de intensidade, taxas anuais, médias móveis ou gráficos de variação percentual podem ser mais informativos. A IA pode ser solicitada a comparar a adequação entre gráfico anual, acumulado e taxa de crescimento, explicitando o que cada um permite e não permite concluir. Essa distinção é especialmente importante em áreas emergentes, nas quais pequenas variações absolutas podem produzir grandes variações percentuais, gerando aparência de explosão temática sem massa crítica suficiente.

A visualização de impacto exige atenção ao tempo de maturação das citações. Documentos recentes tiveram menos tempo para acumular citações do que documentos antigos, o que pode distorcer comparações. Uma série temporal de citações por ano de publicação deve diferenciar impacto acumulado, citações por janela fixa e citações recebidas em cada ano. Gráficos que não controlam a idade dos documentos podem favorecer publicações mais antigas. A IA, quando usada para interpretar esses gráficos, deve ser orientada a verificar o período de observação e a perguntar se há normalização temporal. Sem esse controle, a legenda pode induzir a leituras injustas sobre áreas, autores ou instituições.

Em relatórios institucionais, gráficos de produção e impacto frequentemente são usados para apoiar decisões de planejamento, avaliação ou alocação de recursos. Essa função aumenta a responsabilidade do desenho visual. Um gráfico com cores fortes, títulos avaliativos e ausência de intervalo temporal pode reforçar interpretações políticas não sustentadas pelos dados. Recomenda-se formular títulos descritivos, como 'Distribuição anual de publicações indexadas no *corpus*, 2014–2024', em vez de títulos conclusivos, como 'crescimento da excelência institucional'. A IA pode ser utilizada para revisar títulos e legendas, identificando termos normativos, causalistas ou excessivamente conclusivos.

A escolha de cores em gráficos de produção e impacto deve seguir critérios semânticos e perceptivos. Paletas sequenciais são adequadas para intensidades; paletas divergentes são adequadas para desvios em torno de um ponto de referência; paletas categóricas são úteis para grupos discretos. O uso arbitrário de cores pode criar hierarquias inexistentes. Além disso, visualizações devem considerar acessibilidade, incluindo legibilidade para pessoas com daltonismo e contraste suficiente. Embora modelos de IA possam sugerir paletas, o pesquisador deve verificar compatibilidade com o meio de publicação, impressão em preto e branco, normas institucionais e convenções disciplinares. Em cientometria, cor nunca deve substituir legenda clara.

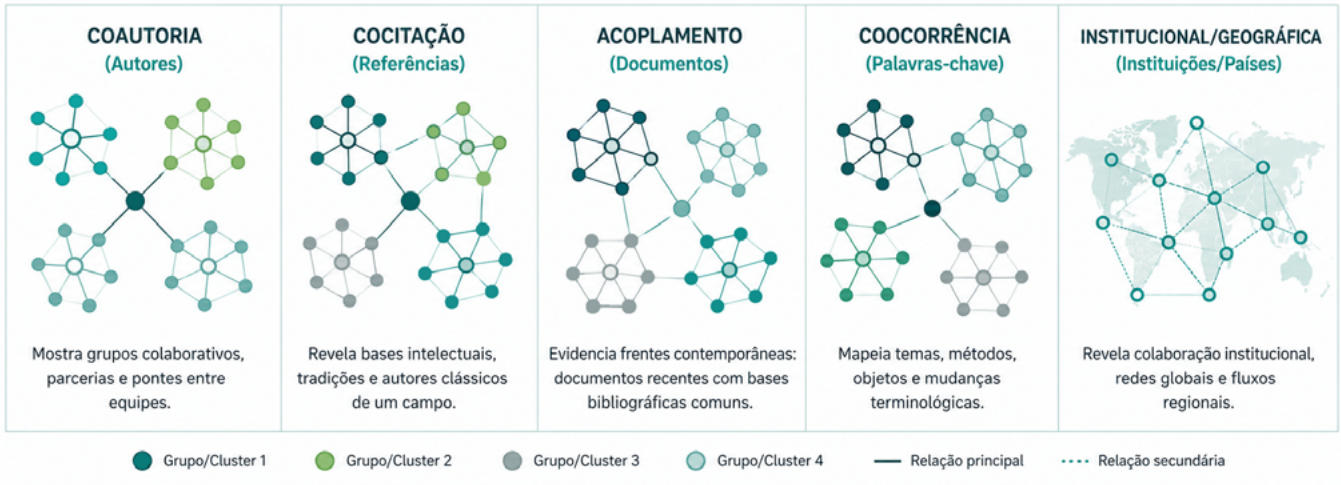
Outro aspecto relevante é o uso de anotações. Gráficos podem indicar mudanças de política científica, criação de programas, início de editais, expansão de bases de dados ou eventos externos que ajudam a contextualizar variações. Contudo, a anotação deve ser tratada como hipótese contextual, não como prova causal. Se a produção aumenta após determinado edital, o gráfico pode sugerir associação temporal, mas não demonstra causalidade sem desenho analítico adicional. A IA pode apoiar a redação de anotações prudentes, utilizando expressões como 'coincide temporalmente com', 'pode estar associado a' ou 'requer análise adicional'. Essa cautela é essencial para evitar narrativas *post hoc*.

Em síntese, gráficos de produção e impacto são fundamentais porque oferecem uma visão inicial da magnitude, distribuição e evolução do *corpus*. Contudo, sua aparente simplicidade não elimina a necessidade de rigor. Uma visualização responsável deve informar unidade de análise, fonte de dados, período, critérios de inclusão, tratamento de duplicatas, normalização, escala e limitação interpretativa. A IA pode atuar como consultora visual, revisora de legibilidade e redatora de legendas, mas não deve ser autorizada a transformar números em juízos de valor. O papel do pesquisador é garantir que a figura responda à pergunta formulada e que a narrativa permaneça compatível com a evidência apresentada.

15.3 Mapas de rede

Mapas de rede são uma das formas mais reconhecíveis de visualização cientométrica. Eles representam relações entre entidades, tais como autores, documentos, referências, periódicos, palavras-chave, instituições ou países. Nós indicam entidades; arestas indicam relações; tamanhos, cores, espessuras e posições codificam atributos derivados de métricas e algoritmos. Essa capacidade de representar estrutura relacional torna os mapas de rede especialmente úteis para investigar colaboração científica, cocitação, acoplamento bibliográfico, coocorrência de termos e conexões institucionais. Contudo, sua força comunicacional também é fonte de risco: mapas de rede podem parecer autoexplicativos, quando na verdade dependem de decisões complexas de construção, normalização, limiarização, layout e agrupamento (SMALL, 1973; KESSLER, 1963; VAN ECK; WALTMAN, 2010; WALTMAN; VAN ECK; NOYONS, 2010).

Figura 15.2 - Exemplos de mapas de rede



Fonte: Imagem gerada por IA, ChatGPT (2026).

A primeira decisão metodológica em um mapa de rede é a definição da relação. Em redes de coautoria, a aresta indica que dois autores assinaram juntos um documento. Em redes de cocitação, indica que duas referências foram citadas conjuntamente por documentos posteriores. Em redes de acoplamento bibliográfico, indica que dois documentos compartilham referências. Em redes de coocorrência, indica que dois termos aparecem no mesmo registro. Essas relações possuem significados distintos. A proximidade entre dois nós em um mapa não tem o mesmo sentido em todos os casos. A IA deve ser instruída a identificar o tipo de rede antes de interpretar a visualização. Uma legenda que não explicita o significado das arestas compromete a validade comunicacional do mapa.

O layout de rede é frequentemente interpretado como se fosse geografia, mas a posição espacial dos nós decorre de algoritmos de disposição. Métodos de força, VOS e outros algoritmos posicionam nós com base em relações e parâmetros, buscando reduzir sobreposição e evidenciar agrupamentos. A distância visual pode sugerir proximidade relacional, mas não deve ser lida como distância métrica absoluta sem conhecer o algoritmo. Em mapas bibliométricos, pequenas mudanças de parametrização podem alterar a aparência, embora a estrutura subjacente permaneça semelhante. A IA pode ajudar a explicar essa limitação em notas metodológicas, destacando que a posição espacial é representação algorítmica e não localização física ou hierarquia científica.

A normalização é outro ponto crítico. Redes bibliométricas podem ser construídas a partir de contagens brutas ou de medidas normalizadas, como força de associação. A escolha afeta o peso das relações e a formação de *clusters*. Contagens brutas favorecem entidades muito produtivas; normalizações procuram ajustar o efeito do tamanho, mas introduzem seus próprios pressupostos. Ao pedir à IA que interprete um mapa de rede, o pesquisador deve informar a medida utilizada, o limiar mínimo de ocorrência, o método de *clustering* e o significado do tamanho dos nós. Sem essas informações, o modelo pode produzir uma interpretação genérica e potencialmente enganosa.

As cores em mapas de rede costumam representar comunidades ou *clusters* detectados por algoritmos. No entanto, *clusters* não são automaticamente escolas teóricas, temas consolidados ou grupos sociais reais. Eles indicam agrupamentos relacionais segundo um conjunto de dados e parâmetros. A IA pode ser útil para propor rótulos para *clusters*, especialmente quando recebe os termos, documentos ou referências mais representativos de cada grupo. Todavia, a rotulagem precisa ser validada por especialistas. Um *cluster* de palavras-chave pode reunir termos metodológicos e substantivos; um *cluster* de cocitação pode refletir uma tradição teórica, mas também pode representar uma base metodológica compartilhada. A visualização mostra estrutura; a interpretação exige conhecimento de domínio.

Mapas de coautoria são particularmente úteis para visualizar colaboração, mas não devem ser confundidos com redes sociais completas. A coautoria é uma forma documentada de colaboração formal, não o conjunto total de interações científicas. Projetos, orientações, participação em bancas, intercâmbios informais e contribuições técnicas podem não aparecer. Além disso, áreas diferem em padrões de autoria. Em algumas disciplinas, grandes equipes são norma; em outras, autoria individual ou pequenos grupos predominam. Assim, comparar mapas de coautoria entre áreas requer cautela. A IA pode revisar uma interpretação para identificar extrapolações, como inferir liderança científica apenas pelo grau de conectividade ou afirmar isolamento sem considerar práticas disciplinares.

Mapas de cocitação e acoplamento bibliográfico oferecem perspectivas complementares sobre a estrutura intelectual da ciência. A cocitação tende a revelar como a literatura posterior associa referências anteriores, aproximando obras que são citadas juntas. O acoplamento bibliográfico aproxima documentos que compartilham referências, sendo sensível a frentes de pesquisa contemporâneas. Visualmente, ambos podem produzir *clusters*, mas os fundamentos temporais são diferentes. A IA deve ser instruída a não tratar cocitação e acoplamento como equivalentes. Em uma análise madura, a combinação de mapas pode mostrar tanto a base intelectual quanto a frente ativa de investigação, desde que as diferenças sejam explicadas no relatório (SMALL, 1973; KESSLER, 1963; BOYACK; KLAVANS, 2010).

Redes de coocorrência de palavras-chave permitem visualizar associações conceituais, mas dependem fortemente do pré-processamento. Sinônimos, siglas, variações linguísticas, plurais, erros de digitação e termos genéricos podem fragmentar ou inflar *clusters*. Uma rede de termos construída sem normalização pode parecer rica, mas refletir ruído terminológico.

Nesse ponto, a IA pode apoiar a revisão semântica, sugerindo agrupamentos e termos preferenciais, mas a decisão final deve ser registrada. Ao interpretar a rede, é recomendável indicar se foram usadas palavras-chave dos autores, descritores indexados, termos extraídos de títulos/resumos ou combinação dessas fontes. Cada opção afeta a visualização (CALLON; COURTIAL; LAVILLE, 1991).

A densidade da rede é um problema recorrente. Mapas com centenas ou milhares de nós podem impressionar, mas não comunicar. Em relatórios, é preferível apresentar uma visualização filtrada, justificada por critérios transparentes, e disponibilizar a rede completa como material suplementar. Filtros podem ser definidos por frequência mínima, peso mínimo de aresta, número de citações, centralidade ou pertinência temática. A IA pode sugerir estratégias de redução visual, mas deve ser orientada a preservar a pergunta analítica. Reduzir a rede para torná-la bonita não é suficiente; a redução deve aumentar a inteligibilidade sem alterar indevidamente a estrutura relevante.

Mapas geográficos e institucionais podem ser representados como redes ou como mapas territoriais. Redes destacam relações entre instituições ou países; mapas territoriais destacam distribuição espacial. O erro comum é interpretar localização geográfica como intensidade relacional. Um país pode ter alta produção absoluta e baixa colaboração internacional; outro pode ter menor produção e alta centralidade em redes transnacionais. A IA pode ajudar a propor visualizações combinadas, como mapa coroplético para produção e rede sobreposta para colaboração, mas deve advertir sobre sobrecarga visual. Em muitos casos, duas figuras simples são mais eficazes do que uma figura composta excessivamente densa.

A interpretação de mapas de rede assistida por IA deve seguir uma estrutura prudente. Primeiro, descrever a estrutura visível: número de *clusters*, concentração, dispersão, nós centrais, pontes e periferias. Segundo, apresentar possíveis explicações, sempre condicionadas ao tipo de rede e aos dados. Terceiro, formular hipóteses interpretativas. Quarto, indicar evidências adicionais necessárias. Quinto, explicitar riscos de interpretação indevida. Esse encadeamento impede que a IA converta aparência visual em causalidade. O *prompt* central deste capítulo adota exatamente essa lógica, exigindo separação entre descrição estrutural, explicação possível, hipótese e limitação.

Em síntese, mapas de rede são instrumentos poderosos para revelar a arquitetura relacional da ciência, mas exigem alfabetização metodológica. Sua interpretação não pode depender apenas da estética do mapa. Cada nó, aresta, cor, distância e *cluster* resulta de uma decisão. A IA pode aumentar a capacidade de análise ao auxiliar na descrição, rotulagem, comparação e comunicação dos mapas; contudo, deve operar sob restrições explícitas. A recomendação mais importante é simples: nunca pedir à IA que 'explique o mapa' sem fornecer o tipo de rede, a fonte de dados, o método de construção, os parâmetros e a finalidade da análise.

15.4 Diagramas de Sankey, mapas temáticos e linhas do tempo

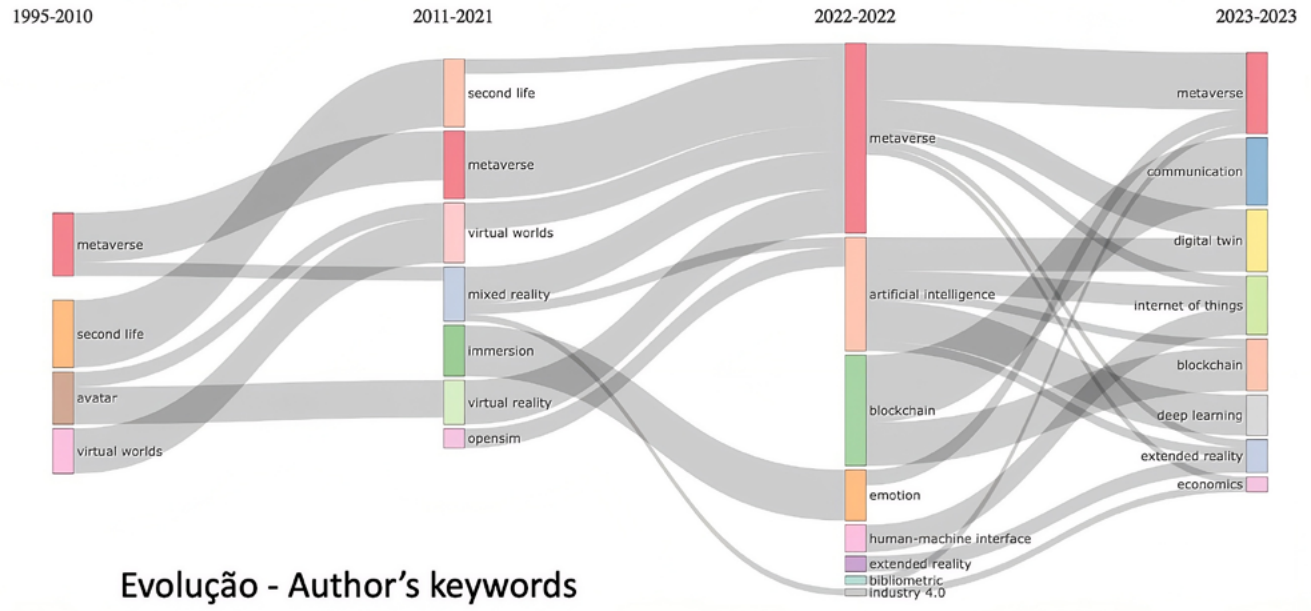
Diagramas de Sankey, mapas temáticos e linhas do tempo são recursos particularmente valiosos quando a cientometria busca comunicar movimento, evolução, transição e estrutura cognitiva. Enquanto gráficos de produção enfatizam quantidades e mapas de rede enfatizam relações, essas visualizações procuram mostrar como temas se deslocam, como categorias se conectam e como trajetórias se desenvolvem. Em estudos assistidos por IA, esses formatos podem ser recomendados para representar mudanças terminológicas, evolução de *clusters*, fluxos de colaboração, migração de temas entre períodos, composição de áreas e conexões entre tópicos. Entretanto, sua interpretação exige cuidado porque a ideia visual de fluxo

pode sugerir continuidade causal mesmo quando os dados apenas indicam associação ou reclassificação (SCHMIDT, 2008a, 2008b; COBO *et al.*, 2012).

Diagramas de Sankey representam fluxos entre categorias por meio de faixas cuja largura corresponde a uma quantidade. Em cientometria, podem mostrar a transição de temas entre períodos, a distribuição de publicações de instituições para áreas, a passagem de palavras-chave para *clusters* ou a composição de colaborações entre países. São úteis quando há uma relação de origem-destino ou quando se deseja mostrar redistribuição entre categorias. Contudo, nem toda relação categorial deve ser transformada em fluxo. Se os temas foram reagrupados por algoritmos diferentes em períodos distintos, o Sankey pode sugerir que um tema 'evoluiu' para outro, quando talvez apenas tenha sido renomeado ou reclassificado. A IA deve ser instruída a verificar se há correspondência real entre categorias antes de recomendar esse tipo de diagrama.

A força de um Sankey está em tornar visível a proporção de fluxos, mas sua fragilidade está na sobrecarga visual. Muitos nós e fluxos produzem emaranhados difíceis de interpretar. Para relatórios cientométricos, recomenda-se limitar o número de categorias, agrupar fluxos residuais em 'outros', apresentar critérios de agregação e evitar cores excessivas. A IA pode sugerir quais fluxos destacar, mas deve receber critérios substantivos, como relevância para a pergunta, volume mínimo ou interesse estratégico. Em contextos de gestão científica, o Sankey pode ser atraente para mostrar trajetórias de financiamento, áreas de aplicação ou redes de cooperação; porém, a atratividade visual não substitui a transparência do modelo de classificação.

Figura 15.3 - Diagramas de Sankey

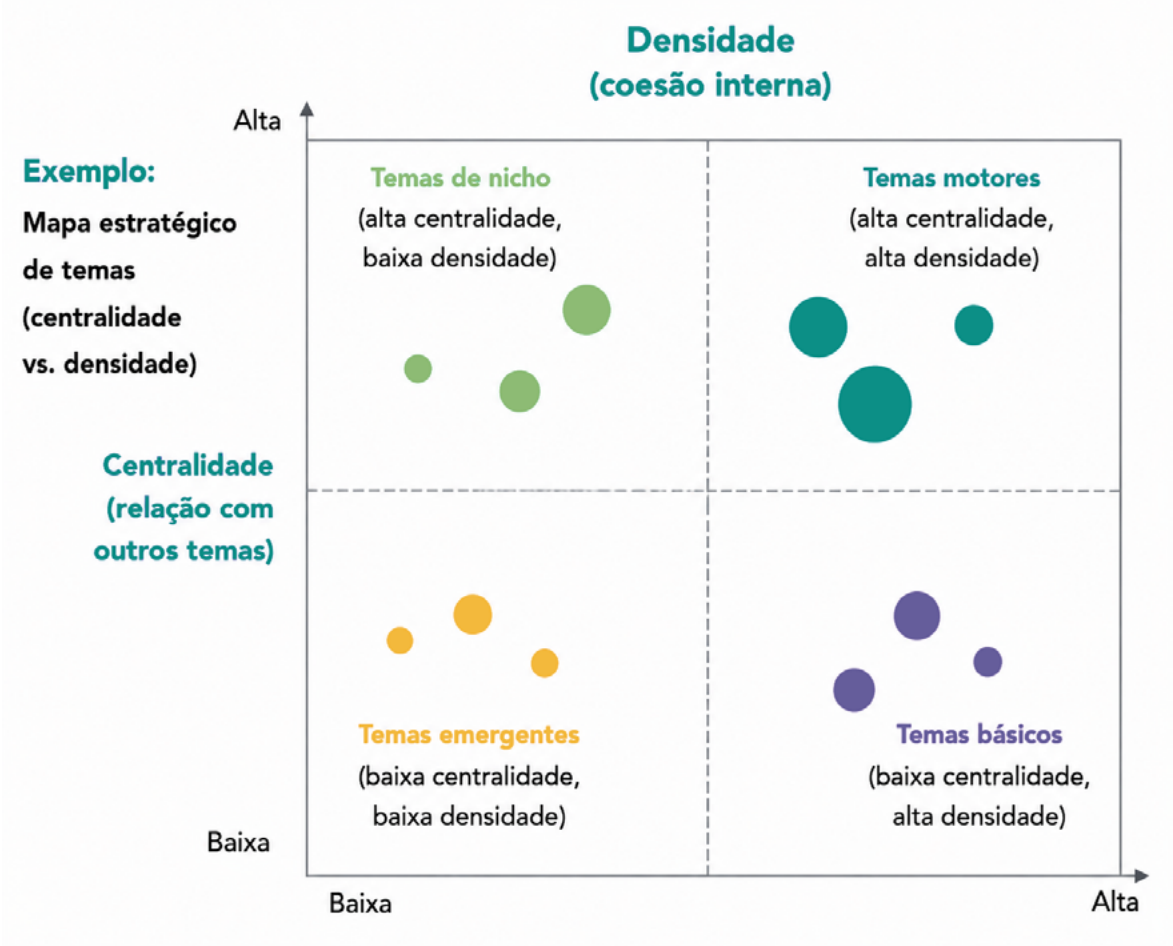


Fonte: Elaborado pelo autor (2026).

Mapas temáticos, por sua vez, representam tópicos ou *clusters* em espaços de associação. Podem ser construídos a partir de coocorrência de palavras-chave, análise fatorial, correspondência múltipla, *embeddings* semânticos, modelos de tópicos ou algoritmos de *cluster*ização. Em cientometria, são usados para identificar temas motores, temas básicos, temas emergentes e temas declinantes; para situar frentes de pesquisa; e para comunicar a estrutura intelectual de uma área. A IA pode auxiliar na nomeação dos *clusters* e na construção de narrativas, mas a figura deve informar como os temas foram

extraídos. Um mapa temático produzido por copalavras não tem o mesmo significado de um mapa produzido por *embeddings* de resumos (CALLON; COURTIAL; LAVILLE, 1991; COBO *et al.*, 2011, 2012; ARIA; CUCCURULLO, 2017).

Figura 15.4 - Mapa temático



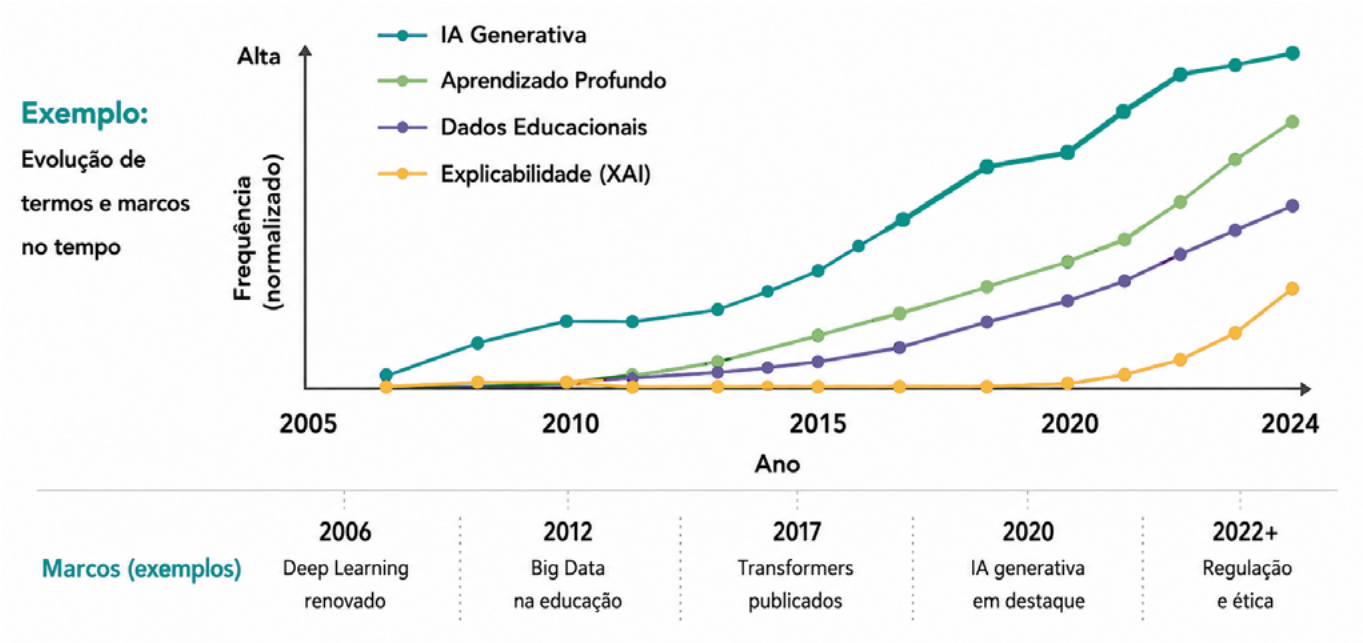
Fonte: Imagem gerada por IA, ChatGPT (2026).

A interpretação de mapas temáticos exige distinguir entre rótulo, conteúdo e posição. O rótulo é uma síntese proposta para um *cluster*; o conteúdo são os termos, documentos ou autores que compõem o grupo; a posição é resultado de algoritmo e métricas. A IA pode propor rótulos sugestivos, mas deve ser obrigada a justificar cada rótulo com evidências textuais: termos mais frequentes, títulos representativos, resumos típicos ou referências centrais. Um *prompt* robusto deve pedir rótulos alternativos, grau de confiança e casos ambíguos. Essa prática evita que um nome elegante oculte heterogeneidade interna. Em mapas temáticos, a nomeação é ato interpretativo, não operação mecânica.

Linhas do tempo são indicadas quando o objetivo é mostrar sequência histórica, períodos de consolidação, emergência de termos, picos de produção, eventos de política científica ou mudanças de foco. Podem representar publicações por ano, citações acumuladas, surgimento de palavras-chave, evolução de *clusters*, marcos institucionais ou publicações seminais citadas. Em estudos cientométricos, uma boa linha do tempo deve combinar escala temporal, eventos anotados e dados quantitativos. A IA pode ajudar a transformar uma série de achados em narrativa temporal, mas deve ser instruída a diferenciar eventos observados no *corpus* de

eventos externos incorporados pelo pesquisador. Essa distinção é importante para não atribuir causalidade sem evidência.

Figura 15.5 - Linhas do tempo



Fonte: Imagem gerada por IA, ChatGPT (2026).

A escolha entre Sankey, mapa temático e linha do tempo depende da pergunta. Se a pergunta é 'como se distribuem os documentos entre áreas e instituições?', um Sankey pode ser útil. Se é 'quais são os principais temas e como se organizam?', um mapa temático é mais adequado. Se é 'quando determinados temas emergem e ganham visibilidade?', uma linha do tempo pode ser superior. A IA deve ser solicitada a comparar alternativas e indicar vantagens, limitações e riscos de cada opção. A pergunta 'qual visualização é melhor?' é incompleta; a pergunta adequada é 'qual visualização responde de modo mais transparente a esta pergunta, para este público, com estes dados?'.

Uma dificuldade recorrente em visualizações de evolução temática é a estabilidade dos rótulos. Termos mudam ao longo do tempo, conceitos se deslocam, siglas surgem, áreas importam vocabulário de outros campos e tecnologias alteram nomes sem alterar necessariamente problemas de pesquisa. A IA pode identificar variantes terminológicas e sugerir equivalências, mas deve ser usada com controle humano. Uma linha do tempo de palavras-chave pode mostrar declínio de um termo antigo e crescimento de um termo novo; isso pode indicar mudança substantiva, mas também substituição lexical. O *prompt* deve pedir explicitamente a distinção entre mudança temática, mudança terminológica e artefato de indexação.

Em mapas temáticos, a visualização de centralidade e densidade é comum para classificar temas como motores, básicos, emergentes ou declinantes. Entretanto, tais categorias dependem do método e dos parâmetros. Um tema 'emergente' não é necessariamente promissor; pode ser apenas recente, pequeno ou mal conectado. Um tema 'declinante' não é necessariamente irrelevante; pode estar consolidado em outra terminologia ou migrado para campos aplicados. A IA pode ser convidada a escrever interpretações prudentes, evitando linguagem determinista. Em vez de afirmar 'o tema X está morrendo', o relatório pode dizer 'no

corpus e no período analisados, o tema X apresenta menor densidade e menor centralidade, o que sugere necessidade de investigação adicional sobre sua trajetória'.

Diagramas aluviais e de Sankey podem representar mudanças de classificação de documentos entre períodos. Para isso, cada documento precisa ter uma identidade rastreável e uma categoria atribuída em cada etapa. Quando a classificação é feita por IA, o risco aumenta: mudanças podem refletir instabilidade do classificador, e não transformação real da área. Portanto, a visualização de transição temática deve ser acompanhada de informação sobre consistência classificatória, amostragem de validação e critérios de atribuição. A IA pode ajudar a gerar o diagrama, mas também deve ser usada para produzir uma auditoria textual dos fluxos: quais são robustos, quais são pequenos demais e quais exigem validação especializada.

A narrativa visual de linhas do tempo deve evitar a ilusão de inevitabilidade. Quando se selecionam apenas eventos que parecem confirmar uma tendência, constrói-se uma história coerente, mas possivelmente enviesada. A IA, se solicitada apenas a 'contar a história da evolução do tema', pode produzir uma narrativa teleológica. Um *prompt* mais adequado pede que o modelo apresente interpretações concorrentes, eventos ausentes, possíveis efeitos de base de dados e incertezas. Visualizações temporais devem ser lidas como evidências parciais de trajetória, não como biografias completas de um campo científico.

Em termos de comunicação, Sankey, mapas temáticos e linhas do tempo são excelentes para relatórios que buscam explicar estrutura e dinâmica a públicos não especialistas. Eles condensam relações complexas em formas visuais memoráveis. Porém, justamente por serem memoráveis, podem fixar interpretações simplificadas. A IA pode apoiar a criação de legendas em camadas: uma legenda curta para leitura rápida, uma legenda metodológica para especialistas e uma nota de cautela para uso institucional. Essa estratégia aumenta a utilidade sem reduzir o rigor. Em livros, artigos e relatórios, a legenda deve funcionar como ponte entre visualização e argumento.

Conclui-se que esses formatos são especialmente adequados para visualizar transformação, mas exigem verificação da natureza da transformação. Fluxo visual não é sinônimo de causalidade; proximidade temática não é equivalência conceitual; sequência temporal não é explicação. A inteligência artificial pode ampliar a capacidade de construir, revisar e narrar essas visualizações, desde que o pesquisador imponha controles: definição de categorias, documentação de parâmetros, validação de rótulos, revisão de ambiguidades e explicitação de limitações. A visualização científica assistida por IA deve transformar complexidade em compreensão, não em simplificação enganosa.

15.5 Dashboards para gestão científica

Dashboards para gestão científica são ambientes visuais que integram múltiplos indicadores e permitem monitorar produção, impacto, colaboração, internacionalização, financiamento, áreas estratégicas, acesso aberto, redes institucionais e tendências temáticas. Diferentemente de uma figura isolada, o *dashboard* organiza um conjunto de visualizações em uma interface orientada à decisão. Em universidades, agências, programas de pós-graduação e centros de pesquisa, *dashboards* podem apoiar planejamento, prestação de contas, identificação de oportunidades e acompanhamento de políticas científicas. Entretanto, a criação de *dashboards* cientométricos requer cuidado redobrado, pois a combinação de muitos indicadores em um mesmo painel pode produzir aparência de controle total e objetividade, mesmo quando os dados são incompletos, desiguais ou sensíveis a decisões metodológicas (SARIKAYA *et al.*, 2019; HICKS *et al.*, 2015).

A primeira questão na construção de um *dashboard* é a finalidade. Um painel para acompanhamento interno de produção institucional não deve ser idêntico a um painel para avaliação externa, prospecção tecnológica ou comunicação pública. Cada finalidade requer seleção distinta de indicadores, níveis de agregação, frequência de atualização e linguagem visual. A IA pode ser solicitada a propor arquitetura de *dashboard* a partir de objetivos, público e decisões esperadas. Contudo, o *prompt* deve impedir que o modelo recomende excesso de métricas. Um bom *dashboard* não é aquele que mostra tudo, mas aquele que apresenta os indicadores necessários para uma decisão ou reflexão específica, com rastreabilidade suficiente para auditoria.

Em gestão científica, *dashboards* frequentemente são construídos em camadas. A primeira camada oferece indicadores sintéticos: número de publicações, citações, colaborações, áreas predominantes e evolução temporal. A segunda camada permite explorar recortes por unidade, departamento, programa, país, área ou período. A terceira camada oferece detalhes: lista de documentos, autores, fontes, critérios de coleta e dados brutos. Essa arquitetura reduz o risco de interpretação superficial, pois permite passar do agregado ao documento. A IA pode apoiar o desenho dessas camadas, sugerindo filtros, hierarquias e perguntas que cada painel deve responder. No entanto, cada camada deve preservar coerência entre unidade de análise e indicador.

Um problema ético surge quando *dashboards* são usados para comparar indivíduos. Indicadores cientométricos possuem limitações conhecidas e podem produzir incentivos perversos se empregados como mecanismos automáticos de avaliação. Visualizações que ordenam pesquisadores por citações, produção ou índice h podem ser utilizadas de modo reducionista, ignorando diferenças de carreira, área, gênero, idioma, tipo de produção e contribuição social. A IA pode ajudar a redigir notas de governança, indicando que *dashboards* devem apoiar diagnóstico institucional e não substituir avaliação qualitativa. Em contextos de pessoas, recomenda-se usar indicadores em agregados, com contextualização e proteção de dados, evitando exposição competitiva indevida.

Dashboards também precisam lidar com atualização e versionamento. Um painel dinâmico pode mudar conforme novas publicações entram nas bases, registros são corrigidos, afiliações são atualizadas ou citações acumulam. Sem registro de data de coleta, versão da base, filtros e scripts, o *dashboard* se torna difícil de auditar. A IA pode gerar checklists de documentação e sugerir campos obrigatórios de metadados do painel: fonte, data de extração, periodicidade, critérios de inclusão, identificadores usados, limitações conhecidas e responsável pela curadoria. Essa documentação deve estar visível ou facilmente acessível no próprio *dashboard*.

A escolha das visualizações internas de um *dashboard* deve respeitar a diversidade de tarefas. Séries temporais respondem a perguntas de evolução; barras ordenadas respondem a perguntas de comparação; mapas geográficos respondem a distribuição territorial; redes respondem a relações; mapas temáticos respondem a estrutura cognitiva; tabelas detalhadas respondem a auditoria. A IA pode propor um conjunto equilibrado, mas deve ser orientada a evitar redundância. Três gráficos diferentes mostrando a mesma contagem podem ocupar espaço que deveria ser usado para explicar limitações. O desenho de *dashboard* é, em parte, uma prática de edição: selecionar, hierarquizar e contextualizar.

A interatividade é uma vantagem dos *dashboards*, mas também um risco. Filtros, seletores, *zoom*, *brushing* e *drill-down* permitem exploração, mas podem gerar interpretações instáveis quando usuários combinam filtros sem compreender o efeito estatístico. Um filtro por área, por exemplo, pode reduzir drasticamente a amostra e tornar percentuais pouco confiáveis. O painel deve sinalizar tamanho da amostra, limites mínimos para exibição e mensagens de alerta quando interpretações forem frágeis. A IA pode auxiliar na formulação desses alertas,

transformando regras estatísticas em mensagens compreensíveis. Assim, o *dashboard* não apenas mostra dados; ele ensina o usuário a interpretá-los com prudência.

Dashboards para gestão científica devem diferenciar monitoramento, avaliação e planejamento. Monitorar é acompanhar sinais; avaliar é julgar valor segundo critérios; planejar é decidir ações. Um painel pode contribuir para os três processos, mas não realiza automaticamente nenhum deles. A visualização de queda de publicações em uma área pode acionar investigação, mas não prova baixa qualidade ou falta de relevância. O aumento de colaboração internacional pode ser positivo em uma estratégia e irrelevante em outra. A IA, ao gerar narrativas automáticas a partir de *dashboards*, deve ser instruída a separar constatações de recomendações. Recomendações precisam estar vinculadas a objetivos institucionais explícitos.

A acessibilidade dos *dashboards* é parte da qualidade científica. Painéis devem ser legíveis em diferentes tamanhos de tela, utilizar contraste adequado, evitar dependência exclusiva de cor, oferecer textos alternativos quando possível e permitir exportação de dados. A IA pode revisar especificações de acessibilidade e sugerir melhorias. Em contextos públicos, a acessibilidade também é política: dados sobre ciência financiada socialmente devem ser comunicados de forma compreensível e inclusiva. A visualização científica assistida por IA deve, portanto, considerar não apenas eficiência analítica, mas direito à interpretação informada.

Outro desafio é a interoperabilidade. *Dashboards* cientométricos podem integrar dados de bases comerciais, repositórios institucionais, ORCID, Crossref, Lens, OpenAlex, sistemas internos e plataformas de fomento. Cada fonte possui estrutura, cobertura, atualização e identificadores distintos. Visualizações integradas podem esconder inconsistências entre bases. A IA pode ajudar a mapear campos, propor dicionários de dados e identificar lacunas, mas não deve assumir equivalência entre fontes. Um painel robusto deve informar de onde vem cada indicador e se há sobreposição, deduplicação ou harmonização. A qualidade do *dashboard* depende da qualidade da integração.

Em termos de design, *dashboards* devem evitar excesso de elementos decorativos. Ícones, gradientes, animações e mapas 3D podem impressionar, mas frequentemente reduzem clareza. A prioridade deve ser a legibilidade, a hierarquia visual e a relação explícita entre indicador e pergunta. A IA pode ser usada para revisar o painel segundo critérios de parcimônia: cada componente responde a qual pergunta? Há elementos redundantes? A escala é apropriada? A cor tem significado? A legenda é suficiente? O usuário consegue identificar a fonte dos dados? Essas perguntas devem fazer parte do *prompt* de revisão.

Dashboards de gestão científica podem incluir narrativas automáticas, geradas por IA, que resumem mudanças recentes e alertas. Essa funcionalidade é útil, mas deve ser controlada. A narrativa automática precisa mencionar incertezas, evitar causalidade indevida, indicar base temporal e sinalizar quando mudanças são pequenas. Recomenda-se que textos gerados sejam revisados por analistas antes de circular institucionalmente. Uma frase como 'a área perdeu relevância' pode ser substituída por 'no *corpus* analisado, houve redução no número de documentos indexados no período recente, o que requer verificação de cobertura, estratégia de busca e dinâmica disciplinar'. Essa diferença é metodologicamente crucial.

Conclui-se que *dashboards* são ferramentas poderosas para gestão científica quando estruturados como instrumentos de inteligência institucional responsável. Eles permitem ver padrões, acompanhar metas, identificar oportunidades e comunicar resultados. Porém, também podem induzir simplificação, competição inadequada e decisões baseadas em métricas parciais. A IA pode apoiar seu desenho, documentação, revisão, explicação e narrativa, desde que subordinada a princípios de avaliação responsável. O *dashboard* deve ser uma interface para perguntas melhores, não uma máquina de respostas automáticas. Em cientometria, boa gestão visual é aquela que aumenta transparência e reduz arbitrariedade.

15.6 Como pedir à IA recomendações de visualização

Pedir à IA uma recomendação de visualização é uma tarefa aparentemente simples, mas metodologicamente exigente. Um *prompt* genérico, como 'qual gráfico devo usar?', raramente produz uma resposta adequada, porque o modelo não conhece a pergunta de pesquisa, a estrutura dos dados, o público, o formato de publicação, as limitações metodológicas nem o grau de detalhe esperado. A qualidade da recomendação depende da qualidade do contexto fornecido. Em cientometria, o *prompt* deve incluir, no mínimo, objetivo analítico, unidade de análise, variáveis disponíveis, tipo de dado, tamanho da amostra, período, público-alvo, mensagem principal, restrições de formato e riscos interpretativos. Sem esses elementos, a IA tende a oferecer listas de gráficos populares, em vez de decisões visualmente justificadas (SHNEIDERMAN, 1996; HEER; BOSTOCK; OGIEVETSKY, 2010).

Um bom *prompt* de recomendação visual deve obrigar a IA a raciocinar em termos de tarefa. A literatura de visualização enfatiza que diferentes gráficos servem a diferentes tarefas: comparar, ordenar, distribuir, correlacionar, localizar, mostrar mudança, revelar composição, representar rede ou comunicar fluxo. Assim, o pesquisador deve perguntar: 'qual visualização é mais adequada para comparar a produção de instituições ao longo de dez anos?', ou 'como representar a evolução de *clusters* temáticos em três períodos sem sugerir causalidade?'. Quanto mais explícita for a tarefa, maior a chance de obter uma recomendação útil. O modelo deve ser instruído a justificar a escolha e a indicar alternativas.

É recomendável pedir que a IA compare pelo menos três opções de visualização. A primeira pode ser a opção principal, a segunda uma alternativa mais simples e a terceira uma alternativa para público especializado. Por exemplo, para representar colaboração internacional, a IA pode comparar mapa geográfico, matriz país-país e rede de colaboração. Cada uma responde a aspectos diferentes: distribuição territorial, intensidade bilateral e estrutura relacional. O pesquisador, então, decide conforme o argumento do estudo. Pedir apenas uma resposta única pode ocultar *trade-offs*. Pedir comparação torna visíveis as vantagens e limitações de cada forma visual.

O *prompt* também deve pedir que a IA identifique visualizações inadequadas. Essa é uma estratégia poderosa de controle de qualidade. Em vez de solicitar apenas 'sugira gráficos', o pesquisador pode pedir: 'indique quais gráficos seriam inadequados para estes dados e explique por quê'. O modelo pode alertar contra gráfico de setores com muitas categorias, eixo duplo sem relação clara, mapa geográfico para dados não espaciais, rede densa demais ou Sankey sem fluxo real. Essa abordagem transforma a IA em revisora crítica, não apenas em geradora de opções. Em estudos cientométricos, evitar uma má visualização é tão importante quanto escolher uma boa.

As recomendações da IA devem incluir especificações práticas. Não basta dizer 'use gráfico de linha'; é necessário indicar eixo horizontal, eixo vertical, unidade, agregação, escala, tratamento de *outliers*, intervalos, rótulos, legenda, cores e notas metodológicas. Um *prompt* robusto pode solicitar uma especificação completa da figura, incluindo título descritivo, subtítulo metodológico, legenda analítica e nota de limitação. Esse nível de detalhe facilita a tradução da recomendação para ferramentas como R, Python, VOSviewer, bibliometrix, Power BI, Tableau ou planilhas. Também aumenta a reprodutibilidade, pois documenta a intenção visual.

Ao pedir recomendações de visualização, o pesquisador deve informar o público-alvo. Uma figura para artigo acadêmico pode privilegiar precisão e densidade; uma figura para relatório de gestão deve favorecer clareza e tomada de decisão; uma figura para apresentação pública deve ser ainda mais sintética. A IA pode adaptar linguagem, rótulos e nível de explicação, mas precisa saber para quem está projetando. Um mapa de rede complexo pode ser adequado

para especialistas em cientometria é inadequado para conselhos universitários. A mesma evidência pode exigir versões distintas: técnica, executiva e pedagógica. Essa adaptação não é simplificação indevida se preservar a rastreabilidade.

Outro elemento essencial é o meio de circulação. Visualizações para impressão precisam de boa legibilidade em escala reduzida, contraste em preto e branco e legendas completas. Visualizações digitais podem usar interação, filtros, *tooltips* e camadas. Slides exigem menos densidade textual e maior hierarquia visual. *Dashboards* exigem responsividade e estados de filtragem. A IA pode sugerir adaptações por meio, mas o *prompt* deve explicitar o destino. Uma figura perfeita para monitor grande pode ser ilegível em artigo impresso. Uma rede interativa pode perder significado se exportada como imagem estática.

O pesquisador deve solicitar que a IA considere incerteza, incompletude e vieses. Em cientometria, dados bibliográficos raramente são completos. Bases têm cobertura desigual, nomes podem estar ambíguos, citações são acumulativas, palavras-chave são heterogêneas e áreas possuem práticas distintas. Visualizações devem, quando possível, incorporar ou mencionar essas limitações. Um *prompt* de recomendação visual pode pedir: 'inclua formas de representar ou comunicar incerteza e limitações dos dados'. A resposta pode sugerir intervalos, notas, filtros mínimos, camadas de qualidade ou anexos metodológicos. Sem essa preocupação, a visualização pode parecer mais exata do que é.

Também é útil pedir que a IA revise uma visualização já planejada. O pesquisador pode descrever a figura: gráfico de barras com 25 países, cores diferentes para cada país, ordenado alfabeticamente, produção total de 2000 a 2024'. A IA pode identificar problemas: excesso de categorias, ordenação pouco informativa, ausência de normalização, dificuldade de leitura, necessidade de agrupar categorias menores ou ordenar por valor. Essa revisão é particularmente eficaz quando o *prompt* pede críticas específicas: legibilidade, adequação ao objetivo, risco de distorção, carga cognitiva, acessibilidade e necessidade de nota metodológica.

Para gerar narrativas visuais, o *prompt* deve proibir extrapolação. A IA deve ser orientada a escrever apenas o que a figura permite afirmar. Em um gráfico de crescimento de publicações, pode-se afirmar aumento de documentos indexados; não se pode afirmar, sem evidência adicional, aumento de qualidade, impacto social ou maturidade teórica. Em uma rede, pode-se afirmar maior conectividade; não se pode afirmar colaboração mais eficiente. Ao solicitar legenda ou interpretação, o pesquisador deve pedir separação entre 'o que a figura mostra', 'o que sugere' e 'o que não permite concluir'. Essa estrutura reduz o risco de conclusões indevidas.

Prompts de visualização devem ser documentados como parte do método. Se a IA sugeriu o tipo de gráfico, revisou a legenda ou propôs agrupamentos visuais, essas interações devem ser registradas em apêndice metodológico ou caderno de pesquisa. A documentação deve incluir versão do modelo, data, *prompt*, dados fornecidos, saída produzida, decisões humanas e alterações posteriores. Essa prática fortalece a reprodutibilidade e evita que a IA opere como caixa-preta invisível. Em trabalhos acadêmicos, não é necessário publicar todas as interações, mas é necessário manter trilha auditável, especialmente quando a IA influencia decisões analíticas.

Um princípio final é pedir à IA para explicar suas recomendações em termos de princípios, não de preferência. O modelo deve responder por que uma visualização é adequada: porque permite comparação precisa, reduz sobrecarga, preserva ordem temporal, explicita distribuição, mostra relação ou comunica fluxo. Deve também indicar quando a recomendação depende de condições. Por exemplo, um Sankey é adequado se as categorias forem comparáveis entre períodos; uma rede é adequada se a relação for substantiva e a densidade controlada; um *dashboard* é adequado se houver atualização e governança de dados. Recomendação condicional é mais útil do que resposta categórica.

15.7 Como evitar visualizações bonitas, mas metodologicamente frágeis

A expressão 'visualizações bonitas, mas metodologicamente frágeis' designa figuras que impressionam pela estética, complexidade ou aparência tecnológica, mas não sustentam interpretações rigorosas. Em cientometria, esse risco é elevado porque mapas de rede, *dashboards* coloridos, diagramas de Sankey e mapas temáticos podem produzir forte impacto visual. A figura pode circular em apresentações, relatórios e mídias institucionais, adquirindo autoridade simbólica. No entanto, se a visualização deriva de dados incompletos, filtros arbitrários, normalização inadequada, rótulos não validados ou interpretação causal indevida, sua beleza torna-se problema, não virtude. A visualização científica deve ser julgada pela coerência entre pergunta, dado, forma visual e interpretação (CLEVELAND; MCGILL, 1984, 1985; HICKS *et al.*, 2015).

O primeiro critério para evitar fragilidade é verificar se a visualização responde a uma pergunta explícita. Figuras produzidas apenas porque a ferramenta permite gerá-las tendem a ser decorativas. Um mapa de rede sem pergunta pode mostrar conexões, mas não necessariamente produzir conhecimento. Um *dashboard* sem decisão associada pode acumular métricas sem orientar análise. Um gráfico de impacto sem contexto disciplinar pode induzir comparações injustas. Antes de construir a figura, o pesquisador deve formular a sentença: 'esta visualização serve para mostrar que...'. Se a sentença não puder ser completada de modo preciso e verificável, a figura ainda não está pronta.

O segundo critério é a adequação da unidade de análise. Muitos erros visuais decorrem de misturar artigos, autores, instituições, países e áreas em uma mesma narrativa sem distinguir níveis. Um gráfico de produção por país não permite inferir desempenho de instituições específicas. Uma rede de coautoria entre autores não descreve necessariamente colaboração entre países. Um mapa de palavras-chave não equivale a mapa de teorias. A IA pode ajudar a auditar essa coerência, respondendo a perguntas como: 'qual é a unidade de análise desta visualização?', 'qual conclusão seria ilegítima?' e 'há mudança de nível interpretativo?'.

O terceiro critério é a transparência dos filtros. Visualizações cientométricas raramente mostram todos os dados. Elas aplicam filtros por frequência, citação, ano, tipo documental, país, idioma, base ou área. Filtros são necessários, mas precisam ser justificados. Uma rede que inclui apenas autores com cinco ou mais documentos mostra o núcleo produtivo, não toda a colaboração. Um mapa temático com palavras-chave acima de dez ocorrências exclui termos emergentes. A IA pode ser usada para revisar notas metodológicas e verificar se os filtros foram comunicados. Sem essa informação, o leitor pode interpretar a figura como representação total do campo.

O quarto critério é a coerência das escalas. Eixos truncados, escalas duplas, áreas proporcionais mal calculadas, mapas coropléticos sem normalização e tamanhos de nós sem legenda podem distorcer a percepção. Em gráficos de produção, iniciar o eixo vertical acima de zero pode exagerar diferenças. Em mapas de rede, tamanhos de nós proporcionais a citações podem tornar invisíveis entidades relevantes por colaboração. Em mapas geográficos, contagens absolutas favorecem sistemas científicos maiores. A IA deve ser solicitada a revisar riscos de distorção de escala e sugerir alternativas mais transparentes.

O quinto critério é a legibilidade. Uma figura cientificamente correta, mas ilegível, falha como comunicação. Rótulos sobrepostos, fontes pequenas, excesso de cores, legendas ausentes, categorias demais e baixa resolução reduzem a capacidade de interpretação. A IA pode apoiar uma revisão de legibilidade com base em descrição da figura ou imagem, indicando elementos que devem ser simplificados. Entretanto, a avaliação humana permanece indispensável, especialmente

porque a legibilidade depende do meio final. O que é legível em tela grande pode não ser legível em página impressa. A revisão deve ocorrer no formato final de circulação.

O sexto critério é evitar codificações visuais ambíguas. Se a cor representa *cluster*, não deve simultaneamente representar área geográfica. Se o tamanho representa produção, não deve ser interpretado como impacto. Se a espessura da aresta representa frequência, a legenda não pode sugerir intensidade qualitativa. Ambiguidade visual produz interpretações múltiplas e reduz confiabilidade. A IA pode ajudar a criar uma 'gramática da figura', descrevendo cada elemento visual e seu significado. Quando o modelo não consegue explicar a codificação de forma simples, é provável que a figura esteja excessivamente complexa.

O sétimo critério é a validação dos rótulos. Em mapas temáticos e redes de termos, rótulos automáticos podem parecer plausíveis, mas ocultar heterogeneidade. A validação por especialistas deve verificar se os termos centrais realmente representam o *cluster*, se há subtemas relevantes, se o rótulo não é amplo demais e se não há viés terminológico. A IA pode propor rótulos alternativos e graus de confiança, mas não deve decidir sozinha. Em relatórios, é recomendável informar que os rótulos foram atribuídos por análise assistida por IA e validados por revisão humana. Essa transparência aumenta confiança.

O oitavo critério é a preservação da incerteza. Muitas visualizações cientométricas apresentam resultados como se fossem definitivos. No entanto, dados podem variar conforme base, estratégia de busca, período e procedimentos de limpeza. Uma visualização bonita pode apagar essa contingência. Formas de comunicar incerteza incluem notas metodológicas, intervalos, comparação entre bases, indicação de cobertura, material suplementar e linguagem prudente. A IA pode ser solicitada a reescrever legendas para incluir incerteza sem tornar o texto excessivamente longo. O objetivo não é enfraquecer a figura, mas impedir que ela prometa mais do que entrega.

O nono critério é evitar causalidade visual indevida. Sequência temporal não prova causa; proximidade em rede não prova influência direta; crescimento de publicações não prova aumento de qualidade; centralidade não prova liderança; colaboração internacional não prova excelência. Visualizações sugerem padrões que podem gerar hipóteses. A IA deve ser instruída a usar verbos adequados: 'indica', 'sugere', 'é compatível com', 'aponta para', 'requer confirmação'. Verbos como 'demonstra', 'comprova' ou 'causa' devem ser reservados para desenhos analíticos que os sustentem. Essa disciplina linguística é parte da qualidade visual.

O décimo critério é alinhar estética e função. A beleza não é inimiga da ciência; visualizações claras, harmônicas e bem desenhadas melhoram a compreensão. O problema surge quando estética substitui evidência. Cores, formas e animações devem servir à interpretação. Uma animação temporal pode ser útil para explorar trajetórias, mas pode dificultar comparação estática. Um gráfico 3D pode parecer sofisticado, mas distorcer proporções. A IA pode sugerir alternativas visualmente limpas e cientificamente adequadas, mas o pesquisador deve priorizar clareza. Em visualização científica, elegância é consequência da precisão comunicacional.

O décimo primeiro critério é testar a figura com leitores. Uma visualização pode parecer clara para o analista que conhece os dados, mas ser confusa para o público. Testes simples, como pedir a um colega que explique o que entendeu em trinta segundos, revelam problemas de legenda, escala ou foco. A IA pode simular perguntas de leitores, listar possíveis interpretações equivocadas e sugerir melhorias. Essa função é útil porque antecipa ruídos comunicacionais. Contudo, simulação não substitui teste com usuários reais, especialmente em *dashboards* institucionais.

O décimo segundo critério é manter rastreabilidade. Toda visualização deve poder ser reconstituída a partir de dados, scripts, parâmetros e decisões documentadas. Se o gráfico foi

gerado por ferramenta interativa, deve-se registrar versão, filtros e exportação. Se a IA sugeriu transformações, deve-se registrar o *prompt* e a decisão humana. Sem rastreabilidade, a figura se torna difícil de auditar e corrigir. Em ambientes de gestão científica, a rastreabilidade protege contra disputas interpretativas e aumenta legitimidade. Visualização responsável é visualização reprodutível.

Quadro 15.1 - Dicas de como evitar visualizações frágeis

Nº	Princípio e recomendações	Perguntas a fazer à IA
01	Comece pela pergunta explícita Toda visualização deve responder a uma pergunta clara. Evite figuras geradas apenas porque a ferramenta permite. Antes de construir, complete a sentença: "Esta visualização serve para mostrar que...".	"Qual pergunta esta figura responde?" "Que conclusão legítima ela permite?" "Que conclusões seriam inegítimas?"
02	Adeque a unidade de análise e o nível de interpretação Defina claramente a unidade: artigo, autor, instituição, país, área, termo, documento etc. Não misture níveis (ex.: autor × país) na mesma narrativa. Cada tipo de aresta representa uma relação específica.	"Qual é a unidade de análise?" "Existe mudança de nível na interpretação?" "Quais inferências seriam indevidas?"
03	Torne transparentes os filtros e parâmetros Informe fontes, bases, período, tipos documentais, idiomas e áreas. Explique limiares (ex.: mínimo de ocorrências, citações, documentos). Filtros moldam a figura: sem transparência, o leitor não entende o que foi incluído ou excluído.	"Quais filtros foram aplicados?" "Eles podem excluir temas emergentes ou atores relevantes?" "Como comunicar isso?"
04	Respeite as escalas e normalizações Use escalas adequadas e inicie eixos em zero, quando pertinente. Normalize por campo, área, população ou tamanho quando comparar. Prefira taxas, proporções ou índices ajustados a contagens absolutas.	"Há risco de distorção de escala?" "A figura favorece sistemas maiores?" "Que normalizações seriam adequadas?"
05	Priorize legibilidade e clareza Reduza ruído visual: rótulos legíveis, cores com propósito, contraste. Evite sobreposição, excesso de categorias e elementos decorativos. Teste a figura no meio final (tela, impressão, slide).	"Esta figura é legível para um leitor não especialista?" "O que pode ser simplificado sem perder informação?"
06	Evite codificações ambíguas e explique tudo Cada cor, tamanho, forma e espessura deve ter um único significado. Inclua legenda completa e unidades de medida, quando aplicável. Explique o que cada elemento representa e como foi calculado.	"A legenda é suficiente e não ambígua?" "Há sobreposição de significados visuais?" "Como tornar a codificação mais clara?"

Nº	Princípio e recomendações	Perguntas a fazer à IA
07	Valide rótulos e agrupamentos Rótulos automáticos podem enganar. Revise *clusters*, temas e comunidades com especialistas. Informe o método de rotulagem e o grau de confiança.	"Quais rótulos alternativos fazem sentido?" "Quão confiável é este rótulo?" "Que subtemas podem estar ocultos?"
08	Explore e comunique a incerteza Mostre limitações de cobertura, dados faltantes e atrasos de indexação. Compare bases e testes de sensibilidade quando possível. Use linguagem prudente e explicite o nível de confiança.	"Quais incertezas esta figura envolve?" "Como comunicar limitações sem reduzir a utilidade?"
09	Não inferir causalidade indevida Correlação, proximidade ou sequência não implicam causa. Use verbos adequados: indica, sugere, é compatível com, aponta para. Evite: demonstra, comprova, causa, prova.	"Que inferências causais podem ser induzidas?" "Como reescrever a legenda usando verbos apropriados?"
10	Alinhe estética à função A forma deve servir ao conteúdo, não o contrário. Evite 3D, efeitos e animações que distorcem ou distraem. Beleza equivale a clareza, harmonia e precisão comunicacional.	"Que alternativas visuais comunicam melhor com menos risco de distorção?" "Como simplificar sem empobrecer?"
11	Teste com leitores reais A clareza depende do público. Peça a alguém que explique o que entendeu em 30 segundos. Antecipe interpretações equivocadas.	"Que perguntas um leitor pode fazer?" "Quais interpretações equivocadas esta figura pode induzir?"
12	Garanta rastreabilidade e reprodutibilidade Registre dados, scripts, parâmetros, versões de software e ferramentas. Guarde *prompts* utilizados, decisões e datas. Permita reconstruir a figura.	"Que informações devem ser registradas para reproduzir esta figura?" "Como organizar um anexo metodológico?"

Fonte: Elaborado pelo autor (2026).

A prevenção de fragilidade metodológica também exige cultura institucional. Relatórios e apresentações devem valorizar notas metodológicas, anexos técnicos e linguagem cautelosa, em vez de premiar apenas impacto visual. Avaliadores e gestores devem ser capacitados para ler gráficos criticamente. A IA pode ajudar na produção de versões pedagógicas e glossários visuais, explicando termos como centralidade, *cluster*, densidade, normalização e intervalo temporal. Assim, a visualização torna-se instrumento de formação, não apenas de convencimento.

Em síntese, a visualização bonita e robusta é possível quando estética, método e interpretação estão alinhados. A IA pode acelerar a geração de alternativas, mas deve ser usada como instrumento de crítica e não apenas de produção. O pesquisador deve pedir ao

modelo que identifique riscos, proponha simplificações, revise legendas, explicite limitações e verifique coerência com a pergunta. O capítulo propõe que toda figura cientométrica passe por uma pergunta final: se o leitor confiar plenamente nesta visualização, o que ele poderá concluir corretamente e o que poderá concluir indevidamente? A resposta a essa pergunta define a responsabilidade da visualização.

15.8 Prompts do capítulo

Os *prompts* apresentados a seguir foram elaborados para funcionar como instrumentos metodológicos de apoio à visualização científica assistida por IA. Eles não devem ser usados como comandos isolados, mas como componentes de um protocolo. Em todos os casos, recomenda-se fornecer dados, contexto, objetivo, unidade de análise, público-alvo, restrições de formato e limitações conhecidas. Quando a IA produzir recomendações, o pesquisador deve revisá-las, registrar alterações e documentar decisões. A finalidade dos *prompts* é ampliar a qualidade da análise visual, não transferir responsabilidade ao modelo (WEI *et al.*, 2022; WANG *et al.*, 2023).

Prompt 15.8.1 Prompt para escolher visualização

Atue como especialista em cientometria, visualização de dados e comunicação científica. Tenho a seguinte pergunta de pesquisa: [PERGUNTA].
A unidade de análise é [ARTIGO/AUTOR/INSTITUIÇÃO/PAÍS/TEMA/REDE].
As variáveis disponíveis são [LISTA]. O período analisado é [PERÍODO].
O público-alvo é [PÚBLICO].
Recomende três opções de visualização, indicando para cada uma: objetivo analítico, tipo de gráfico ou mapa, variáveis nos eixos ou codificações visuais, vantagens, limitações, riscos de interpretação indevida, legenda mínima necessária e alternativa mais simples.
Não recomende visualizações apenas pela estética; priorize adequação à pergunta e transparência metodológica.

Prompt 15.8.2 Prompt para revisar legibilidade de gráfico

Atue como revisor metodológico de visualizações científicas.

A figura planejada é descrita assim: [DESCRIÇÃO DA FIGURA].

Avalie a legibilidade considerando: clareza do título, adequação dos eixos, escala, rótulos, legenda, número de categorias, uso de cores, acessibilidade, sobrecarga visual, coerência entre dado e codificação visual e risco de distorção.

Apresente uma tabela com problemas identificados, gravidade, justificativa e recomendação de correção.

Ao final, proponha uma versão revisada da figura em linguagem operacional.

Prompt 15.8.3 Prompt para escrever legenda analítica

Atue como redator científico especializado em cientometria.

Escreva uma legenda analítica para a seguinte visualização: [DESCRIÇÃO].

A legenda deve conter: o que a figura mostra diretamente, fonte e período dos dados, unidade de análise, método de cálculo ou construção, significado das cores/tamanhos/eixos, principal padrão observado, cautela interpretativa e uma frase sobre o que a figura não permite concluir.

Use tom formal e evite causalidade indevida.

Prompt 15.8.4 Prompt para transformar achados em narrativa visual

Atue como analista de comunicação científica.

A partir dos achados abaixo [LISTA DE ACHADOS], construa uma narrativa visual para relatório técnico.

Organize a narrativa em sequência de figuras, indicando para cada uma: mensagem principal, visualização recomendada, dados necessários, título descritivo, legenda curta, nota metodológica e transição textual para a próxima figura.

Separe constatações de hipóteses e indique onde é necessária validação adicional.

Prompt 15.8.5 Prompt para auditar causalidade visual

Atue como auditor de inferências em visualizações cientométricas.

Examine a seguinte interpretação de uma figura: [INTERPRETAÇÃO].

Identifique afirmações descritivas, inferências plausíveis, hipóteses, causalidades não demonstradas e extrapolações indevidas.

Reescreva a interpretação em versão metodologicamente prudente, preservando a utilidade comunicacional, mas eliminando conclusões que a figura não sustenta.

Prompt 15.8.6 Prompt para especificar dashboard científico

Atue como especialista em *dashboards* para gestão científica.

Com base nos objetivos [OBJETIVOS], público [PÚBLICO], indicadores [INDICADORES] e fontes [FONTES], proponha a arquitetura de um *dashboard* cientométrico.

Defina páginas, filtros, visualizações, indicadores-chave, notas metodológicas, níveis de drill-down, alertas de baixa confiabilidade e campos de documentação.

Inclua recomendações para evitar rankings simplistas e uso avaliativo indevido.

15.9 Estudo de caso ilustrativo: visualização de uma área emergente

Para ilustrar a aplicação dos princípios discutidos, considere um estudo cientométrico sobre inteligência artificial na saúde digital entre 2014 e 2024. O *corpus* contém artigos indexados em duas bases bibliográficas, após deduplicação e padronização de palavras-chave. O objetivo é comunicar a evolução do tema a um comitê de pesquisa que precisa identificar áreas consolidadas, temas emergentes e oportunidades de colaboração. O analista dispõe de publicações por ano, citações normalizadas, países de afiliação, redes de coautoria, *clusters* de palavras-chave e classificação temática assistida por IA. A tentação inicial seria construir um *dashboard* com todos os gráficos possíveis. Contudo, a metodologia recomendada começa pela pergunta: que decisões o comitê precisa tomar?

A primeira figura do relatório poderia ser uma série temporal de publicações anuais, com anotação de eventos relevantes, como mudanças regulatórias, pandemia ou lançamento de políticas nacionais de IA. A legenda deveria afirmar que a figura mostra crescimento de documentos indexados, não crescimento de qualidade científica. A segunda figura poderia ser um gráfico de dispersão comparando produção e impacto normalizado por país, com destaque para países de interesse institucional. A terceira poderia ser uma rede de colaboração internacional, filtrada por peso mínimo de coautoria e acompanhada de nota sobre o significado das arestas. A quarta poderia ser um mapa temático de palavras-chave, com rótulos validados por especialistas. A quinta poderia ser uma linha do tempo de emergência de termos, diferenciando mudança temática de mudança terminológica.

Nesse caso, a IA poderia ser usada em quatro momentos. Primeiro, para recomendar a sequência de visualizações com base nas perguntas do comitê. Segundo, para revisar se cada gráfico é legível e se a legenda explicita fonte, método e limitação. Terceiro, para propor rótulos alternativos aos *clusters* temáticos, indicando grau de confiança e termos que justificam cada rótulo. Quarto, para transformar os achados em narrativa visual, separando constatações, hipóteses e recomendações. Em nenhum momento a IA deveria decidir sozinha que tema é estratégico ou que país é prioritário. Essa decisão depende de contexto institucional, conhecimento substantivo e objetivos de política científica.

Um possível erro seria usar um diagrama de Sankey para mostrar que 'telemedicina evoluiu para IA generativa', apenas porque termos associados aparecem em períodos posteriores. Tal figura poderia sugerir uma transição causal inexistente. Uma abordagem mais prudente seria afirmar que há redistribuição de termos e coexistência de *clusters*, exigindo leitura qualitativa dos documentos centrais. Outro erro seria usar tamanho de nó em rede para representar citações e, ao mesmo tempo, interpretar o nó como líder de colaboração. A mesma marca visual não pode representar duas dimensões sem ambiguidade. A IA, quando bem orientada, pode identificar esses riscos antes da publicação do relatório.

O estudo de caso mostra que visualização científica assistida por IA não significa automatizar a comunicação de resultados. Significa estruturar uma colaboração controlada entre pesquisador, modelo, dados e público. O pesquisador define a pergunta, fornece contexto, valida recomendações e assume responsabilidade. A IA sugere, critica, redige e compara alternativas. Os dados impõem limites. O público orienta o nível de detalhe e a forma de explicação. Essa articulação é o núcleo da visualização responsável em cientometria. O produto final deve ser visualmente claro, metodologicamente rastreável e interpretativamente prudente.

15.10 Checklist de qualidade e recomendações finais

A qualidade de uma visualização cientométrica pode ser avaliada por um checklist simples, mas rigoroso. Primeiro, a figura responde a uma pergunta explícita? Segundo, a unidade de análise

está clara? Terceiro, a fonte de dados, período e critérios de inclusão foram informados? Quarto, filtros e normalizações foram documentados? Quinto, o tipo de visualização é adequado à tarefa analítica? Sexto, eixos, legendas, cores e tamanhos estão explicados? Sétimo, a figura é legível no formato final? Oitavo, há nota sobre limitações e incertezas? Nono, a legenda separa descrição, inferência e hipótese? Décimo, o leitor pode reconstituir a figura a partir dos dados e parâmetros?

Quadro 15.2 - Quadro-síntese: relação entre pergunta, dado e visualização

Pergunta cientométrica	Dados necessários	Visualizações possíveis	Risco principal
Como evoluiu a produção no tempo?	Ano de publicação e contagem de documentos	Série temporal, barras anuais, média móvel	Confundir crescimento de registros com aumento de qualidade científica.
Quem colabora com quem?	Relações de coautoria, instituição ou país	Mapa de rede, matriz, mapa geográfico com arestas	Inferir colaboração informal ou causalidade apenas por coautoria.
Quais temas estruturam a área?	Palavras-chave, títulos, resumos ou *clusters*	Mapa temático, rede de coocorrência, MCA	Tratar rótulos automáticos como categorias definitivas.
Como temas mudam entre períodos?	*Clusters* por período, termos emergentes, fluxos	Linha do tempo, Sankey, diagrama aluvial	Interpretar fluxo visual como evolução causal.
Como monitorar indicadores institucionais?	Indicadores agregados e filtros por unidade	*Dashboard* com camadas e drill-down	Transformar monitoramento em avaliação automática.

Fonte: Elaborado pelo autor (2026).

Esse checklist também pode ser convertido em *prompt* de auditoria. O pesquisador descreve a figura, informa os dados e pede que a IA avalie os dez critérios, atribuindo risco baixo, médio ou alto. A resposta não substitui revisão humana, mas ajuda a identificar pontos cegos. Em projetos complexos, recomenda-se aplicar a auditoria antes de inserir figuras em relatório final. Quando houver *dashboards*, o checklist deve ser aplicado a cada componente e ao conjunto da interface. Uma figura isolada pode ser correta, mas o painel como um todo pode sugerir narrativa indevida se ordenar, colorir ou agrupar indicadores de modo enviesado.

A principal recomendação prática é adotar um ciclo iterativo: pergunta, escolha visual, construção, revisão, validação, legenda, teste com leitor e documentação. A IA pode participar de todas as etapas, mas sempre com papéis definidos. Como consultora, sugere alternativas. Como revisora, identifica problemas. Como redatora, cria legendas. Como auditora, aponta inferências indevidas. Como tradutora, adapta linguagem para públicos distintos. Essa multiplicidade de papéis só é segura quando *prompts* explicitam restrições e quando o pesquisador mantém controle metodológico. A visualização científica assistida por IA deve ser vista como uma disciplina de projeto, não como geração instantânea de figuras.

Em conclusão, a visualização é uma das etapas em que a cientometria se torna socialmente inteligível. Indicadores e redes ganham forma, padrões se tornam discutíveis e resultados passam a circular em artigos, relatórios, apresentações e decisões institucionais. Essa potência demanda responsabilidade. A IA pode tornar a visualização mais rápida, acessível e reflexiva, mas também pode amplificar equívocos se operar sem contexto e validação. O objetivo deste capítulo foi oferecer princípios, critérios e *prompts* para que a visualização

científica assistida por IA seja clara, rigorosa, rastreável e útil. O bom gráfico não é aquele que impressiona primeiro, mas aquele que continua correto depois de ser examinado criticamente.

15.11 Protocolo operacional para visualização científica assistida por IA

Um protocolo operacional de visualização científica assistida por IA deve começar antes da construção da primeira figura. A etapa inicial consiste em registrar a pergunta de pesquisa, a decisão que a visualização deve apoiar e a unidade de análise que será representada. Essa orientação evita que a figura seja criada por disponibilidade de ferramenta ou por preferência estética. Em seguida, o pesquisador deve listar os dados necessários, seus campos, suas limitações e seus tratamentos prévios. Só depois deve solicitar à IA recomendações de visualização. Essa ordem é importante porque modelos generativos tendem a responder de acordo com padrões textuais aprendidos, e não com inspeção real dos dados, a menos que o usuário forneça informações suficientes sobre estrutura, distribuição e contexto metodológico.

A segunda etapa do protocolo consiste em gerar alternativas visuais. Para cada pergunta, recomenda-se solicitar ao modelo pelo menos três possibilidades, incluindo uma opção simples, uma opção relacional ou temporal quando pertinente e uma opção adequada a públicos especializados. A resposta da IA deve ser avaliada segundo critérios de adequação, legibilidade, risco de distorção, necessidade de normalização e compatibilidade com o meio de publicação. Essa etapa transforma a IA em parceira de desenho analítico, mas não em decisor final. O pesquisador compara alternativas e seleciona aquela que melhor responde à pergunta com menor custo cognitivo para o leitor.

A terceira etapa é a construção da visualização em ferramenta apropriada. A IA pode auxiliar com pseudocódigo, comandos em R ou Python, especificações para VOSviewer, bibliometrix, Gephi, Power BI ou Tableau, mas o código deve ser testado. Em cientometria, recomenda-se preservar scripts, parâmetros e versões de pacotes. Quando a visualização depende de ferramenta gráfica interativa, como *dashboards* ou mapas de rede ajustados manualmente, deve-se registrar capturas de parâmetros, filtros, *thresholds* e algoritmos. O valor científico da figura depende da possibilidade de reconstrução. Uma visualização que não pode ser reproduzida deve ser tratada como ilustração exploratória, não como evidência central.

A quarta etapa envolve revisão de legibilidade e coerência. O pesquisador deve verificar títulos, eixos, unidades, legendas, escala, cores, contrastes, rótulos, posição de anotações e densidade visual. Nessa etapa, um *prompt* de revisão pode pedir que a IA identifique elementos que confundem, omitem ou exageram. O modelo pode produzir uma lista de problemas potenciais, mas a verificação deve ocorrer sobre a figura real, preferencialmente no formato final. A revisão também deve incluir acessibilidade: legibilidade para impressão, contraste, leitura por pessoas com deficiência visual parcial e independência semântica da cor.

A quinta etapa é a redação da legenda analítica. A legenda deve ser mais do que uma descrição breve. Ela deve explicar o que a figura mostra, quais dados foram usados, qual período foi considerado, que transformação foi aplicada e que interpretação é permitida. Uma legenda robusta também informa o que não se pode concluir. Essa prática reduz o risco de leituras indevidas, sobretudo quando figura circular fora do texto completo, como em slides ou relatórios executivos. A IA é particularmente útil nessa etapa, pois pode gerar versões em diferentes níveis de detalhe, mas a redação final deve preservar linguagem científica e evitar extrapolações.

A sexta etapa é a validação por pares ou por especialistas. Visualizações temáticas, redes e *dashboards* institucionais devem ser examinados por pessoas que conhecem a área e por pessoas que representam o público-alvo. O especialista avalia coerência substantiva; o usuário avalia compreensibilidade. A IA pode simular perguntas de revisão, mas não

substitui leitura crítica real. Em estudos com impacto institucional, recomenda-se registrar comentários de validação e ajustes realizados. Essa prática aproxima a visualização de procedimentos de qualidade científica e reduz a chance de que decisões sejam tomadas com base em figuras ambíguas.

A sétima etapa é a documentação do uso de IA. O protocolo deve registrar que *prompts* foram utilizados para escolha visual, revisão, legenda, narrativa ou auditoria. Deve indicar se a IA teve acesso aos dados brutos, a resumos estatísticos ou apenas a descrições. Deve informar quais sugestões foram aceitas, rejeitadas ou modificadas. Essa documentação não precisa ocupar o corpo principal de um artigo, mas deve existir como registro de pesquisa. Em uma ciência cada vez mais assistida por modelos generativos, a transparência sobre o papel da IA é parte da integridade metodológica.

A etapa final é a integração narrativa. A figura deve entrar no texto no momento em que responde a uma pergunta ou sustenta um argumento. Figuras soltas, sem articulação com a análise, tendem a ser decorativas. A IA pode ajudar a construir transições entre visualizações, criando uma sequência argumentativa: primeiro, magnitude, depois distribuição, depois relação, depois evolução e, por fim, implicações. Essa sequência deve ser revisada para garantir que o texto não adicione conclusões que as figuras não sustentam. A narrativa visual é boa quando conduz o leitor de evidência em evidência, sem saltos interpretativos.

15.12 Boas práticas por público e meio de comunicação

A visualização para artigos científicos deve priorizar precisão, transparência metodológica e reprodutibilidade. O leitor acadêmico espera compreender a fonte dos dados, os filtros, os métodos e as limitações. Figuras em artigos podem ser mais densas do que em relatórios executivos, desde que sejam acompanhadas de legendas completas e, quando necessário, materiais suplementares. A IA pode auxiliar na redação de legendas e na seleção de figuras principais, mas deve evitar transformar a seção de resultados em narrativa promocional. O tom adequado é analítico, não persuasivo. Em artigos, cada figura deve estar diretamente vinculada a uma afirmação do texto.

A visualização para gestores de pesquisa deve equilibrar clareza e cautela. Gestores precisam de sínteses, mas decisões de política científica não devem se basear em simplificações métricas. Recomenda-se utilizar visualizações que indiquem tendências, distribuição, riscos e oportunidades, sem hierarquizar indevidamente pessoas ou grupos. *Dashboards* e relatórios executivos devem incluir notas metodológicas visíveis e explicações curtas sobre indicadores. A IA pode gerar versões executivas de achados, mas deve ser instruída a não usar linguagem competitiva ou avaliativa quando os dados apenas descrevem padrões. Expressões como ‘lidera’, ‘fica atrás’ ou ‘desempenho fraco’ devem ser cuidadosamente controladas.

A visualização para apresentações orais deve ser ainda mais seletiva. Slides exigem baixo volume de informação por figura, títulos que expressem a mensagem principal e elementos visuais que possam ser compreendidos rapidamente. Uma rede muito densa pode ser útil em artigo, mas inadequada em apresentação. O apresentador pode dividir a análise em sequência: primeiro mostrar a rede completa como contexto, depois destacar um subgrupo, depois apresentar a interpretação. A IA pode ajudar a transformar uma figura complexa em roteiro de apresentação, indicando que partes devem ser explicadas verbalmente e quais devem aparecer no slide.

A visualização para *dashboards* interativos deve considerar autonomia do usuário. Diferentemente do artigo, o usuário pode explorar filtros em ordem não prevista. Portanto, a interface precisa orientar a leitura. Títulos dinâmicos, *tooltips*, notas de rodapé, mensagens

de alerta e painéis de metodologia são componentes essenciais. A IA pode sugerir textos de ajuda e alertas contextuais, como 'a amostra filtrada contém menos de 30 documentos; interprete percentuais com cautela'. Essa capacidade de transformar regras metodológicas em mensagens compreensíveis é uma das contribuições mais úteis dos modelos generativos à visualização científica aplicada à gestão.

A visualização para divulgação pública deve ser simples, honesta e contextualizada. Dados cientométricos podem ser mal interpretados por públicos não especializados, especialmente quando envolvem rankings, países, instituições ou temas sensíveis. A figura deve evitar jargão, mas não deve esconder limitações. A IA pode produzir versões em linguagem acessível, glossários e explicações em camadas. Entretanto, simplificar não significa eliminar incerteza. Uma comunicação pública responsável explica que bases bibliográficas não cobrem toda a ciência, que citações não equivalem automaticamente a qualidade e que mapas de rede mostram relações documentadas, não a totalidade das interações científicas.

O meio impresso impõe restrições específicas: tamanho fixo, ausência de interação, dependência de resolução e possível impressão em escala de cinza. Figuras com muitas cores ou rótulos pequenos podem perder legibilidade. Para impressão, gráficos de barras, linhas, mapas simples e tabelas bem desenhadas costumam ser mais robustos do que visualizações densas. A IA pode ser solicitada a adaptar uma figura digital para versão impressa, reduzindo categorias, sugerindo agrupamentos e reescrevendo legenda. Essa adaptação deve preservar a mensagem analítica, não apenas converter formato.

O meio digital permite interação, mas nem toda interação é necessária. Filtros, *zoom*, *tooltips* e animações devem responder a necessidades analíticas. Interatividade excessiva pode aumentar a carga cognitiva e dificultar comparação. Em *dashboards* cientométricos, o usuário deve saber qual filtro está ativo, quantos registros restam e como retornar ao estado inicial. A IA pode propor fluxos de interação e mensagens de orientação, mas a usabilidade deve ser testada com usuários. Visualização interativa responsável é aquela que aumenta controle interpretativo, não aquela que oferece possibilidades infinitas sem orientação.

Em qualquer público e meio, a boa prática central é alinhar figura, texto e decisão. O gráfico não deve dizer uma coisa enquanto a legenda sugere outra. O *dashboard* não deve ordenar por um indicador e afirmar conclusão baseada em outro. A apresentação não deve dramatizar diferença pequena. A IA deve ser utilizada para revisar essa consistência, funcionando como leitora crítica. Um *prompt* pode pedir: 'compare a mensagem do título, a legenda e os dados descritos; identifique inconsistências'. Essa verificação simples pode evitar muitos problemas de comunicação científica.

15.13 Ética, governança e responsabilidade nas visualizações assistidas por IA

A ética da visualização científica em cientometria começa pelo reconhecimento de que figuras influenciam decisões. Um gráfico pode afetar a reputação de uma instituição, a prioridade de uma área, a percepção sobre um grupo ou a justificativa de uma política. Quando a IA participa da produção visual, a responsabilidade não se desloca para o sistema. O pesquisador continua responsável por escolhas, interpretações e comunicação. Essa responsabilidade inclui verificar dados, corrigir erros, evitar rótulos estigmatizantes, proteger informações sensíveis e impedir que métricas parciais sejam apresentadas como avaliação completa (HICKS *et al.*, 2015).

A governança de visualizações institucionais deve definir quem pode produzir, revisar, publicar e atualizar painéis. *Dashboards* cientométricos não são apenas ferramentas técnicas; são artefatos de governança. Devem ter responsáveis, documentação, periodicidade, processo

de correção e critérios de acesso. Quando a IA é usada para gerar narrativas automáticas, é necessário estabelecer revisão humana antes da divulgação. Em contextos sensíveis, como avaliação de departamentos ou pesquisadores, a narrativa automática pode produzir danos se interpretar mal os dados incompletos. A prudência metodológica deve ser incorporada à infraestrutura, não depender apenas da boa vontade do analista.

A justiça interpretativa é outro princípio relevante. Visualizações que comparam áreas distintas devem reconhecer diferenças disciplinares. Uma área com menor número de artigos pode ter outras formas de contribuição científica, como livros, relatórios técnicos, software, bases de dados, patentes, normas ou impacto social. Um *dashboard* baseado apenas em artigos indexados pode favorecer determinadas disciplinas e invisibilizar outras. A IA pode ajudar a produzir notas de contextualização disciplinar e sugerir indicadores complementares, mas o desenho final deve refletir pluralidade de práticas científicas. Métricas e visualizações não são neutras em relação aos campos que representam.

A privacidade também deve ser considerada. Dados bibliográficos são públicos em muitos casos, mas visualizações agregadas podem expor indivíduos, grupos pequenos ou trajetórias sensíveis. Ao representar pesquisadores por produtividade, colaboração ou impacto, o analista deve avaliar finalidade, proporcionalidade e contexto de uso. Em alguns casos, é preferível agregar por programa, área ou período, evitando rankings individuais. A IA pode ser solicitada a revisar riscos de exposição e sugerir agregações mais seguras. Essa preocupação é especialmente importante quando painéis são disponibilizados publicamente ou usados em processos decisórios.

A transparência sobre IA é parte da ética. Quando rótulos de *clusters*, legendas, narrativas ou recomendações visuais foram produzidas com apoio de modelos generativos, essa participação deve ser documentada. A transparência não significa desvalorizar o trabalho; ao contrário, mostra que houve método. O leitor não precisa conhecer cada interação, mas deve poder saber que decisões substantivas foram validadas por humanos. Em áreas nas quais a confiança na evidência é central, esconder o papel da IA pode comprometer credibilidade. A boa prática é explicitar o uso de IA em seção metodológica, apêndice ou documentação técnica.

A responsabilidade visual também envolve evitar manipulação retórica. Figuras podem ser desenhadas para persuadir de modo indevido: escalas exageradas, cores alarmistas, exclusão seletiva de períodos, agrupamentos convenientes ou títulos conclusivos. A IA, se orientada a maximizar impacto, pode reforçar essas práticas. Por isso, *prompts* devem priorizar clareza, precisão e cautela, não persuasão. Em vez de pedir 'crie uma visualização impactante', o pesquisador deve pedir 'crie uma visualização clara, proporcional e metodologicamente defensável'. A linguagem do *prompt* molda a ética da resposta.

A governança responsável recomenda separar visualizações exploratórias de visualizações publicáveis. Durante a exploração, é legítimo testar mapas densos, filtros alternativos, escalas e layouts. Contudo, nem toda visualização exploratória deve entrar no relatório. Figuras publicáveis precisam passar por validação. A IA pode organizar um registro das versões testadas e explicar por que certas alternativas foram descartadas. Esse registro ajuda a mostrar que a escolha final não foi arbitrária. Em ciência aberta e avaliação responsável, documentar o caminho visual fortalece a confiabilidade dos resultados.

Finalmente, a ética da visualização assistida por IA exige reconhecer limites. Alguns fenômenos científicos não são bem capturados por dados bibliométricos. Nem toda contribuição é indexada; nem toda influência é citada; nem todo conhecimento circula por artigos; nem toda colaboração aparece em coautoria. Visualizações cientométricas são mapas parciais, não

territórios completos. A IA deve ajudar a comunicar essa parcialidade. Uma figura responsável não enfraquece a análise ao reconhecer limites; ela a torna mais confiável. A maturidade metodológica consiste em mostrar padrões sem esquecer as condições que os produziram.

15.14 Exemplo aplicado de sequência de prompts para um relatório visual

Uma sequência de *prompts* é mais adequada do que um *prompt* único quando o objetivo é produzir relatório visual cientométrico. O primeiro *prompt* deve solicitar diagnóstico do material disponível: quais variáveis existem, que perguntas podem ser respondidas, quais visualizações são possíveis e quais análises seriam metodologicamente frágeis. O segundo *prompt* deve selecionar a sequência narrativa: que figura vem primeiro, que padrão ela mostra e como prepara a próxima visualização. O terceiro *prompt* deve revisar cada figura segundo critérios de legibilidade e risco de interpretação. O quarto *prompt* deve redigir legendas analíticas. O quinto *prompt* deve auditar a narrativa final, verificando se há extrapolações. Essa cadeia reduz o risco de uma resposta única, excessivamente confiante, transformar dados parciais em relatório aparentemente completo.

Considere um *corpus* com 4.000 documentos, 12.000 autores, 75 países, 40.000 referências citadas e 8.000 palavras-chave. Um *prompt* único provavelmente sugerirá gráficos genéricos: linha temporal, rede de coautoria, mapa de palavras-chave e ranking de países. Uma cadeia de *prompts*, ao contrário, pode perguntar se há dados suficientes para cada figura, se os autores foram desambiguados, se as palavras-chave foram normalizadas, se os países devem ser contados por autoria fracionária ou inteira e se o impacto deve ser normalizado por área e ano. A diferença entre uma sugestão genérica e uma recomendação metodológica está justamente na exigência de verificação prévia.

O primeiro *prompt* da cadeia pode ser formulado assim: 'Avalie os dados disponíveis para um relatório visual cientométrico. Para cada variável, indique que pergunta pode responder, que visualização seria adequada, que pré-processamento é necessário e que limitação deve ser comunicada'. Esse *prompt* transforma a IA em examinadora de prontidão dos dados. Se a resposta indicar que redes de coautoria dependem de desambiguação ou que mapas temáticos dependem de normalização terminológica, o pesquisador ganha uma lista de tarefas antes de produzir figuras. Essa etapa é especialmente importante quando o *corpus* foi coletado de múltiplas bases.

O segundo *prompt* pode solicitar planejamento narrativo: 'Organize uma sequência de seis figuras para relatório executivo, começando por panorama geral e avançando para temas, colaboração e implicações'. A IA pode sugerir uma trajetória lógica: produção no tempo, distribuição por áreas, impacto, colaboração internacional, mapa temático e tendências emergentes. O pesquisador deve então verificar se essa sequência corresponde à pergunta do estudo. Em alguns casos, a narrativa deve começar pela estrutura temática, não pela produção. Em outros, a colaboração pode ser central. A IA propõe; o desenho de pesquisa decide.

O terceiro *prompt* deve revisar a adequação de cada figura. Ele pode pedir: 'Para cada visualização planejada, indique se há risco de ambiguidade, se a escala é apropriada, se a legenda deve incluir normalização e se existe alternativa mais simples'. Essa revisão frequentemente revela que figuras desejadas não são necessárias. Um mapa geográfico pode ser substituído por barras se a pergunta for apenas comparação de produção por país. Uma rede pode ser substituída por matriz se o objetivo for comparar intensidade bilateral. A simplificação é uma virtude quando preserva a mensagem.

O quarto *prompt* deve produzir legendas analíticas com linguagem prudente. A legenda deve dizer: 'A figura apresenta...', 'Os dados correspondem a...', 'O tamanho representa...', 'As

cores indicam...', 'O padrão observado sugere...', 'A figura não permite concluir...'. Essa estrutura evita duas falhas: legenda meramente descritiva, que não ajuda o leitor, e legenda interpretativa demais, que excede a evidência. A IA é útil para manter consistência entre legendas de várias figuras, mas o analista deve conferir se os termos técnicos estão corretos e se a legenda reflete o gráfico real.

O quinto *prompt* deve auditar o texto final: 'Leia a narrativa visual e identifique afirmações que não são sustentadas pelas figuras descritas'. Esse *prompt* pode revelar adjetivos excessivos, causalidade indevida, mudanças de unidade de análise ou generalizações. Por exemplo, a frase 'a instituição tornou-se referência internacional' pode não ser sustentada por um gráfico de aumento de colaboração; seria mais prudente afirmar 'a instituição ampliou sua presença em coautorias internacionais no *corpus* analisado'. A auditoria de linguagem é parte da auditoria visual, porque interpretação e figura circulam juntas.

A sequência de *prompts* deve ser arquivada com o relatório. O registro permite que outro pesquisador compreenda como as visualizações foram escolhidas e revisadas. Também permite atualizar o relatório no futuro, substituindo dados e mantendo lógica analítica. Em ambientes institucionais, essa prática cria memória metodológica. Em vez de depender de decisões tácitas de um analista, a organização passa a possuir um protocolo visual. A IA, nesse contexto, não é apenas ferramenta de produtividade; torna-se componente documentado de uma rotina de inteligência científica responsável.

Um benefício adicional da cadeia de *prompts* é a criação de versões para diferentes públicos. A mesma sequência pode gerar relatório técnico, sumário executivo e apresentação. A IA pode adaptar narrativa, densidade visual e glossário, preservando a mesma base de dados e as mesmas limitações. Essa capacidade é valiosa para universidades e centros de pesquisa que precisam comunicar resultados a comissões acadêmicas, gestores, financiadores e sociedade. Entretanto, toda adaptação deve preservar equivalência substantiva. Simplificar a linguagem não autoriza simplificar a evidência. A visualização assistida por IA deve ampliar acesso à compreensão, não reduzir a complexidade metodológica a slogans.

Portanto, a prática recomendada é substituir a pergunta isolada 'que gráfico usar?' por um encadeamento de perguntas: que problema visual preciso resolver, que dados sustentam a figura, que alternativa é mais adequada, que riscos existem, que legenda comunica sem exagerar e que conclusões permanecem proibidas? Essa sequência transforma a IA em parceira de reflexão metodológica. O resultado esperado não é apenas um gráfico melhor, mas uma cultura de visualização mais consciente, auditável e alinhada aos princípios da cientometria responsável.

15.15 Conclusão

A visualização científica assistida por IA deve ser compreendida como uma prática de tradução rigorosa entre dados bibliográficos e interpretação socialmente circulável. Ela não se reduz à escolha de gráficos, mas envolve seleção de evidências, codificação visual, definição de escalas, documentação de filtros, construção de legendas, adaptação a públicos e preservação de limites metodológicos. Em cientometria, a figura final é resultado de uma cadeia de decisões que começa na pergunta de pesquisa e passa por busca, coleta, limpeza, indicador, rede, *cluster*, normalização e narrativa. Quando essa cadeia é invisibilizada, a visualização assume autoridade excessiva.

A IA pode ampliar a capacidade de produzir visualizações melhores ao sugerir alternativas, identificar fragilidades, revisar legibilidade, transformar achados em narrativas e apoiar a comunicação para públicos diversos. Porém, sua utilidade depende da qualidade dos

prompts e da disciplina metodológica do pesquisador. *Prompts* eficazes fornecem contexto, tarefa, dados, restrições, público e critérios de validação. *Prompts* frágeis geram respostas genéricas. O uso responsável da IA em visualização científica exige que o pesquisador permaneça no comando das decisões analíticas e que registre como a IA contribuiu para o processo.

A principal contribuição deste capítulo é propor uma ética prática da visualização cientométrica: figuras devem ser claras, rastreáveis, proporcionais, acessíveis, interpretativamente prudentes e alinhadas à pergunta. Visualizações bonitas são desejáveis quando servem à compreensão; tornam-se problemáticas quando mascaram incerteza, impõem causalidade ou obscurecem decisões metodológicas. O futuro da cientometria com IA dependerá da capacidade de combinar automação, julgamento humano e alfabetização visual. O objetivo não é produzir mais gráficos, mas produzir gráficos que permitam pensar melhor a ciência.

Capítulo 16

Interpretação cientométrica com IA

16.1 Introdução

A interpretação é o ponto em que o estudo cientométrico deixa de ser apenas uma sequência de operações técnicas e passa a produzir significado sobre a organização, a dinâmica e a visibilidade da ciência. Até essa etapa, o pesquisador formulou perguntas, delimitou escopo, construiu estratégias de busca, coletou metadados, limpou registros, calculou indicadores, extraiu redes, produziu mapas e visualizações. Nenhuma dessas operações, isoladamente, responde ao que os resultados significam. Uma curva de crescimento, uma comunidade de coautoria ou um conjunto de artigos altamente citados são sinais empíricos que exigem contextualização. A interpretação cientométrica, portanto, é o exercício de converter padrões observados em enunciados prudentes sobre produção, circulação, influência, colaboração e transformação do conhecimento científico.

No contexto de uma trilha dedicada à cientometria com inteligência artificial, a interpretação assume importância ainda maior. A *IA generativa* pode produzir textos fluentes, organizar argumentos, sugerir explicações, apontar inconsistências e formular hipóteses alternativas. Contudo, essa fluência pode ocultar fragilidades metodológicas. Um modelo de linguagem pode redigir uma interpretação convincente de um mapa temático sem saber se os dados foram adequadamente coletados, se a estratégia de busca gerou viés, se a base bibliográfica tem cobertura suficiente ou se o algoritmo de *cluster*ização produziu agrupamentos estáveis. Por essa razão, este capítulo trata a IA como instrumento de apoio interpretativo, não como autoridade epistêmica autônoma. Esse cuidado é coerente com a literatura sobre alucinações em geração de linguagem natural e com os debates sobre limites sociais dos modelos de linguagem (BENDER *et al.*, 2021; JI *et al.*, 2023).

A interpretação cientométrica com IA deve ser orientada por princípios de métricas responsáveis. O Manifesto de Leiden recomenda que a avaliação quantitativa apoie, e não substitua, o julgamento qualitativo especializado; que os processos de coleta e análise sejam transparentes; que as diferenças entre campos sejam consideradas; e que se evite falsa precisão. Esses princípios são particularmente relevantes quando *prompts* são usados para transformar achados numéricos em textos interpretativos, porque o risco de extrapolação aumenta quando a linguagem natural atribui coerência causal a padrões que são apenas associativos. O problema não está em usar IA para redigir, mas em permitir que a redação apague as condições de produção dos dados. (HICKS *et al.*, 2015).

Quadro 16.1 - Níveis de interpretação cientométrica

Nível	Pergunta orientadora	Produto textual	Risco principal
Descrição objetiva	O que os dados mostram dentro do recorte definido?	Frases sobre valores, relações e padrões observáveis.	Omitir base, período, unidade de análise ou método.
Interpretação provável	O que os padrões sugerem quando contextualizados?	Explicações cautelosas, comparações e relações com literatura.	Transformar associação em causalidade.

Nível	Pergunta orientadora	Produto textual	Risco principal
Hipótese interpretativa	Que mecanismo poderia explicar o padrão?	Proposições que exigem evidência adicional.	Apresentar hipótese como conclusão.
Recomendação	Que ação seria coerente com os achados e objetivos?	Cenários, implicações e cautelas para decisão.	Derivar política diretamente de uma métrica.

Fonte: Elaborado pelo autor (2026).

Interpretação também não é sinônimo de avaliação. Uma análise cientométrica pode descrever que determinada instituição ampliou sua produção internacional, que um tema ganhou centralidade em mapas de coocorrência ou que uma comunidade de autores se tornou mais conectada. Avaliar se isso é bom, estratégico, inovador ou socialmente relevante exige critérios adicionais, missão institucional, revisão qualitativa de conteúdo, conhecimento de política científica e, frequentemente, comparação com objetivos normativos. A IA pode ajudar a explicitar esses níveis, mas deve ser instruída a separar descrição objetiva, interpretação provável e hipótese que requer evidência adicional. Essa distinção acompanha a crítica ao uso automático de indicadores como substitutos de avaliação qualitativa (DE RIJCKE *et al.*, 2016; LEYDESDORFF *et al.*, 2016).

Este capítulo propõe uma abordagem operacional para a interpretação cientométrica assistida por *prompts*. O capítulo está organizado em seis seções. A primeira discute a passagem da descrição dos dados à explicação prudente. A segunda examina a interpretação de indicadores em contexto. A terceira trata da formulação de hipóteses sem exagero conclusivo. A quarta apresenta a triangulação entre métricas, literatura e conhecimento especializado. A quinta discute o papel da IA como interlocutor crítico. A sexta orienta a redação interpretativa com transparência metodológica. Ao final, são apresentados *prompts*, matrizes de validação, um estudo de caso ilustrativo e referências bibliográficas.

16.2 Da descrição dos dados à explicação prudente

No plano conceitual, a passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos exige reconhecer que dados bibliográficos são registros parciais da atividade científica. Eles não capturam todas as formas de produção, aprendizagem, circulação e influência, mas oferecem indícios sistemáticos sobre práticas de publicação, citação e colaboração. Quando o objeto analisado envolve séries de publicação, distribuições de citação, mapas de rede, *clusters* temáticos e perfis institucionais, o intérprete deve perguntar que parte do fenômeno científico está representada e que parte permanece invisível. A IA pode ajudar a formular essa pergunta de modo recorrente, sobretudo quando o pesquisador solicita que o modelo apresente limites de cobertura, pressupostos de mensuração e possíveis dimensões ausentes. Essa operação impede que o resultado seja tomado como espelho completo da ciência.

A primeira tarefa interpretativa consiste em separar dado bruto, dado tratado, indicador e inferência. Um registro bibliográfico é um dado; uma contagem anual de publicações é um indicador simples; uma taxa de crescimento é um indicador derivado; a afirmação de que uma área está se consolidando é uma inferência. O risco específico desta etapa é confundir padrão observado com explicação causal ou transformar uma regularidade estatística em diagnóstico substantivo. Por isso, a interpretação deve declarar explicitamente a passagem de uma camada a outra. *Prompts* bem formulados podem exigir que a IA produza uma tabela com

quatro colunas: evidência observada, transformação realizada, interpretação possível e grau de incerteza. Essa estrutura reduz a ambiguidade retórica e favorece auditoria. Essa separação é compatível com a literatura sobre indicadores de citação, impacto e avaliação responsável (AKSNES; LANGFELDT; WOUTERS, 2019; WALTMAN, 2016).

A descrição objetiva não é trivial. Descrever cientometricamente significa informar valores, unidades, períodos, fontes, critérios de inclusão e procedimentos de normalização. Uma frase aparentemente simples, como "houve crescimento da literatura", precisa ser qualificada: crescimento absoluto ou relativo? crescimento de documentos, autores, periódicos ou citações? crescimento em qual base? crescimento comparado a que campo ou universo? A IA pode auxiliar propondo perguntas de verificação antes da redação de resultados, mas não sabe, por si mesma, quais dados foram realmente usados. O pesquisador deve fornecer tabelas, metadados e decisões metodológicas, e não apenas solicitar uma interpretação genérica.

A explicação prudente emerge quando o pesquisador aceita que uma mesma configuração de resultados admite diferentes leituras. Por exemplo, um aumento de publicações sobre inteligência artificial em saúde pode refletir crescimento real do campo, alteração terminológica, mudança de indexação, ampliação de financiamento ou combinação desses fatores. Nenhuma dessas explicações pode ser escolhida apenas pela elegância narrativa. A interpretação responsável apresenta alternativas e indica que evidências adicionais seriam necessárias para discriminá-las. Ao usar IA, o *prompt* deve solicitar explicitamente explicações concorrentes e pedir que o modelo não selecione uma delas como conclusão final sem fundamento. Esse procedimento aproxima a IA de uma função heurística, capaz de ampliar o repertório interpretativo, mas não de fechar o argumento.

A literatura de métricas responsáveis recomenda que indicadores sejam lidos em relação aos objetivos de avaliação, às missões institucionais e às diferenças de campo. Esse princípio é decisivo na passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos. Uma mesma taxa de colaboração internacional pode ter significados distintos em física de partículas, saúde coletiva, direito, educação ou ciências agrárias. A IA pode listar fatores contextuais relevantes para cada domínio, mas a pertinência dessa lista precisa ser verificada por especialistas. Assim, a interpretação não se reduz ao reconhecimento automático de padrões, mas combina conhecimento estatístico, conhecimento bibliográfico e compreensão substantiva da área. (HICKS *et al.*, 2015; DE RIJCKE *et al.*, 2016).

Outro cuidado diz respeito à temporalidade. Resultados cientométricos carregam marcas do tempo de publicação, indexação, citação e difusão. Uma publicação recente pode ser invisível em métricas de citação, ainda que seja relevante para uma frente de pesquisa. Um artigo antigo pode concentrar muitas citações por ser método, revisão, base de dados ou objeto de controvérsia. Em séries de publicação, distribuições de citação, mapas de rede, *clusters* temáticos e perfis institucionais, a IA pode sugerir leituras temporais, mas deve ser instruída a distinguir maturidade bibliográfica de relevância epistemológica. A interpretação deve perguntar se o indicador favorece literatura consolidada, literatura emergente, redes antigas ou modismos de curto prazo.

A base de dados é também parte da interpretação. Scopus, Web of Science, PubMed, Dimensions, OpenAlex e repositórios institucionais diferem em cobertura disciplinar, regional, linguística, documental e citacional. Qualquer interpretação sobre séries de publicação, distribuições de citação, mapas de rede, *clusters* temáticos e perfis institucionais deve ser lida à luz dessas diferenças. A IA pode apoiar a construção de uma seção de limitações, mas não deve inventar características da base nem atribuir cobertura inexistente. O *prompt* precisa informar a fonte usada e solicitar que a resposta trate a cobertura como condição analítica. Isso

é especialmente importante em áreas com forte presença de livros, capítulos, literatura cinzenta ou produção local.

A normalização constitui uma mediação interpretativa. Indicadores de citação, colaboração e produção podem mudar significativamente quando ajustados por área, ano, tipo documental ou número de autores. Em a passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos, comparar valores brutos pode ser útil para descrever volume, mas inadequado para inferir influência relativa. A IA pode elaborar explicações didáticas sobre a diferença entre indicadores absolutos e normalizados, ajudando a comunicar resultados a gestores e públicos não especialistas. Entretanto, a escolha da normalização deve ser metodológica, documentada e consistente com a pergunta de pesquisa, não uma decisão estética tomada após observar o resultado. Indicadores normalizados por fonte, área ou período devem ser relatados com seus pressupostos (WALTMAN; VAN ECK, 2013).

A visualização não substitui a interpretação. Mapas, linhas do tempo, diagramas Sankey e *dashboards* facilitam a percepção de padrões, mas também podem sugerir proximidades, hierarquias ou causalidades que não decorrem diretamente dos dados. Quando a análise envolve séries de publicação, distribuições de citação, mapas de rede, *clusters* temáticos e perfis institucionais, a IA pode ajudar a redigir legendas analíticas e alertas de leitura, indicando que distância visual, tamanho de nó ou espessura de ligação dependem do algoritmo e da normalização escolhida. A interpretação deve explicar como a visualização foi construída antes de extrair conclusões substantivas de sua aparência. A interpretação de mapas bibliométricos depende de medidas, algoritmos e parâmetros de visualização (VAN ECK; WALTMAN; DEKKER; VAN DEN BERG, 2010; WALTMAN; VAN ECK; NOYONS, 2010).

A interpretação cientométrica deve evitar o que se pode chamar de "narrativa retrospectiva da inevitabilidade". Depois que um mapa mostra *clusters*, é tentador contar uma história em que cada agrupamento parece natural, coerente e previsível. Entretanto, *clusters* são produtos de decisões de dados, medidas de similaridade, limiares e algoritmos. A IA tende a produzir narrativas coesas e, por isso, precisa ser contida por *prompts* que solicitem incertezas, rótulos alternativos e riscos de interpretação. Em a passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos, a boa narrativa não é a que elimina ambivalências, mas a que mostra como os dados sustentam certas leituras e não outras.

O uso de IA deve ser documentado como parte do método interpretativo. Isso inclui o modelo utilizado, a data de uso, os *prompts* principais, os dados fornecidos, as instruções de restrição, as versões da saída aproveitadas e as decisões humanas de edição. Em a passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos, a documentação dos *prompts* torna a interpretação reprodutível em sentido fraco: outro pesquisador pode não obter exatamente a mesma redação, mas poderá entender o procedimento, avaliar suas premissas e refazer a análise crítica. A ausência dessa documentação transforma a IA em caixa-preta retórica, dificultando a revisão por pares.

Um procedimento útil consiste em solicitar à IA três produtos separados: uma descrição neutra, uma interpretação cautelosa e uma lista de hipóteses que exigem evidência adicional. Essa divisão é particularmente adequada quando há risco de confundir padrão observado com explicação causal ou transformar uma regularidade estatística em diagnóstico substantivo. A descrição deve conter apenas padrões observáveis. A interpretação cautelosa deve relacionar os padrões a literatura, contexto e decisões metodológicas. As hipóteses devem declarar o que ainda precisa ser testado. Essa arquitetura textual ajuda a impedir que o relatório apresente como conclusão aquilo que é apenas possibilidade analítica.

A validação humana não é etapa final decorativa; ela atravessa todo o processo. Especialistas devem verificar se os termos usados para descrever uma área fazem sentido, se os rótulos de *clusters* são reconhecíveis, se os indicadores escolhidos são adequados, se os autores citados são efetivamente centrais e se as conclusões respeitam práticas do campo. A IA pode organizar fichas de validação e sintetizar comentários de especialistas, mas não possui responsabilidade científica. Em a passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos, o pesquisador continua responsável por cada afirmação substantiva, mesmo quando a redação foi inicialmente gerada por um modelo.

O controle da linguagem é parte do rigor. Termos como "liderança", "excelência", "relevância", "inovação", "impacto", "maturidade" e "declínio" carregam implicações avaliativas. Usá-los sem critério pode produzir interpretações mais fortes que os dados. A IA, por ser treinada para produzir textos fluentes, tende a empregar vocabulário valorizador. Por isso, *prompts* interpretativos devem solicitar linguagem modalizada: "sugere", "é compatível com", "pode indicar", "não permite concluir", "exige confirmação". Em a passagem da descrição objetiva dos dados para explicações prudentes em estudos cientométricos, a precisão retórica evita que o relatório converta padrões bibliográficos em julgamentos substantivos indevidos.

A interpretação cientométrica com IA deve incorporar uma camada de auditoria argumentativa. Após redigir a análise, o pesquisador pode pedir ao modelo que identifique afirmações sem evidência, inferências causais indevidas, generalizações excessivas, ausência de limitações e termos avaliativos não justificados. Essa função crítica está no centro de organizar a interpretação em níveis, exigir justificativas explícitas e propor hipóteses alternativas sem apagar incertezas. O resultado da auditoria não deve ser aceito automaticamente, mas serve como lista de verificação. A qualidade da interpretação aumenta quando a IA é usada não apenas para produzir texto, mas para tensionar a coerência entre dados, método e conclusão.

A distinção entre descrição e explicação pode ser compreendida por meio de uma sequência analítica. Primeiro, identifica-se o que foi observado: por exemplo, o número de publicações cresceu, a densidade da rede aumentou ou certos termos passaram a ocupar posição central. Em seguida, pergunta-se que operação produziu essa observação: busca, deduplicação, recorte temporal, normalização, algoritmo de *cluster*ização ou filtro de frequência. Só depois se avalia o que esse padrão pode significar. A IA deve ser mantida dentro dessa sequência, pois, sem esse controle, tende a saltar diretamente para explicações gerais sobre expansão de campos, consolidação teórica ou emergência de temas.

A explicação prudente é especialmente necessária porque estudos cientométricos são geralmente observacionais. Eles analisam registros produzidos por práticas sociais e editoriais, não experimentos controlados. Assim, mesmo quando há associação entre aumento de financiamento e crescimento de publicações, ou entre colaboração internacional e maior citação, a interpretação deve evitar causalidade direta. A IA pode ajudar a redigir enunciados condicionais, mas o pesquisador deve especificar que a evidência mostra compatibilidade, não demonstração causal. Esse cuidado é central para relatórios que serão usados em gestão institucional ou política científica.

Um recurso prático é o uso de níveis de assertividade. A descrição objetiva pode ser escrita em termos fortes: "o conjunto analisado contém 2.340 documentos". A interpretação provável deve ser modalizada: "o crescimento observado sugere maior visibilidade do tema na base analisada". A hipótese deve ser explicitamente marcada: "é possível que políticas de financiamento tenham contribuído para essa expansão, mas essa hipótese exige análise documental". A IA pode estruturar textos nessa gradação quando recebe instrução clara para evitar afirmações monolíticas.

16.3 Como interpretar indicadores em contexto

Interpretar indicadores em contexto significa reconhecer que produção, citação, colaboração, interdisciplinaridade e atenção social medem dimensões diferentes da atividade científica. O número de publicações informa presença documental; as citações indicam circulação bibliográfica; a coautoria expressa uma forma formal de colaboração; e métricas alternativas sinalizam atenção em ambientes distintos. Nenhuma dessas medidas, isoladamente, autoriza juízos automáticos de qualidade, excelência ou relevância social (BORNMANN; DANIEL, 2008; TAHAMTAN; BORNMANN, 2019).

Indicadores absolutos e normalizados devem ser usados para finalidades distintas. Valores brutos descrevem volume e visibilidade no *corpus*; indicadores normalizados procuram reduzir efeitos de área, ano, tipo documental ou fonte. A decisão entre um e outro precisa ser anterior à interpretação e coerente com a pergunta de pesquisa, e não uma escolha posterior para tornar o resultado mais favorável (WALTMAN, 2016; WALTMAN; VAN ECK, 2013).

Indicadores de produção devem ser lidos como medidas de presença documental, não como equivalentes automáticos de contribuição científica. Um programa com muitos artigos pode estar em área de publicação intensiva; outro com poucos livros e capítulos pode ter influência relevante em campos nos quais a produção monográfica é central. A IA pode ajudar a explicar esse ponto a gestores, evitando interpretações simplistas. Contudo, o *prompt* precisa informar o tipo de produção relevante para o campo, pois modelos de linguagem tendem a assumir que artigos indexados são a forma universal de produção científica.

Indicadores de citação exigem ainda mais contextualização. Citações podem expressar uso metodológico, reconhecimento conceitual, crítica, obrigação de campo, autocitação, citação negativa, citação ritual ou simples visibilidade editorial. A literatura sobre comportamento citacional mostra que contagens de citação não medem uma entidade única. Assim, quando a IA interpreta impacto, deve ser instruída a evitar a equivalência "muito citado = alta qualidade". O texto interpretativo deve tratar citação como evidência de circulação e influência bibliográfica, cuja relação com qualidade depende de leitura complementar. Por isso, a interpretação deve apresentar citações como sinais de uso e circulação, e não como prova direta de qualidade científica (AKSNES; LANGFELDT; WOUTERS, 2019; LEYDESDORFF *et al.*, 2016).

Indicadores de colaboração também são ambíguos. Aumento de coautoria pode indicar integração científica, divisão de trabalho, internacionalização, exigências de financiamento, grandes equipamentos, autoria honorária ou práticas editoriais. Uma rede densa não é automaticamente mais inovadora; uma rede dispersa não é necessariamente frágil. A IA pode produzir uma matriz de possíveis significados, mas a decisão interpretativa exige conhecer normas de autoria, escala de projetos e organização institucional. O indicador só ganha sentido quando situado em uma ecologia de práticas científicas. A análise de redes de colaboração deve considerar padrões disciplinares, tamanho de equipes e convenções de autoria (NEWMAN, 2001).

16.4 Como formular hipóteses sem exagerar conclusões

A formulação de hipóteses é necessária quando os indicadores sugerem padrões que ainda não podem ser convertidos em conclusão. Hipóteses cientométricas devem declarar o padrão observado, o mecanismo possível, a evidência adicional necessária e as explicações alternativas. Esse procedimento impede que o texto apresente como diagnóstico aquilo que é apenas uma interpretação plausível.

Em estudos observacionais, a causalidade raramente pode ser afirmada apenas por indicadores bibliográficos. A aproximação entre dois *clusters*, o crescimento de uma série temporal ou o aumento de colaboração internacional podem ser compatíveis com diferentes mecanismos: mudança terminológica, expansão de financiamento, alteração de indexação, amadurecimento cognitivo ou reorganização institucional.

Hipóteses cientométricas devem ser formuladas como proposições investigáveis. Uma boa hipótese não apenas afirma que um padrão "ocorreu", mas indica que mecanismo poderia explicá-lo e que evidência adicional poderia sustentá-lo ou refutá-lo. Por exemplo, diante do crescimento de publicações em uma área, pode-se hipotetizar expansão de financiamento, mudança terminológica, aumento de colaboração internacional ou incorporação do tema por periódicos de alto impacto. A IA pode listar hipóteses, mas o pesquisador deve priorizar aquelas que podem ser confrontadas com dados adicionais.

A prudência conclusiva exige reconhecer a diferença entre hipótese plausível e narrativa sedutora. Um mapa temático pode convidar à leitura de que duas tradições teóricas se aproximaram, mas, sem análise de conteúdo, essa afirmação pode ser exagerada. O correto seria dizer que os termos ou referências associadas a essas tradições passaram a coocorrer com maior frequência. A IA pode ajudar a reescrever conclusões em linguagem mais controlada, substituindo afirmações ontológicas sobre o campo por enunciados sobre os dados observados.

Uma técnica útil é solicitar que a IA atribua grau de evidência a cada conclusão. Não se trata de probabilizar mecanicamente a verdade, mas de classificar a força do suporte empírico: forte quando há múltiplos indicadores convergentes, moderado quando há um padrão consistente, mas dependente de uma única fonte, fraco quando há apenas indício exploratório. Essa classificação torna visível a distância entre resultado e conclusão. O pesquisador pode então revisar o texto, fortalecendo evidências ou rebaixando o nível de assertividade.

A IA pode auxiliar ao listar hipóteses concorrentes e rebaixar o nível de assertividade das conclusões, mas deve ser usada com instruções explícitas de linguagem modalizada. Expressões como "sugere", "é compatível com" e "exige confirmação" são preferíveis a formulações deterministas.

16.5 Triangulação entre métricas, literatura e conhecimento especializado

A triangulação é a estratégia que impede que a interpretação dependa de uma única métrica, de um único mapa ou de uma única síntese gerada por IA. Ela combina evidências quantitativas, leitura documental, literatura especializada e julgamento de domínio para verificar se uma afirmação se sustenta por diferentes caminhos.

A triangulação é uma resposta metodológica à incompletude das métricas. Ela não significa simplesmente acrescentar muitas fontes, mas combinar evidências de naturezas distintas para verificar se uma interpretação se sustenta sob diferentes ângulos. Em um estudo sobre emergência temática, por exemplo, crescimento anual, explosão de termos, presença em periódicos centrais, citações recentes, financiamento e validação por especialistas podem convergir. Quando convergem, a interpretação ganha robustez. Quando divergem, a divergência deve ser analisada e não escondida por uma narrativa produzida pela IA.

A leitura de literatura é indispensável porque indicadores mostram padrões de circulação, mas não necessariamente conteúdo teórico, método, controvérsia ou contribuição. Um *cluster* pode parecer coerente por palavras-chave, mas reunir abordagens incompatíveis. A IA pode sumarizar resumos e organizar tópicos, porém a validação deve incluir leitura de documentos

centrais. Essa leitura permite verificar se o rótulo dado ao *cluster* representa de fato uma escola, uma frente de pesquisa, um método ou apenas uma associação terminológica.

O conhecimento especializado atua como filtro semântico e histórico. Especialistas sabem que certos termos mudaram de significado, que determinadas bases sub-representam periódicos locais, que alguns autores são centrais apesar de pouca citação internacional e que certos temas aparecem com rótulos diferentes em comunidades distintas. A IA pode sistematizar essas informações quando elas são fornecidas, mas não deve substituir a memória sociocognitiva do campo. Por isso, o protocolo de interpretação deve prever consulta ou revisão especializada quando o estudo sustenta decisões relevantes.

Em termos operacionais, a triangulação pode ser registrada em matriz: cada achado quantitativo é confrontado com evidência bibliométrica, evidência textual ou documental, conhecimento especializado necessário, convergências, divergências e decisão interpretativa. Quando as fontes não convergem, a decisão recomendada deve ser "não concluir ainda".

16.6 O papel da IA como "interlocutor crítico"

A IA pode ser mais útil como interlocutor crítico do que como redatora obediente. Nesse papel, o modelo é solicitado a questionar afirmações, identificar lacunas de evidência, apontar causalidade indevida, propor explicações alternativas e simular objeções de revisão por pares. Esse uso desloca a IA da produção automática de conclusões para a auditoria argumentativa.

A noção de interlocutor crítico desloca a IA da posição de redatora obediente para a posição de questionadora metodológica. Em vez de pedir apenas "interprete os resultados", o pesquisador pode solicitar: "critique minha interpretação, identifique saltos causais, aponte dados ausentes e proponha explicações alternativas". Esse uso explora a capacidade dos modelos de linguagem de reorganizar argumentos e reconhecer padrões retóricos. O ganho não está em obter uma verdade, mas em ampliar a revisão crítica antes de submeter o texto a pares ou gestores.

A IA como interlocutor crítico é particularmente útil para detectar assimetrias de argumentação. Um relatório pode enfatizar indicadores favoráveis e minimizar limitações; pode usar linguagem mais forte para achados positivos e mais vaga para achados negativos; pode citar mapas quando eles confirmam a hipótese e ignorá-los quando a complicam. *Prompts* de auditoria podem pedir que o modelo identifique seletividade, inconsistência, ausência de contraprovas e uso desigual de evidências. Essa função aproxima a IA de uma revisão interna de qualidade.

No entanto, a crítica gerada por IA pode ser genérica. Modelos frequentemente produzem listas padronizadas de limitações: viés de base, recorte temporal, idioma, autocitação, normalização e cobertura. Embora úteis, essas listas podem não dialogar com o estudo real. Para evitar crítica superficial, o *prompt* deve incluir pergunta de pesquisa, base usada, período, indicadores, principais resultados e decisões metodológicas. Quanto mais situada for a entrada, mais específica tende a ser a crítica. A responsabilidade final por incorporá-la, porém, continua humana.

A crítica gerada por IA deve ser situada. Para evitar respostas genéricas, o pesquisador precisa fornecer a pergunta, a base utilizada, o período, os indicadores, os principais resultados, as decisões metodológicas e a finalidade do relatório. Quanto mais precisa for a entrada, maior será a utilidade da crítica.

Mesmo como interlocutor crítico, a IA não decide a validade da interpretação. Ela amplia a lista de perguntas e fragilidades, mas o pesquisador continua responsável por verificar dados, referências, leitura de literatura e pertinência substantiva.

16.7 Redação interpretativa com transparência metodológica

A redação interpretativa é parte do método, porque define como a força e os limites dos achados serão compreendidos pelo leitor. Um texto cientométrico transparente deve separar descrição, interpretação, hipótese e recomendação, além de declarar base, período, recorte, procedimentos de limpeza e participação da IA.

A redação transparente começa antes da primeira frase de resultado. Ela depende de um inventário de decisões: onde os dados foram coletados, qual *string* de busca foi usada, que filtros foram aplicados, como duplicatas foram tratadas, como autores e instituições foram normalizados, que indicadores foram calculados, que software foi usado e como a IA participou da interpretação. A seção de resultados deve remeter a essas decisões quando elas condicionam a leitura. A IA pode transformar esse inventário em texto fluente, mas não deve omitir detalhes para tornar a narrativa mais elegante.

Uma boa seção interpretativa evita tanto o excesso de tecnicismo quanto a simplificação abusiva. Para públicos especializados, pode discutir normalização, medidas de similaridade e limitações de *cluster*ização. Para gestores, pode traduzir esses aspectos em implicações: que resultados são robustos, que achados exigem cautela, que decisões não deveriam ser tomadas apenas com esses dados. A IA pode adaptar o nível de linguagem, mas o pesquisador deve revisar se a adaptação preserva o sentido técnico. Comunicação clara não significa apagar incertezas.

A transparência metodológica inclui declarar o uso de IA. Em contextos científicos, relatórios e artigos devem informar se a IA foi usada para sumarização, classificação, revisão crítica ou redação. Também devem indicar que saídas foram revisadas por humanos. Essa declaração não diminui o rigor; ao contrário, aumenta a rastreabilidade. A interpretação cientométrica assistida por IA deve ser apresentada como procedimento híbrido, no qual modelos apoiam organização textual e reflexão metodológica, enquanto pesquisadores assumem responsabilidade pelas conclusões e referências.

Na prática, a IA pode transformar tabelas, mapas e matrizes em texto claro, mas o *prompt* deve proibir extrapolações. Toda conclusão precisa permanecer vinculada aos dados informados, às referências efetivamente utilizadas e às limitações metodológicas do estudo.

A declaração de uso de IA deve informar se o modelo apoiou sumarização, classificação, revisão crítica, tradução ou redação. Também deve indicar que as saídas foram revisadas por humanos. Essa transparência fortalece a rastreabilidade e a responsabilidade autoral.

16.8 Prompts do capítulo: interpretação, auditoria e redação

Os *prompts* a seguir foram desenhados para funcionar como instrumentos de pesquisa. Eles não devem ser usados como comandos isolados, mas como componentes de um protocolo interpretativo. Em todos os casos, recomenda-se fornecer à IA os dados efetivamente analisados, a fonte, o período, a estratégia de busca, as decisões de limpeza e a finalidade do relatório. Quando esses elementos são omitidos, a IA tende a preencher lacunas com generalizações plausíveis, mas não necessariamente verdadeiras para o estudo. Os *prompts* também devem ser acompanhados de revisão humana, pois a qualidade formal da resposta não garante adequação metodológica.

Prompt 16.8.1 Prompt central: interpretação em três níveis

A partir dos resultados abaixo, escreva uma interpretação cientométrica em três níveis:

(1) descrição objetiva dos dados;

(2) interpretação provável, indicando o que os padrões sugerem sem afirmar causalidade;

(3) hipóteses que exigem evidência adicional.

Para cada afirmação interpretativa, indique: evidência usada, grau de incerteza, possível viés de base de dados, possível viés de recorte temporal e risco de extrapolação.

Não use termos como excelência, liderança, declínio ou maturidade sem justificar empiricamente.

Prompt 16.8.2 Prompt para interpretar indicadores em contexto

Atue como especialista em cientometria responsável. Receba os indicadores abaixo e explique seu significado para um público de gestores de pesquisa.

Para cada indicador, apresente:

- definição;
- unidade de análise;
- o que ele permite inferior;
- o que ele não permite inferior;
- condições de comparação;
- riscos de uso indevido e uma formulação prudente para relatório.

Não trate citação como sinônimo automático de qualidade.

Prompt 16.8.3 Prompt para formular hipóteses sem exagerar conclusões

Com base nos achados cientométricos abaixo, gere até cinco hipóteses interpretativas. Para cada hipótese, apresente:

- padrão que a sustenta;
- mecanismo explicativo plausível;
- evidência adicional necessária;
- explicação alternativa e nível de confiança.

Use linguagem modalizada e indique quando os dados são apenas compatíveis com a hipótese, sem demonstrá-la.

Prompt 16.8.4 Prompt para triangulação

Construa uma matriz de triangulação para a interpretação abaixo.

Use as colunas: achado quantitativo, evidência bibliométrica, evidência textual ou documental, conhecimento especializado necessário, convergência entre fontes, divergências observadas e decisão interpretativa recomendada.

Caso a evidência seja insuficiente, escreva explicitamente "não concluir ainda".

Prompt 16.8.5 Prompt para IA como interlocutor crítico

Leia a interpretação cientométrica abaixo como revisor metodológico cético.

Identifique afirmações sem evidência, causalidades indevidas, generalizações excessivas, termos avaliativos não justificados, ausência de limitações, vieses de base de dados e alternativas interpretativas ignoradas.

Em seguida, reescreva a conclusão em tom mais prudente.

Prompt 16.8.6 Prompt para redação transparente

Transforme os resultados abaixo em uma seção de discussão cientométrica.

Estruture o texto em:

- síntese dos achados;
- interpretação contextualizada;
- comparação com literatura, implicações;
- limitações metodológicas e recomendações cautelosas.

Declare claramente que as conclusões se restringem à base, ao período, à estratégia de busca e aos procedimentos de limpeza informados.

16.9 Matriz prática de auditoria interpretativa

A interpretação cientométrica assistida por IA pode ser submetida a uma matriz de auditoria antes de ser incorporada a relatórios, artigos ou pareceres. A matriz abaixo transforma princípios de prudência metodológica em perguntas verificáveis. Seu uso é simples: para cada afirmação relevante, o pesquisador identifica a evidência, classifica a força do suporte, registra limitações e decide se a afirmação deve ser mantida, enfraquecida, removida ou transformada em hipótese. A IA pode preencher uma versão preliminar da matriz, mas a decisão final deve ser humana.

Quadro 16.2 - Matriz prática de auditoria interpretativa

Dimensão auditada	Pergunta de controle	Sinal de alerta	Ação recomendada
Evidência	Qual dado sustenta a afirmação?	Afirmação sem tabela, indicador ou mapa associado.	Reescrever ou remover.

Dimensão auditada	Pergunta de controle	Sinal de alerta	Ação recomendada
Unidade de análise	A conclusão corresponde à unidade analisada?	Inferir sobre qualidade de autores a partir de periódicos.	Reformular no nível correto.
Causalidade	Há evidência causal ou apenas associação?	Uso de "causou", "determinou", "explica" sem desenho causal.	Trocar por linguagem modalizada.
Contexto de campo	A comparação considera diferenças disciplinares?	Comparar áreas com práticas citacionais distintas.	Normalizar ou contextualizar.
Base de dados	A cobertura é adequada ao fenômeno?	Omitir literatura local, livros ou conferências relevantes.	Declarar limitação e triangular.
IA	O uso de IA foi documentado?	Saída textual sem *prompt*, versão ou revisão humana.	Registrar procedimento e auditoria.

Fonte: Elaborado pelo autor (2026).

16.10 Estudo de caso ilustrativo

Considere um estudo hipotético sobre a produção científica em "*IA generativa* na educação superior" entre 2018 e 2026. A busca foi realizada em uma base internacional, com termos em inglês e português, e resultou em 1.250 registros após deduplicação. Observou-se crescimento acentuado de publicações a partir de 2023, aumento de colaborações internacionais e emergência de *clusters* relacionados a avaliação automatizada, integridade acadêmica, personalização da aprendizagem e formação docente. A interpretação apressada poderia afirmar que o campo se consolidou globalmente e que determinadas instituições lideram a inovação educacional. Uma interpretação prudente, entretanto, começaria reconhecendo que o crescimento coincide com a difusão pública de modelos generativos e com a ampliação de vocabulário indexado.

Nesse exemplo, o aumento de publicações é um achado robusto dentro da base, mas sua explicação permanece aberta. Pode refletir expansão real de pesquisa, resposta editorial a tema emergente, criação de chamadas especiais, mudança nos termos de busca ou maior indexação de anais de conferências. A IA pode ajudar listando essas possibilidades, mas não deve escolher uma delas sem análise adicional. A conclusão adequada seria afirmar que os dados sugerem intensificação bibliográfica do tema, compatível com sua emergência recente, mas que a consolidação científica exigiria análise de continuidade temporal, diversidade metodológica, estabilidade conceitual e incorporação em currículos e políticas institucionais.

A interpretação dos *clusters* também deve ser cuidadosa. Um *cluster* de "integridade acadêmica" pode reunir estudos sobre plágio, detecção de texto gerado por IA, ética discente e políticas institucionais, mas não necessariamente representar uma escola teórica. O rótulo do *cluster* é uma hipótese semântica. A IA pode propor rótulos alternativos, como "avaliação da autoria", "ética do uso estudantil" ou "governança acadêmica da IA". A validação por especialistas deve verificar qual rótulo melhor representa documentos centrais. Sem essa etapa, o relatório corre o risco de estabilizar artificialmente uma nomenclatura ainda em disputa.

Por fim, recomendações institucionais devem decorrer de triangulação. Se os dados mostram crescimento do tema e baixa participação de determinada universidade, pode-se sugerir monitoramento estratégico ou formação de grupos de pesquisa, mas não concluir automaticamente atraso institucional. Talvez a instituição publique em bases não indexadas, atue por projetos aplicados, produza políticas internas ou utilize terminologia diferente. A IA pode redigir recomendações alternativas condicionadas a cenários: se o objetivo for visibilidade internacional, ampliar publicação indexada; se o objetivo for impacto pedagógico local, mapear

práticas e políticas internas; se o objetivo for colaboração, identificar redes temáticas com afinidade substantiva.

16.11 Checklist de redação interpretativa

Quadro 16.3 - Checklist de redação interpretative para estudos cientométricos assistidos por IA

Nº	Item de verificação	Por que é importante	Como atender	Verificado
01	A pergunta interpretativa está explicitamente retomada antes da interpretação dos resultados.	Mantém o foco do leitor e vincula os achados ao objetivo do estudo.	Reapresente a pergunta em uma frase clara antes de iniciar cada seção interpretativa.	☐
02	A base de dados, o período, a estratégia de busca e os critérios de inclusão são mencionados quando afetam a leitura dos achados.	Diferentes escolhas ampliam ou reduzem temas, atores e métricas.	Informe base, período, *string*, campos, filtros e critérios sempre que influenciarem a interpretação.	☐
03	Indicadores absolutos e normalizados são distinguidos, evitando comparações inadequadas.	Evita conclusões enviesadas entre áreas, períodos e tipos documentais distintos.	Declare o tipo de indicador e a normalização utilizada. Justifique as comparações realizadas.	☐
04	Citações são tratadas como sinais de circulação e uso, não como prova automática de qualidade.	Citações refletem visibilidade e uso, não necessariamente mérito ou rigor.	Explique o que a métrica de citação indica e o que ela não indica.	☐
05	Mapas e *clusters* são interpretados como produtos de escolhas metodológicas, não como retratos naturais do campo.	Resultados dependem de dados, parâmetros e algoritmos: não são entidades naturais.	Declare métodos, parâmetros e limiares. Evite linguagem que reifique *clusters*.	☐
06	A redação separa descrição objetiva, interpretação provável, hipótese e recomendação.	Clarifica o nível de evidência e evita saltos interpretativos.	Use marcadores como: (a) descrição; (b) interpretação; (c) hipótese; (d) recomendação.	☐
07	Termos avaliativos fortes são substituídos por linguagem mais prudente.	Evita exageros, vieses e afirmações não sustentadas pelos dados.	Prefira “sugere”, “indica”, “é compatível com”. Justifique termos fortes quando usados.	☐
08	Explicações alternativas são mencionadas quando os dados não permitem discriminação conclusiva.	Mostra rigor analítico e evita interpretações unívocas e deterministas.	Apresente pelo menos uma explicação alternativa plausível para os achados.	☐
09	A triangulação com literatura e conhecimento especializado é indicada quando a interpretação sustenta decisões relevantes.	Fortalece a validade, reduz vieses e incorpora contexto teórico e empírico.	Cite literatura relevante, teorias e evidências externas que apoiem ou contrariem sua interpretação.	☐
10	O uso de IA, os *prompts* principais e a validação humana são documentados.	Garante transparência, reprodutibilidade e responsabilidade metodológica.	Registre ferramentas, versões, *prompts*-chave, etapas assistidas por IA e validações realizadas por quem e quando.	☐

Lembrete final
Interpretar é explicar padrões com base em evidências, métodos e contexto — não é afirmar verdades absolutas. Seja claro sobre o que os dados permitem concluir e sobre o que permanece incerto.

Dica de uso
Utilize este checklist antes da submissão de artigos, de relatórios institucionais ou de apresentações, para garantir rigor e clareza na comunicação dos seus achados.

Fonte: Elaborado pelo autor (2026).

16.12 Conclusão

A interpretação cientométrica é o momento em que os resultados quantitativos e relacionais se transformam em argumentos sobre a ciência. Esse processo exige mais do que competência técnica; exige prudência epistemológica, conhecimento contextual, clareza sobre unidades de análise e consciência dos limites dos indicadores. A inteligência artificial pode ampliar a capacidade de organizar e revisar interpretações, mas também pode intensificar o risco de excesso de confiança. Textos gerados por IA costumam parecer coerentes mesmo quando os dados disponíveis não sustentam suas conclusões. Por isso, a principal competência do pesquisador não é apenas saber pedir uma interpretação, mas saber restringi-la, auditá-la e validá-la.

O capítulo defendeu que a interpretação cientométrica com IA deve ser estruturada em camadas: descrição objetiva, interpretação provável, hipótese dependente de evidência adicional e recomendação condicionada. Essa arquitetura impede que números sejam convertidos diretamente em julgamentos de qualidade, que mapas sejam tratados como entidades naturais e que tendências sejam apresentadas como previsões seguras. A IA é útil quando ajuda a explicitar incertezas, comparar explicações, localizar lacunas de evidência, redigir em linguagem clara e revisar argumentos. Ela é inadequada quando substitui leitura especializada, validação de dados, compreensão de campo ou responsabilidade autoral.

A contribuição metodológica central deste capítulo é tratar *prompts* como instrumentos de interpretação responsável. Um *prompt* cientométrico bem desenhado não pede apenas uma resposta elegante; ele define níveis de evidência, exige distinção entre dado e inferência, solicita alternativas, obriga a declarar limitações e orienta a redação transparente. Dessa forma, a IA deixa de operar como máquina de conclusões e passa a funcionar como ambiente de reflexão assistida. A cientometria com IA será tanto mais rigorosa quanto mais tornar visíveis as mediações que sustentam seus enunciados: dados, métricas, métodos, contexto, literatura, especialistas e decisões humanas.

Capítulo 17

Escrita de relatórios, artigos e pareceres cientométricos

17.1 Introdução

A escrita de relatórios, artigos e pareceres cientométricos constitui uma etapa decisiva da pesquisa sobre a ciência. Depois da formulação da pergunta, da coleta dos dados, do pré-processamento, da escolha dos indicadores, da construção de mapas, da análise temporal e da interpretação contextualizada, é no documento final que as decisões metodológicas se tornam visíveis, que os resultados ganham inteligibilidade e que as conclusões se submetem ao escrutínio de leitores especializados e não especializados. Em estudos assistidos por inteligência artificial, essa etapa adquire uma relevância adicional: o texto precisa comunicar não apenas o que foi encontrado, mas também como a IA foi utilizada, quais *prompts* orientaram as tarefas, quais saídas foram aceitas ou rejeitadas, quais validações humanas foram executadas e quais limites permanecem abertos.

A escrita cientométrica não deve ser compreendida como simples ornamentação posterior da análise. Ela é parte constitutiva do método. Um estudo com dados corretos, indicadores pertinentes e visualizações tecnicamente adequadas pode produzir efeitos equivocados se for comunicado com ambiguidade, extrapolação, jargão indevido ou ocultamento de limitações. Do mesmo modo, um relatório dirigido a gestores pode induzir decisões inadequadas se converter correlações em causalidades, rankings em juízos absolutos de qualidade ou mapas exploratórios em diagnósticos definitivos. Por essa razão, escrever sobre ciência exige disciplina argumentativa: cada afirmação deve estar ancorada em evidência, cada inferência deve ser marcada como interpretação, cada recomendação deve declarar suas condições de validade e cada visualização deve ser acompanhada de legenda analítica suficiente.

No contexto deste livro, a inteligência artificial é tratada como apoio à produção de documentos cientométricos, não como autora autônoma de conclusões. Modelos de linguagem podem auxiliar na organização de seções, na transformação de tabelas em narrativas, na tradução de resultados técnicos para linguagem executiva, na revisão de coerência argumentativa, na identificação de lacunas metodológicas e na adaptação do texto para públicos distintos. Entretanto, tais usos exigem controle epistemológico. A IA tende a produzir textos fluentes mesmo quando os dados são insuficientes, pode apagar nuances metodológicas, pode sugerir causas sem evidência causal e pode criar uma aparência de autoridade textual que não corresponde à robustez analítica do estudo. Por isso, os *prompts* de escrita devem conter restrições explícitas: não inventar dados, não extrapolar resultados, separar descrição de interpretação, informar incertezas, preservar o vocabulário técnico quando necessário e adaptar a linguagem sem deformar o conteúdo. Essa cautela é coerente com a literatura sobre modelos de linguagem, métodos de *prompting* e riscos de alucinação em geração de linguagem natural (BROWN *et al.*, 2020; LIU *et al.*, 2023; JI *et al.*, 2023).

A literatura sobre relato científico e transparência metodológica reforça a importância de descrever com clareza desenho, fontes, procedimentos e limitações. A estrutura *IMRAD*, amplamente difundida na comunicação científica, organiza artigos em introdução, métodos, resultados e discussão, facilitando a localização de informações essenciais pelo leitor. Diretrizes como *PRISMA* e *PRISMA-S*, embora voltadas a revisões sistemáticas,

oferecem princípios transferíveis para estudos cientométricos, sobretudo no que se refere à explicitação de estratégias de busca, fontes consultadas, critérios de seleção e fluxo de exclusões. Em bibliometria e cientometria, diretrizes e boas práticas recentes têm enfatizado a necessidade de relatar bases utilizadas, data da coleta, campos exportados, procedimentos de limpeza, escolhas de normalização, parâmetros de rede e limitações dos indicadores. A escrita, portanto, funciona como dispositivo de reprodutibilidade e responsabilidade. A estrutura *IMRAD* e as diretrizes de relato reforçam essa exigência de transparência, especialmente quando articuladas a recomendações de documentação de busca e seleção de estudos (SOLLACI; PEREIRA, 2004; MOHER *et al.*, 2009; PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021; SIMERA *et al.*, 2010).

Este capítulo examina a escrita de relatórios, artigos e pareceres cientométricos em seis dimensões: a estrutura do relatório cientométrico; a escrita da metodologia; a escrita dos resultados; a escrita da discussão; a produção de sumários executivos; e a comunicação para gestores, pesquisadores e formuladores de políticas. Em cada dimensão, são discutidas funções analíticas da escrita, riscos recorrentes e possibilidades de uso de *prompts*. O objetivo é oferecer ao leitor um repertório de princípios e modelos que permitam transformar resultados cientométricos em documentos claros, verificáveis, prudentes e úteis para diferentes contextos de decisão.

17.2 Estrutura de um relatório cientométrico

Um relatório cientométrico é um gênero híbrido. Ele combina elementos de artigo científico, relatório técnico, diagnóstico institucional, produto de inteligência estratégica e documento de apoio à decisão. Sua estrutura deve refletir essa natureza. Quando destinado à publicação acadêmica, tende a aproximar-se da organização *IMRAD*, com introdução, metodologia, resultados, discussão e conclusão. Quando produzido para uma agência, pró-reitoria, instituto, programa de pós-graduação ou unidade de pesquisa, necessita também de sumário executivo, mensagens-chave, recomendações, quadros de risco e anexos metodológicos. Em ambos os casos, a arquitetura do documento deve orientar o leitor desde o problema de pesquisa até as implicações práticas, sem ocultar as incertezas do processo.

A primeira função da estrutura é estabelecer rastreabilidade. O leitor deve conseguir reconstruir o percurso do estudo: qual pergunta foi formulada, que universo documental foi considerado, quais bases foram consultadas, que filtros temporais e temáticos foram aplicados, quais campos bibliográficos foram utilizados, quais indicadores foram calculados, quais redes foram geradas e como se chegou às interpretações apresentadas. Essa rastreabilidade não depende apenas de anexos técnicos; ela deve atravessar o documento inteiro. Uma boa seção introdutória antecipa a pergunta e o escopo. Uma boa metodologia registra decisões. Uma boa seção de resultados evita esconder exceções. Uma boa discussão explicita até onde os dados permitem ir.

A segunda função da estrutura é modular a leitura. Nem todos os públicos leem o relatório da mesma forma. Gestores podem iniciar pelo sumário executivo e pelas recomendações. Pesquisadores tendem a examinar metodologia, parâmetros e tabelas. Avaliadores institucionais podem buscar indicadores comparativos. Bibliotecários e analistas de informação podem concentrar-se em bases, estratégias de busca e qualidade dos metadados. Um relatório robusto deve permitir múltiplas trajetórias de leitura sem perder coerência. Isso justifica a presença de quadros de síntese, glossários, notas metodológicas, apêndices e seções de limitações.

A terceira função da estrutura é distinguir camadas argumentativas. Em estudos cientométricos, é comum misturar resultados numéricos, visualizações, interpretações

e recomendações no mesmo parágrafo. Essa prática aumenta o risco de extrapolação. Uma estrutura mais rigorosa separa, ainda que de modo articulado, descrição dos dados, interpretação provável, hipótese explicativa, implicação prática e recomendação. Por exemplo, dizer que uma instituição aumentou sua produção em determinado período é uma descrição; sugerir que esse aumento decorre de política institucional é hipótese; recomendar manutenção de investimentos é decisão que exige evidência adicional. A estrutura do relatório deve tornar essas diferenças visíveis.

O sumário executivo, quando presente, não deve ser uma versão superficial do estudo. Ele deve condensar a questão, o método, os principais achados, as limitações e as implicações de modo fiel. Deve evitar frases genéricas como "a área está crescendo rapidamente" sem indicar base temporal, universo documental e comparação de referência. Deve também indicar quando um resultado depende de base específica, como Scopus, Web *of Science*, OpenAlex, *Lens* ou Dimensions. Em contextos de gestão, a precisão do sumário é crucial, pois muitas decisões serão tomadas por leitores que não examinarão integralmente os anexos técnicos.

A introdução deve cumprir quatro tarefas. Primeiro, apresentar o problema que justifica o estudo. Segundo, delimitar o objeto: área, tema, instituição, país, período ou conjunto documental. Terceiro, explicitar a pergunta ou objetivo cientométrico. Quarto, posicionar o estudo em relação à literatura metodológica e ao contexto de aplicação. Uma introdução frágil costuma tratar o tema como se a relevância fosse autoevidente. Uma introdução forte demonstra porque determinado mapeamento, indicador ou diagnóstico é necessário naquele momento e para aquele público.

A seção metodológica é o núcleo de credibilidade. Deve informar fontes, data da coleta, estratégia de busca, critérios de inclusão e exclusão, formatos de exportação, campos utilizados, procedimentos de limpeza, tratamento de duplicatas, padronização de nomes, escolha de indicadores, normalizações, parâmetros de visualização e softwares empregados. Quando houver uso de IA, é necessário registrar tarefas realizadas, modelos utilizados, *prompts* centrais, versões quando disponíveis, critérios de validação e decisões humanas. A omissão desses elementos compromete a reprodutibilidade e pode transformar o relatório em narrativa não auditável.

A seção de resultados deve apresentar evidências de forma ordenada. Em geral, recomenda-se começar por uma visão descritiva do *corpus*: número de documentos, período, tipos documentais, fontes, idiomas e distribuição temporal. Depois, podem ser apresentados indicadores de produção, impacto, colaboração, redes, temas, internacionalização e tendências. A ordem deve acompanhar a lógica da pergunta, e não apenas a sequência de saídas do software. Um erro recorrente é escrever o relatório como inventário de gráficos: primeiro o gráfico de publicações, depois o mapa de coautoria, depois a nuvem de palavras, sem uma linha argumentativa que una os achados.

A discussão interpreta os resultados em contexto. Ela deve comparar achados com literatura, explicar possíveis padrões, considerar vieses de base, avaliar limitações e formular implicações. Em cientometria, a discussão é particularmente delicada porque indicadores são facilmente convertidos em juízos de qualidade. Uma seção de discussão responsável evita afirmar que mais publicações significam automaticamente melhor ciência ou que mais citações significam maior relevância social. Ela trabalha com indicadores como sinais parciais, dependentes de campos disciplinares, práticas de citação, cobertura das bases, idioma, idade dos documentos e natureza da produção científica.

A conclusão deve ser proporcional aos resultados. Deve responder à pergunta do estudo sem reabrir discussões extensas. Deve também indicar limitações e próximos passos.

Conclusões como "o tema é promissor" ou "a instituição é líder" exigem critérios explícitos. Melhor seria afirmar, por exemplo, que "no *corpus* recuperado na base X, no período Y, a instituição Z apresentou maior número de publicações entre as instituições analisadas, embora a interpretação de liderança dependa de indicadores adicionais de impacto, colaboração e qualidade". A precisão protege o documento contra usos indevidos.

Os anexos devem preservar a auditabilidade. Neles podem constar *strings* de busca completas, logs de coleta, lista de campos exportados, versões de softwares, parâmetros de rede, tabelas suplementares, *prompts* utilizados, amostras revisadas manualmente e decisões de normalização. Em estudos assistidos por IA, anexar *prompts* e critérios de validação é uma prática recomendável. Não se trata de expor toda interação conversacional, mas de registrar os comandos metodologicamente relevantes que influenciaram classificação, sumarização, rotulagem ou redação.

A IA pode ajudar a planejar a estrutura do relatório, mas o *prompt* deve ser exigente. Um comando genérico, como "escreva um relatório bibliométrico", tende a produzir textos padronizados e pouco auditáveis. Um *prompt* adequado especifica objetivo, público, tipo de estudo, base de dados, indicadores, limitações e formato esperado. Também deve exigir separação entre achados, interpretações e recomendações. A função da IA nesse ponto é sugerir arquitetura textual e verificar lacunas, não substituir o desenho metodológico. Em tarefas de escrita e revisão, técnicas de *prompting* e raciocínio assistido devem ser tratadas como apoio metodológico, não como garantia automática de validade (WEI *et al.*, 2022; WANG *et al.*, 2023; LIU *et al.*, 2023).

17.3 Escrita da metodologia

A metodologia é a seção em que o estudo cientométrico se torna verificável. Ela deve permitir que outro pesquisador compreenda as decisões adotadas e, dentro das condições de acesso às bases, reproduza ou audite o percurso analítico. Em estudos bibliométricos, a metodologia precisa ser mais detalhada do que em textos descritivos convencionais, pois pequenas diferenças de busca, data de coleta, tipo documental, idioma, campo pesquisado ou critério de exclusão podem alterar substancialmente os resultados. Em estudos assistidos por IA, essa exigência aumenta, porque é necessário documentar também como os modelos foram utilizados e validados.

Uma metodologia cientométrica começa pela delimitação do objeto. Deve indicar se o estudo analisa um tema, uma área, uma instituição, uma revista, um programa, um país, um grupo de autores ou uma tecnologia. Também deve esclarecer o período analisado e justificar esse recorte. Um período de dez anos pode ser adequado para observar evolução temática, mas insuficiente para áreas consolidadas ou excessivo para temas muito recentes. A justificativa temporal não deve ser mero formalismo; ela condiciona a interpretação de crescimento, impacto, maturidade e emergência.

A descrição das fontes de dados é obrigatória. Bases comerciais e abertas diferem em cobertura, curadoria, atualização, campos disponíveis, normalização de afiliações, indexação de referências e acesso a citações. Assim, não basta afirmar que os dados foram coletados em "uma base bibliográfica". Deve-se informar nome da base, data da coleta, modalidade de acesso, campos exportados, filtros utilizados e eventuais limitações. Quando mais de uma fonte for combinada, é necessário explicitar como se tratou a duplicação, a harmonização de metadados e a divergência entre identificadores.

A estratégia de busca deve ser apresentada de modo completo. Em relatórios técnicos, recomenda-se incluir a *string* principal no corpo do texto e versões completas em apêndice.

A estratégia deve indicar campos pesquisados, operadores booleanos, truncamentos, proximidade quando utilizada, filtros por ano, idioma, tipo documental e área. Em bases diferentes, a mesma lógica conceitual pode exigir sintaxes diferentes. Por isso, comparar resultados de Scopus, Web *of Science*, PubMed, Dimensions e OpenAlex exige relatar adaptações, não apenas repetir a pergunta de pesquisa.

Os critérios de inclusão e exclusão precisam ser precisos. Termos como "documentos relevantes" são insuficientes se não houver critérios operacionais. Em um estudo cientométrico, incluir apenas artigos e revisões pode ser adequado para medir produção consolidada, mas pode excluir anais de conferências importantes em ciência da computação ou engenharia. Excluir documentos sem resumo pode melhorar análise textual, mas pode distorcer períodos antigos ou bases com metadados incompletos. A metodologia deve informar não apenas o que foi excluído, mas porque a exclusão foi metodologicamente aceitável.

O pré-processamento deve ser descrito como etapa analítica. Padronizar autores, desambiguar instituições, tratar duplicatas, normalizar palavras-chave e agrupar sinônimos são decisões que influenciam indicadores e mapas. A metodologia deve informar critérios, ferramentas, regras de decisão e amostras verificadas. Se a IA foi usada para sugerir agrupamentos terminológicos, o texto deve declarar que a sugestão foi revisada por humanos, informar como os casos ambíguos foram tratados e preservar versões dos termos originais. A limpeza de dados não é uma operação neutra; ela produz o *corpus* analisável.

A escolha de indicadores deve ser justificada pela pergunta de pesquisa. Indicadores de produção respondem a perguntas sobre volume e evolução; indicadores de citação respondem parcialmente a perguntas sobre reconhecimento acadêmico; redes de colaboração respondem a perguntas relacionais; métricas alternativas podem indicar atenção social ou circulação digital. A metodologia deve evitar listas longas de indicadores sem função analítica. Cada métrica utilizada deve ter propósito, fórmula ou descrição, unidade de análise, fonte dos dados e limitação reconhecida. Essa vinculação entre indicador, pergunta e limitação é destacada em orientações metodológicas para análises bibliométricas e em estudos sobre comportamento citacional (DONTHU *et al.*, 2021; BORNMANN; DANIEL, 2008; BORNMANN; MARX, 2013).

A descrição de softwares e parâmetros é essencial. Ferramentas como bibliometrix, VOSviewer, CiteSpace, *SciMAT*, *Gephi*, *Pajek* ou scripts próprios podem gerar resultados diferentes a partir dos mesmos dados, porque usam parâmetros de normalização, limiares e algoritmos distintos. Ao relatar mapas de rede, deve-se indicar tipo de rede, unidade de nó, definição de aresta, método de normalização, limiar mínimo, algoritmo de layout, método de *cluster*ização e critérios de rotulagem. A ausência desses detalhes reduz a capacidade de auditoria e dificulta comparação entre estudos. A documentação desses parâmetros é particularmente relevante em ferramentas de mapeamento e análise bibliométrica como bibliometrix, VOSviewer, CiteSpace e outras soluções de *science mapping* (ARIA; CUCCURULLO, 2017; VAN ECK; WALTMAN, 2010; CHEN, 2006; COBO *et al.*, 2011).

Quando a metodologia envolve IA generativa, a transparência deve ser ampliada. É recomendável informar quais tarefas foram realizadas com apoio de IA: expansão de termos de busca, revisão de *strings*, normalização de palavras-chave, classificação de resumos, rotulagem de *clusters*, redação preliminar, revisão de legibilidade ou geração de sumário executivo. Para cada tarefa, devem ser registrados *prompt*, entrada fornecida, saída esperada, critérios de aceitação, revisão humana e eventuais amostras de controle. O modelo utilizado e sua versão devem ser informados quando disponíveis, reconhecendo que modelos podem mudar ao longo do tempo.

A metodologia também deve explicitar como se lidou com incerteza. Em classificação assistida por IA, por exemplo, é prudente exigir grau de confiança, justificativa textual e marcação de casos duvidosos. Depois, uma amostra deve ser revisada por especialistas ou por dupla codificação. Em rotulagem de *clusters*, rótulos sugeridos devem ser comparados com palavras-chave, títulos de documentos centrais e conhecimento de domínio. Em redação interpretativa, frases geradas pela IA devem ser conferidas contra os resultados. A seção metodológica deve relatar esses mecanismos de controle.

A escrita metodológica deve ser precisa, mas não hermética. O leitor técnico precisa de parâmetros; o leitor gestor precisa entender a confiabilidade do processo. Uma boa estratégia é combinar descrição narrativa com quadros de síntese. Um quadro metodológico pode conter fonte, período, busca, filtros, *corpus* final, campos exportados, limpeza, indicadores, ferramentas, uso de IA e validação. Essa combinação torna a metodologia acessível sem sacrificar rigor.

A IA pode ajudar a redigir a metodologia a partir de um protocolo estruturado. Contudo, o *prompt* deve instruir o modelo a não preencher lacunas com suposições. Deve também pedir que marque campos ausentes. Por exemplo: se o pesquisador não informou a data de coleta, a IA deve escrever "[informar data de coleta]" e não inventar uma data plausível. Esse tipo de restrição transforma a IA em assistente de documentação, não em geradora de ficções metodológicas.

17.4 Escrita dos resultados

A seção de resultados deve apresentar os achados de forma clara, ordenada e proporcional à evidência. Em estudos cientométricos, os resultados frequentemente incluem tabelas extensas, séries temporais, rankings, mapas de rede, diagramas temáticos, indicadores normalizados e descrições textuais. O desafio não é apenas mostrar muitos dados, mas construir uma narrativa que responda à pergunta de pesquisa. A escrita de resultados deve selecionar, hierarquizar e conectar evidências, evitando tanto a enumeração mecânica quanto a interpretação excessiva.

Uma prática recomendável é iniciar a seção de resultados pela caracterização do *corpus*. O leitor precisa saber quantos documentos foram analisados, quais tipos documentais predominam, qual período foi coberto, quais idiomas aparecem, quais bases contribuíram para o conjunto e como o *corpus* se distribui no tempo. Essa descrição funciona como contexto para todos os resultados posteriores. Se o *corpus* é pequeno, mapas de rede podem ser instáveis. Se o *corpus* é dominado por uma disciplina ou país, generalizações devem ser cautelosas. Se houve crescimento recente, é necessário considerar efeitos de indexação, políticas editoriais e expansão da base.

A escrita de resultados deve evitar juízos avaliativos prematuros. Em vez de afirmar que "o pesquisador A é o mais importante da área", o texto deve dizer que "o pesquisador A apresentou o maior número de documentos no *corpus* analisado" ou "recebeu o maior número de citações brutas no período". A importância científica é uma categoria interpretativa mais ampla, que pode envolver qualidade, originalidade, influência, impacto social e reconhecimento por pares. Resultados devem ser apresentados na linguagem exata do indicador utilizado.

Tabelas e gráficos exigem texto analítico. Um erro comum consiste em inserir uma tabela e apenas repetir seus números. A função do texto é destacar padrões, contrastes, inflexões, concentração, dispersão, mudanças de posição e exceções. Uma tabela de países, por exemplo, pode revelar concentração geográfica, assimetria entre produção e citação,

presença de países emergentes ou dependência de colaboração internacional. O texto deve orientar o olhar do leitor, mas sem introduzir explicações ainda não sustentadas.

Séries temporais devem ser narradas com atenção a escalas. Crescimento absoluto pode parecer expressivo em áreas pequenas, enquanto crescimento percentual pode exagerar variações de bases reduzidas. O texto deve indicar valores iniciais e finais, mudanças de tendência, anos de inflexão e possíveis efeitos de cobertura. Quando o estudo usa IA para interpretar séries temporais, o *prompt* deve pedir distinção entre padrão observado, interpretação plausível e evidência necessária. A IA não deve concluir que uma política causou crescimento sem dados sobre a política e sem desenho analítico adequado. A leitura de crescimento da literatura deve considerar escalas e séries temporais para evitar inferências desproporcionais sobre expansão científica (BORNMANN; MUTZ, 2015).

Rankings devem ser tratados com cautela. Listas de autores, instituições, países ou periódicos podem ser úteis para identificação de atores proeminentes, mas também podem induzir leituras competitivas simplificadoras. A escrita deve informar se o ranking é por número de documentos, citações, média de citações, colaboração, centralidade ou outro indicador. Deve também indicar empates, diferenças pequenas e sensibilidade a filtros. Em relatórios institucionais, é recomendável evitar linguagem celebratória ou punitiva e preferir formulações analíticas.

Mapas de rede exigem uma escrita diferente da escrita de tabelas. Em redes de coautoria, o texto deve descrever densidade, componentes, grupos, atores centrais e padrões de conexão. Em redes de cocitação, deve discutir estrutura intelectual. Em redes de coocorrência de palavras-chave, pode descrever agrupamentos temáticos. Entretanto, a escrita deve lembrar que proximidade visual não é necessariamente proximidade epistemológica absoluta; depende de algoritmo, normalização, limiar e layout. A legenda analítica deve informar essas condições.

A escrita de resultados também deve registrar ausências significativas. Às vezes, a falta de colaboração entre instituições, a baixa presença de países periféricos, a ausência de determinados termos ou a inexistência de conexão entre subcampos é tão relevante quanto os elementos presentes. A IA pode ajudar a procurar lacunas em tabelas e mapas, mas deve ser instruída a basear observações nos dados fornecidos. Um bom *prompt* solicita que o modelo indique padrões observáveis e também perguntas que os resultados não respondem.

Ao transformar tabela em narrativa, o pesquisador deve evitar três erros: explicar demais, interpretar de menos e perder rastreabilidade. Explicar demais significa atribuir causas sem análise causal. Interpretar de menos significa deixar o leitor sozinho diante de números. Perder rastreabilidade significa escrever afirmações sem indicar qual tabela, gráfico ou indicador as sustenta. Uma narrativa de resultados deve manter ligação clara com as evidências. Expressões como "conforme a Tabela 2", "o mapa de coocorrência sugere" e "no período observado" ajudam a preservar essa ligação.

A IA é especialmente útil para gerar versões alternativas da redação dos resultados. Pode produzir uma versão técnica, uma versão sintética, uma versão executiva e uma versão didática. Porém, todas devem ser verificadas contra os dados. Recomenda-se fornecer ao modelo tabelas ou resumos estruturados, não apenas imagens de gráficos, e pedir que cada frase interpretativa cite a evidência que a sustenta. A saída deve ser revisada para corrigir exageros, omissões e inconsistências terminológicas.

Em artigos científicos, a seção de resultados deve resistir à tentação de discutir todos os achados. A seleção deve obedecer à pergunta central. Resultados suplementares podem ir para apêndices ou material complementar. Em relatórios técnicos, é possível incluir mais

detalhes, mas ainda assim convém criar hierarquia: mensagens-chave, evidências principais, detalhamento e anexos. A escrita deve ser desenhada para evitar que o leitor confunda volume de informação com força de argumento.

Por fim, resultados não devem ocultar fragilidades. Se uma análise de citações exclui publicações recentes por baixa janela de citação, isso deve ser informado. Se a rede de colaboração é sensível ao limiar de corte, isso deve ser mencionado. Se a normalização de palavras-chave agrupou termos ambíguos, deve-se indicar revisão humana. A transparência sobre limites não enfraquece o relatório; ao contrário, aumenta sua credibilidade.

17.5 Escrita da discussão

A discussão é a seção em que os resultados são colocados em contexto. Enquanto a seção de resultados responde ao que foi observado, a discussão pergunta o que esses achados podem significar, em que condições são plausíveis, como se relacionam com literatura anterior, que limitações afetam a interpretação e quais implicações podem ser derivadas. Em cientometria, a discussão é particularmente importante porque os dados bibliográficos são sinais indiretos de processos científicos. Publicações, citações, colaborações e palavras-chave não são a própria ciência, mas rastros documentais de atividades científicas.

Uma discussão responsável começa pela retomada da pergunta. O leitor deve perceber como os principais achados respondem ao objetivo do estudo. Se a pergunta é sobre evolução temática, a discussão deve concentrar-se em mudanças de temas, continuidade, emergência e declínio. Se a pergunta é sobre colaboração institucional, a discussão deve tratar padrões relacionais, centralidade e possível dependência de poucos atores. Se a pergunta é sobre impacto, a discussão deve considerar janelas de citação, diferenças disciplinares e normalização. A discussão perde força quando se dispersa em comentários genéricos sobre todos os gráficos.

A interpretação de indicadores exige contextualização. Um aumento de publicações pode indicar expansão da área, mas também pode refletir mudança de estratégia de busca, crescimento geral da literatura, indexação de novos periódicos ou aumento de documentos de conferência. Um aumento de citações pode indicar reconhecimento, mas também envelhecimento do *corpus*, autocitação, concentração em poucos artigos ou práticas disciplinares intensivas em citação. A discussão deve examinar alternativas e não apenas confirmar a narrativa desejada.

A formulação de hipóteses é legítima, desde que marcada como tal. Estudos cientométricos muitas vezes revelam padrões que sugerem explicações futuras: a emergência de um tema pode estar associada a políticas públicas, avanços tecnológicos, crises sanitárias ou financiamento internacional; a centralidade de uma instituição pode decorrer de infraestrutura, tradição disciplinar ou liderança de pesquisadores-chave. No entanto, a discussão deve diferenciar hipótese plausível de conclusão demonstrada. Um bom texto usa expressões como "pode sugerir", "é compatível com", "exige investigação adicional" e "não permite afirmar causalidade".

A triangulação é um princípio central da discussão cientométrica. Resultados métricos devem ser confrontados com literatura substantiva, conhecimento de especialistas, documentos de política científica, dados institucionais e informações qualitativas. Uma rede de coautoria pode revelar colaboração formal, mas entrevistas ou documentos institucionais podem explicar sua origem. Uma tendência temática pode ser confirmada pela literatura de revisão. Um indicador de impacto pode ser relativizado por avaliação qualitativa de contribuições. A IA pode ajudar a estruturar essa triangulação, mas não deve substituir a consulta a fontes.

A discussão deve lidar explicitamente com vieses. Bases de dados têm cobertura desigual por idioma, região, área, tipo documental e período. O uso exclusivo de bases internacionais pode sub-representar produção local, livros, capítulos, periódicos nacionais e documentos técnicos. A análise de citações pode favorecer áreas com maior densidade citacional e publicações em inglês. A altmetria pode privilegiar temas midiáticos ou comunidades digitais ativas. A discussão deve reconhecer essas assimetrias e impedir que resultados parciais sejam apresentados como retrato total da ciência. Estudos comparativos de cobertura e viés linguístico em bases bibliográficas mostram que esses efeitos precisam ser explicitados na interpretação (VAN LEEUWEN *et al.*, 2001; MARTÍN-MARTÍN *et al.*, 2021).

Quando a IA participa da interpretação, o risco de fluência enganosa aumenta. Modelos de linguagem tendem a produzir explicações coerentes mesmo sem evidência suficiente. Por isso, o *prompt* de discussão deve exigir três níveis: descrição objetiva, interpretação provável e hipótese que exige evidência adicional. Esse arranjo obriga o texto a declarar seus limites. Além disso, deve pedir uma lista de possíveis vieses e uma revisão de frases que pareçam causais. A IA pode funcionar como interlocutor crítico quando é instruída a questionar a força das conclusões, e não apenas a embelezá-las. Essa precaução dialoga com a literatura sobre alucinação e confiabilidade de textos gerados por modelos de linguagem (JI *et al.*, 2023).

A discussão de relatórios técnicos deve conectar resultados a decisões possíveis, mas sem ultrapassar a evidência. Por exemplo, um relatório institucional pode sugerir fortalecer colaboração internacional se os dados mostram baixa conectividade externa. Entretanto, a recomendação deve considerar objetivos institucionais, recursos disponíveis, áreas estratégicas e riscos de dependência. A discussão deve transformar resultados em implicações condicionais: "caso o objetivo seja aumentar internacionalização, os dados sugerem..." Essa forma evita prescrição universal.

Em artigos acadêmicos, a discussão deve dialogar com o estado da arte metodológico. É importante explicar por que determinadas escolhas analíticas são adequadas e como os achados se comparam a estudos semelhantes. Se o estudo usa IA para classificação temática, deve discutir confiabilidade, validação e limites. Se usa mapas de coocorrência, deve mencionar que *clusters* são representações dependentes de parâmetros. Se usa indicadores de impacto, deve relacionar-se com literatura sobre métricas responsáveis.

A redação da discussão deve evitar linguagem promocional. Expressões como "revolucionário", "sem precedentes", "comprova a excelência" ou "demonstra liderança absoluta" raramente são adequadas em cientometria. A linguagem científica deve ser precisa e proporcional. Isso não significa escrever de forma tímida; significa sustentar cada afirmação com evidência adequada. Uma conclusão forte é aceitável quando o desenho e os dados a sustentam. Uma conclusão forte sem base é fragilidade retórica.

A IA pode ser usada para revisar a discussão em busca de extrapolações. Um *prompt* útil pede que o modelo identifique frases que afirmam causalidade, generalizam além do *corpus*, confundem produção com qualidade, confundem citação com impacto social ou ignoram vieses de base. Essa revisão deve ser tratada como apoio, pois o modelo também pode errar, mas funciona como camada adicional de controle. A versão final deve ser validada pelo autor humano.

A discussão se encerra melhor quando retorna ao problema inicial e indica próximos passos. Em estudos cientométricos, esses próximos passos podem incluir análise qualitativa de documentos centrais, ampliação para outras bases, comparação internacional, entrevistas com especialistas, atualização periódica do *corpus*, validação de *clusters* ou análise de políticas. Dessa forma, a discussão não apenas interpreta resultados passados, mas orienta investigações futuras sem converter prognóstico em certeza.

17.6 Produção de sumários executivos

O sumário executivo é uma peça estratégica em relatórios cientométricos. Ele traduz um estudo complexo em um conjunto de mensagens compreensíveis para tomadores de decisão, sem eliminar o rigor metodológico. Diferentemente de um resumo acadêmico, que costuma apresentar objetivo, método, resultados e conclusão de modo compacto, o sumário executivo deve destacar achados acionáveis, implicações, limites e recomendações. Seu público pode incluir gestores universitários, dirigentes de agências, coordenadores de programas, equipes de planejamento, conselhos científicos e formuladores de políticas.

A principal virtude de um bom sumário executivo é a fidelidade. Ele não deve simplificar a ponto de distorcer. Se o estudo analisou apenas artigos indexados em determinada base, o sumário não pode afirmar que avaliou toda a produção científica da instituição. Se a análise de citações não foi normalizada por área, não deve comparar diretamente campos com comportamentos citacionais distintos. Se a rede de colaboração foi construída com coautoria, não deve representar todas as formas de parceria científica. A clareza executiva depende de precisão metodológica.

Um sumário executivo eficaz pode ser estruturado em seis partes: finalidade do estudo, escopo e método, principais achados, interpretação prudente, implicações para decisão e limitações. Essa estrutura permite que o gestor compreenda rapidamente o que foi feito e o que pode ser feito com os resultados. Também protege contra uso indevido, pois torna explícito que recomendações dependem do *corpus*, do período e da pergunta analisada.

Mensagens-chave devem ser curtas, mas não vagas. Em vez de "a produção cresceu", pode-se escrever: "no *corpus* recuperado, a produção anual aumentou entre 2015 e 2024, com maior aceleração após 2020; a interpretação requer cautela porque a base ampliou a cobertura de periódicos no período". Essa frase comunica padrão, período e limitação. Em vez de "há pouca colaboração", pode-se escrever: "a rede de coautoria apresenta poucos vínculos interinstitucionais fora do núcleo principal, sugerindo oportunidade de diversificação de parcerias".

Sumários executivos também devem distinguir achado e recomendação. Um achado descreve o estado observado. Uma recomendação propõe ação. A passagem de um a outro exige justificativa. Se o achado é concentração da produção em poucos grupos, a recomendação pode ser apoiar formação de novos núcleos, mas isso depende de objetivos institucionais. Se o achado é baixa visibilidade internacional, a recomendação pode envolver colaboração externa, publicação em periódicos internacionais ou participação em redes, mas cada opção tem custos e riscos. O sumário deve evitar automatismo.

A IA pode auxiliar muito na produção de sumários executivos. Pode reduzir extensão, organizar mensagens, adaptar linguagem e propor títulos. Entretanto, deve ser alimentada com resultados verificados e instruída a não criar números ou recomendações não presentes no relatório. Um *prompt* adequado deve exigir que cada mensagem-chave esteja vinculada a uma evidência e que cada recomendação seja formulada como condicional. Também deve pedir a identificação de limitações em linguagem acessível.

A linguagem do sumário executivo deve ser precisa e não técnica em excesso. Termos como modularidade, centralidade de intermediação, normalização de citações ou acoplamento bibliográfico podem ser explicados quando necessários, mas não devem dominar o texto. Uma alternativa é manter termos técnicos acompanhados de explicação breve: "centralidade de intermediação, isto é, a posição de um ator como ponte entre grupos". A adaptação da linguagem não deve apagar a complexidade, mas torná-la inteligível.

Para gestores, é útil apresentar implicações em níveis. O primeiro nível envolve achados consolidados, com forte suporte nos dados. O segundo envolve sinais que exigem monitoramento. O terceiro envolve hipóteses que demandam investigação adicional. Essa organização evita que todos os resultados tenham o mesmo peso. A IA pode ajudar a classificar achados nesses níveis, desde que o pesquisador defina critérios explícitos de força da evidência.

Sumários executivos podem incluir quadros de decisão. Por exemplo, uma matriz pode cruzar achado, evidência, implicação, ação possível, risco e dado adicional necessário. Esse formato é especialmente útil em política científica, pois transforma resultados em insumos para planejamento sem ocultar incertezas. Ele também facilita a prestação de contas, pois deixa claro que uma decisão não decorre automaticamente de uma métrica isolada.

A produção de sumários executivos em estudos assistidos por IA deve declarar o uso de IA quando ela tiver contribuído para redação, síntese ou adaptação de linguagem. Essa declaração pode ser breve, mas deve informar que o conteúdo foi revisado e validado por autores humanos. Em contextos institucionais sensíveis, essa transparência aumenta a confiança no documento e evita a percepção de que recomendações foram terceirizadas a um sistema opaco.

Um risco comum é transformar o sumário executivo em propaganda institucional. Em relatórios internos, há pressão para destacar sucessos e minimizar problemas. A escrita científica, no entanto, deve preservar a função diagnóstica. Um bom sumário pode reconhecer pontos fortes, vulnerabilidades e oportunidades. A credibilidade de uma análise cientométrica depende da disposição de mostrar resultados desconfortáveis quando eles são metodologicamente sustentados.

Outro risco é produzir sumários excessivamente extensos. O sumário executivo deve permitir leitura rápida. Quando muitos detalhes são necessários, eles devem ser deslocados para quadros, anexos ou seções técnicas. A IA pode ser útil para gerar versões em diferentes extensões: 300 palavras, uma página, duas páginas ou apresentação de dez pontos. Mas a versão final deve ser revisada para garantir que nenhuma simplificação comprometa a exatidão

17.7 Comunicação para gestores, pesquisadores e formuladores de políticas

A comunicação cientométrica precisa reconhecer a diversidade de públicos. Gestores buscam informações para decisão; pesquisadores buscam validade metodológica e interpretação substantiva; formuladores de políticas buscam evidências para programas, prioridades e alocação de recursos. O mesmo conjunto de dados pode ser apresentado de formas distintas para esses públicos, desde que a adaptação preserve a integridade dos resultados. A IA pode ajudar nessa adaptação, mas deve ser guiada por critérios explícitos de fidelidade, transparência e adequação de linguagem.

Para pesquisadores, a comunicação deve enfatizar metodologia, limitações, parâmetros e diálogo com a literatura. Esse público tende a questionar escolhas de base, critérios de busca, normalização, indicadores, algoritmos e inferências. Portanto, textos destinados a pesquisadores devem conter detalhe técnico suficiente. Também devem apresentar referências, justificar decisões e expor incertezas. A IA pode ajudar a revisar consistência metodológica e clareza argumentativa, mas não deve suavizar controvérsias metodológicas relevantes.

Para gestores, a comunicação deve enfatizar achados acionáveis, riscos e alternativas. Gestores não precisam conhecer todos os detalhes de normalização de uma rede, mas precisam saber se o diagnóstico é robusto, se há incerteza relevante, que decisões podem ser apoiadas e quais dados adicionais são necessários. A linguagem deve ser direta, com mensagens-chave,

mapas interpretados, quadros de implicação e recomendações condicionais. A comunicação para gestores falha quando se limita a reproduzir gráficos técnicos sem explicar seu significado.

Para formuladores de políticas, a comunicação deve considerar escala, equidade, viabilidade e efeitos sistêmicos. Indicadores podem orientar políticas de financiamento, internacionalização, inovação, formação de redes e avaliação institucional. Contudo, métricas mal comunicadas podem reforçar desigualdades, punir áreas com menor densidade citacional, invisibilizar produção local ou induzir comportamentos estratégicos. A comunicação para políticas deve, portanto, incluir cautelas éticas, efeitos indesejados potenciais e alternativas de monitoramento.

A adaptação de linguagem não significa redução de rigor. Um conceito técnico pode ser traduzido sem perder precisão. "Cocitação" pode ser explicada como "ocorrência de dois trabalhos sendo citados conjuntamente por outros documentos, sinalizando proximidade na estrutura intelectual da área". "Acoplamento bibliográfico" pode ser explicado como "similaridade entre documentos que compartilham referências". "Normalização por área" pode ser descrita como "ajuste para comparar campos com diferentes padrões de citação". A IA pode gerar essas explicações, mas deve ser instruída a preservar o termo técnico acompanhado da explicação.

A comunicação visual é parte da comunicação textual. Gráficos, mapas e *dashboards* devem ser acompanhados de títulos informativos, legendas analíticas e notas metodológicas. Um mapa de rede sem explicação pode ser esteticamente atraente e cognitivamente opaco. Um gráfico de barras sem unidade de análise pode ser enganoso. Um *dashboard* sem definição de indicadores pode induzir decisões equivocadas. A escrita deve explicar o que a visualização mostra, o que não mostra e quais cuidados são necessários.

Na comunicação para públicos não técnicos, é recomendável trabalhar com analogias moderadas. Redes podem ser explicadas como padrões de conexão; *clusters* como grupos de documentos ou atores mais relacionados entre si; indicadores como sinais parciais. Contudo, analogias não devem sugerir propriedades inexistentes. Dizer que um *cluster* é uma "comunidade natural" pode ser enganoso, pois comunidades dependem de algoritmo e parâmetros. É melhor dizer que o *cluster* é "um agrupamento produzido pelo método utilizado, que sugere proximidade temática ou relacional".

A IA pode ser utilizada como simulador de audiência. O pesquisador pode pedir ao modelo que leia um trecho como gestor, como avaliador técnico ou como formulador de política, identificando termos obscuros, inferências fortes demais e lacunas de explicação. Esse uso é produtivo porque transforma a IA em ferramenta de revisão comunicacional. Porém, as sugestões devem ser julgadas pelo autor, pois o modelo pode propor simplificações inadequadas.

A comunicação de incerteza é um desafio central. Muitos relatórios apresentam resultados com aparência de certeza absoluta, enquanto as análises dependem de filtros, bases, janelas temporais e escolhas de normalização. A linguagem deve incluir marcadores de incerteza sem parecer evasiva. Expressões como "no *corpus* analisado", "a evidência sugere", "há indícios", "a interpretação requer cautela" e "os dados não permitem afirmar" são úteis quando usadas de modo consistente. A IA deve ser instruída a manter esses marcadores.

Comunicar para decisão implica também evitar o fetichismo do número. Indicadores são úteis porque condensam informação, mas podem esconder processos, desigualdades e contextos. Um relatório pode mostrar que uma área tem baixo número de publicações internacionais, mas a decisão deve considerar idioma, missão social, financiamento, infraestrutura, relevância local e objetivos estratégicos. A comunicação cientométrica deve convidar à decisão informada, não ao automatismo métrico.

Em contextos de política científica, recomendações devem ser acompanhadas de mecanismos de monitoramento. Se o relatório recomenda ampliar colaboração internacional, deve sugerir indicadores de acompanhamento, como número de coautorias internacionais, diversidade de países parceiros, redes de projetos, mobilidade acadêmica e impacto em formação. Se recomenda apoiar temas emergentes, deve sugerir atualização periódica da busca, análise de tendência e validação por especialistas. A comunicação deve conectar diagnóstico, ação e avaliação futura.

Por fim, a comunicação cientométrica assistida por IA deve ser ética. Deve declarar limitações, evitar manipulação visual, não ocultar resultados desfavoráveis, não produzir autoridade artificial e não atribuir à IA decisões humanas. O texto final deve deixar claro que a responsabilidade interpretativa é dos autores e da instituição que utiliza o relatório. A IA é instrumento de apoio à escrita, revisão e adaptação; a responsabilidade científica permanece humana.

17.8 Prompts do capítulo

Os *prompts* a seguir foram concebidos como instrumentos metodológicos de escrita e revisão. Eles não substituem julgamento humano, mas ajudam a estruturar relatórios, artigos, pareceres e sumários executivos, desde que sejam usados com dados verificados e critérios explícitos de validação.

Prompt 17.8.1 Prompt para redigir metodologia

Atue como especialista em escrita cientométrica.

Redija uma seção de metodologia para um estudo com os seguintes dados: [objetivo], [base de dados], [data da coleta], [string de busca], [período], [tipos documentais], [campos exportados], [procedimentos de limpeza], [indicadores], [softwares], [uso de IA] e [procedimentos de validação].

Não invente informações ausentes; marque lacunas com colchetes.

Separe a metodologia em: desenho do estudo, fonte e coleta, critérios de seleção, pré-processamento, análise, uso de IA, validação e limitações metodológicas.

Prompt 17.8.2 Prompt para transformar tabela em resultado narrativo

Atue como redator científico.

A partir da tabela abaixo, escreva uma subseção de resultados. Faça apenas afirmações sustentadas pelos dados.

Destaque padrões, contrastes, concentração, mudanças temporais e exceções.

Não use termos avaliativos como melhor, pior, líder ou excelência, a menos que o critério esteja definido.

Para cada afirmação interpretativa, indique a evidência correspondente. Ao final, liste cautelas de interpretação.

Prompt 17.8.3 Prompt para gerar sumário executivo

Atue como analista de política científica. Com base nos resultados verificados abaixo, produza um sumário executivo de até [número] palavras para gestores.

Inclua:

- objetivo do estudo;
- escopo;
- método em linguagem acessível;
- cinco mensagens-chave;
- implicações;
- recomendações condicionais e limitações.

Não introduza números que não estejam nos resultados.

Separe achados de recomendações.

Use linguagem clara, formal e não promocional.

Prompt 17.8.4 Prompt para adaptar linguagem para público não técnico

Reescreva o texto abaixo para público não técnico, preservando o conteúdo científico.

Mantenha termos técnicos essenciais, mas explique-os entre parênteses ou em frases curtas.

Não simplifique conclusões além do que os dados permitem.

Preserve expressões de incerteza, limitações e escopo.

Ao final, liste os termos técnicos que foram explicados e as afirmações que exigem cuidado interpretativo.

17.9 Protocolos de escrita assistida por IA

A escrita assistida por IA deve seguir protocolo, não improviso. Um protocolo mínimo inclui definição do produto textual, público-alvo, dados de entrada, restrições, formato de saída, critérios de validação e registro da versão final. Sem esse protocolo, a IA tende a ser usada como ferramenta de redação genérica, produzindo textos fluentes, mas metodologicamente frágeis. Em cientometria, o risco é maior porque as análises combinam números, grafos, mapas e inferências. O protocolo deve impedir que a fluência substitua a evidência.

O primeiro passo é definir o papel da IA. Ela pode atuar como organizadora, redatora preliminar, revisora crítica, tradutora de linguagem, verificadora de coerência ou geradora de alternativas. Cada papel exige *prompt* diferente. Um *prompt* de organização pode pedir estrutura de relatório; um *prompt* de revisão pode pedir identificação de extrapolações; um *prompt* de comunicação pode pedir adaptação para gestores. Misturar todos os papéis em um único comando reduz controle e aumenta o risco de saída inconsistente.

O segundo passo é controlar a entrada. A IA deve receber dados limpos, tabelas verificadas e descrições precisas. Quando recebe resultados incompletos, tende a preencher lacunas com plausibilidade textual. Por isso, recomenda-se fornecer tabelas em formato estruturado, indicar nomes exatos de indicadores e delimitar o escopo. Se o texto deve explicar uma rede de coautoria, o *prompt* deve incluir tipo de rede, número de nós, número de arestas, densidade, comunidades, principais atores e limitações. Quanto mais estruturada a entrada, mais auditável a saída.

O terceiro passo é estabelecer regras de redação. Em relatórios cientométricos, regras úteis incluem: não afirmar causalidade sem evidência causal; não confundir produção com qualidade; não usar citação como único sinônimo de impacto; não generalizar além do *corpus*; mencionar base e período; separar descrição, interpretação e hipótese; preservar incerteza; indicar limitações. Essas regras devem aparecer no *prompt*, não apenas na intenção do pesquisador.

O quarto passo é validar a saída. A validação deve comparar texto gerado com tabelas, gráficos e protocolo metodológico. Frases com números devem ser conferidas. Frases interpretativas devem ser classificadas quanto ao suporte. Recomendações devem ser rastreadas até evidências. Em documentos sensíveis, uma revisão por pares ou especialistas é recomendável. A IA pode auxiliar nessa validação, mas não deve ser a única camada de controle.

O quinto passo é registrar o processo. O relatório final pode incluir uma nota metodológica sobre uso de IA, e os anexos podem preservar *prompts* centrais. Esse registro fortalece transparência e permite que leitores compreendam em que etapas a IA participou. Não é necessário registrar cada interação exploratória, mas os *prompts* que influenciaram classificação, interpretação, redação ou recomendação devem ser documentados.

A reprodutibilidade em escrita assistida por IA é relativa. Modelos podem produzir saídas diferentes para o mesmo *prompt*, e versões podem mudar. Por isso, o registro deve incluir data, modelo quando disponível, parâmetros relevantes e saída validada. Em vez de prometer reprodutibilidade perfeita, o pesquisador deve prometer auditabilidade: tornar visíveis entradas, *prompts*, decisões e validações.

A escrita com IA também exige controle de autoria. O documento final deve refletir julgamento humano. A IA não deve ser citada como autora, pois não assume responsabilidade, não responde por integridade científica e não pode aprovar a versão final. Autores humanos devem revisar, corrigir e responder pelo texto. Essa posição é consistente com orientações editoriais contemporâneas sobre uso de ferramentas de IA em produção científica. A responsabilização humana pela versão final é coerente com princípios de integridade e transparência do relato científico (SIMERA *et al.*, 2010).

Modelos de seções redacionais

Modelo de metodologia:

"Este estudo adotou abordagem cientométrica descritiva e relacional, com o objetivo de analisar [objeto] no período [período]. Os dados foram coletados na base [base] em [data], por meio da estratégia de busca [string], aplicada aos campos [campos]. Foram incluídos [tipos documentais] e excluídos [critérios].

Os registros foram exportados em [formato] com os campos [campos]. O pré-processamento incluiu remoção de duplicatas, padronização de nomes de autores e instituições, normalização de palavras-chave e revisão manual de casos ambíguos. As análises contemplaram indicadores de produção, impacto, colaboração e estrutura temática. O uso de IA ocorreu em [tarefas], com validação humana por [procedimento]."

Modelo de resultados:

"O *corpus* final reuniu [n] documentos publicados entre [ano inicial] e [ano final]. Observou-se [padrão temporal], com [valor] no início do período e [valor] no final. A distribuição por [unidade] indica concentração em [atores], enquanto [grupo] apresentou crescimento recente. A rede de [tipo] revelou [número] comunidades, com maior conectividade em [*cluster*]. Esses resultados descrevem padrões de produção e relação no *corpus* analisado; não permitem, isoladamente, inferir qualidade científica ou causalidade institucional."

Modelo de discussão:

"Os achados sugerem que [interpretação prudente], em consonância com [literatura/contexto]. Contudo, essa interpretação deve ser relativizada por [limitações]. A concentração observada pode refletir [hipótese 1], [hipótese 2] ou [hipótese 3], exigindo dados adicionais para confirmação. Do ponto de vista institucional, os resultados indicam oportunidade de [implicação], desde que considerada a missão da organização e a diversidade de áreas envolvidas. Assim, os indicadores devem ser lidos como sinais complementares à avaliação qualitativa."

Modelo de conclusão:

"O estudo respondeu à pergunta proposta ao demonstrar que [síntese do achado principal] no *corpus* analisado. A contribuição metodológica reside em [contribuição], enquanto a contribuição aplicada consiste em [uso possível]. As principais limitações decorrem de [base], [período], [metadados] e [indicadores]. Recomenda-se que estudos futuros ampliem [fonte/período/método], validem [*clusters*/indicadores] com especialistas e combinem análise cientométrica com evidências qualitativas."

Modelo de parecer:

"Com base na análise realizada, conclui-se que o estudo apresenta [forças], especialmente em [aspectos]. Entretanto, requer atenção em [fragilidades], notadamente [questões metodológicas]. Recomenda-se revisar [seção], explicitar [parâmetro], reduzir [extrapolação] e acrescentar [limitação]. O parecer é favorável à utilização dos resultados para [uso], desde que as recomendações sejam interpretadas como condicionais ao *corpus* e ao período analisado."

Esses modelos não devem ser copiados mecanicamente. Funcionam como estruturas de apoio, a serem preenchidas com dados reais. A IA pode ajudar a adaptar os modelos, mas o autor deve verificar cada afirmação. Em particular, campos entre colchetes devem ser substituídos por informações verificadas; se a informação não existir, a lacuna deve permanecer indicada até que seja resolvida.

17.10 Estudo de caso ilustrativo

Imagine um relatório sobre a produção científica de uma universidade em inteligência artificial aplicada à saúde entre 2014 e 2024. O *corpus* foi coletado em Scopus e OpenAlex, deduplicado por DOI e título, e analisado quanto a produção anual, periódicos, instituições parceiras, países, citações, palavras-chave e redes de coautoria. A IA foi utilizada para revisar a estratégia de busca, sugerir agrupamentos de palavras-chave, gerar rótulos preliminares de *clusters* e produzir versões iniciais do sumário executivo. Todas as saídas foram revisadas por equipe humana.

A metodologia desse estudo deveria informar as duas bases utilizadas, a data de coleta, as *strings* adaptadas a cada base, os filtros aplicados e o procedimento de combinação. Deveria explicar que documentos sem DOI foram deduplicados por similaridade de título e autores, que afiliações foram normalizadas manualmente nos casos de maior frequência e que palavras-chave foram agrupadas por equivalência semântica com revisão de especialistas da área. Também deveria declarar que a IA não classificou documentos de forma autônoma, mas sugeriu agrupamentos revisados.

Na seção de resultados, o texto poderia afirmar que a produção cresceu de forma contínua após 2018, com aceleração entre 2020 e 2022, mas deveria evitar atribuir automaticamente esse crescimento à pandemia sem análise adicional. Poderia indicar que os *clusters* temáticos se concentram em diagnóstico por imagem, prontuários eletrônicos, aprendizado profundo e suporte à decisão clínica. Poderia também observar que a colaboração internacional é concentrada em poucos países parceiros, sugerindo dependência de redes específicas.

Na discussão, o relatório poderia levantar a hipótese de que a expansão do tema se relaciona a investimentos em saúde digital, disponibilidade de dados clínicos e avanço de métodos de aprendizado profundo. Contudo, deveria reconhecer que os dados bibliográficos não demonstram causalidade. Poderia recomendar diversificação de parcerias e fortalecimento de governança de dados, mas sempre como implicações condicionais, não como conclusões inevitáveis.

O sumário executivo apresentaria cinco mensagens-chave: crescimento da produção; concentração temática; centralidade de alguns grupos; oportunidades de colaboração internacional; e necessidade de políticas de dados e ética em IA. Cada mensagem deveria ser

acompanhada de evidência e cautela. Esse estudo de caso mostra como a escrita transforma análise em documento decisório sem abandonar prudência metodológica.

17.11 Checklist de qualidade para escrita cientométrica

Um checklist de qualidade ajuda a revisar o documento final. A primeira pergunta é: a pergunta de pesquisa está explicitada? Sem pergunta clara, o relatório tende a se tornar coleção de gráficos. A segunda é: o *corpus* está descrito? O leitor precisa saber o que foi analisado. A terceira é: as fontes e a data de coleta estão informadas? Bases mudam, e a data é parte do método. A quarta é: a estratégia de busca está registrada? Sem ela, não há auditabilidade.

Outras perguntas verificam a metodologia: os critérios de inclusão e exclusão são claros? Os procedimentos de limpeza foram descritos? Os indicadores têm justificativa? Os parâmetros de rede e visualização foram informados? O uso de IA foi documentado? Houve validação humana? As limitações metodológicas estão no corpo do texto e não apenas no final? Essas perguntas protegem a credibilidade.

Na revisão dos resultados, deve-se perguntar: cada afirmação numérica corresponde a tabela ou gráfico? O texto evita transformar rankings em juízos absolutos? As visualizações têm legendas analíticas? As diferenças pequenas são tratadas com cautela? Os resultados ausentes ou negativos foram considerados? A redação separa resultado de interpretação? Essas perguntas reduzem risco de extrapolação.

Na revisão da discussão, deve-se perguntar: hipóteses estão marcadas como hipóteses? Causalidade é afirmada apenas quando há desenho adequado? Indicadores são interpretados em contexto? A literatura substantiva foi mobilizada? Vieses de base e período foram discutidos? Recomendações são condicionais? O uso de IA não produziu autoridade artificial? Essas perguntas ajudam a preservar responsabilidade epistêmica.

Na revisão de comunicação, deve-se perguntar: o sumário executivo é fiel ao relatório? A linguagem está adequada ao público? Termos técnicos foram explicados sem simplificação indevida? As recomendações são acionáveis? As limitações são compreensíveis para gestores? Há risco de uso punitivo ou para gerar *rank* indevido? A resposta a essas perguntas define a qualidade social do documento, não apenas sua qualidade técnica.

17.12 Escrita de pareceres cientométricos

O parecer cientométrico é um gênero específico de comunicação técnico-científica. Ele pode avaliar a consistência de um estudo já realizado, apoiar decisões institucionais, examinar a pertinência de indicadores, revisar metodologia de um relatório ou interpretar resultados para um comitê. Diferentemente do artigo, que busca contribuir para o debate científico, e do relatório, que apresenta diagnóstico amplo, o parecer precisa formular juízo fundamentado sobre adequação, suficiência, fragilidade e uso possível de evidências. Seu valor depende da clareza dos critérios de avaliação e da capacidade de distinguir problemas metodológicos graves de aspectos aperfeiçoáveis.

Um parecer deve começar com identificação do objeto avaliado: título do relatório ou estudo, objetivo declarado, escopo, fonte de dados, período, indicadores principais e contexto de uso. Em seguida, deve apresentar critérios de análise. Esses critérios podem incluir aderência entre pergunta e método, completude da descrição metodológica, qualidade da estratégia de busca, adequação dos indicadores, consistência dos procedimentos de limpeza,

transparência dos parâmetros de rede, proporcionalidade das interpretações e clareza das limitações. Sem critérios explícitos, o parecer se torna opinião impressionista.

A seção de méritos deve reconhecer contribuições efetivas. Um estudo pode apresentar boa delimitação temática, estratégia de busca bem documentada, *corpus* robusto, visualizações adequadas, transparência no uso de IA ou cautela interpretativa. Registrar méritos é importante porque o parecer não deve ser apenas inventário de falhas. Além disso, reconhecer acertos ajuda os autores a preservar partes fortes ao revisar o documento. A IA pode auxiliar a organizar méritos em categorias, mas a avaliação substantiva deve ser humana.

A seção de fragilidades deve ser precisa e acionável. Em vez de escrever "a metodologia está fraca", o parecer deve indicar que a data de coleta não foi informada, que os campos pesquisados na *string* não foram especificados, que a deduplicação não foi descrita, que os mapas não apresentam parâmetros de normalização ou que as conclusões extrapolam os dados. Uma crítica útil informa o problema, explica sua consequência e sugere correção. Essa estrutura aumenta a utilidade do parecer e reduz conflito interpretativo.

Pareceres sobre relatórios assistidos por IA devem incluir avaliação do uso da IA. É necessário verificar se o documento declara tarefas realizadas, se *prompts* relevantes foram preservados, se houve validação humana e se a IA foi usada para gerar afirmações substantivas sem evidência. Um parecer pode recomendar que o relatório inclua anexo com *prompts* metodológicos, matriz de validação de classificações ou descrição de amostragem revisada. Esse tipo de recomendação contribui para governança de IA em pesquisa.

A conclusão do parecer deve indicar grau de confiabilidade e usos recomendados. Por exemplo: "o relatório é adequado para diagnóstico exploratório, mas não para ranqueamento de desempenho individual"; "os resultados podem orientar discussão estratégica, desde que complementados por avaliação qualitativa"; ou "as conclusões sobre impacto devem ser revistas porque as citações não foram normalizadas por área e idade". Essa formulação evita respostas binárias quando o uso apropriado é condicional.

Um *prompt* para parecer deve instruir a IA a agir como revisora crítica, não como revisora complacente. Deve pedir análise por critérios, identificação de lacunas, frases potencialmente exageradas e recomendações de correção. Também deve proibir a invenção de dados e exigir que toda crítica seja associada a trecho ou ausência observável no documento. A IA pode ajudar a detectar problemas de coerência, mas não substitui conhecimento especializado sobre cientometria e contexto institucional.

Em contextos de gestão, pareceres podem ter consequências relevantes. Uma avaliação negativa pode impedir uso de relatório em decisão; uma avaliação positiva pode legitimar política institucional. Por isso, o tom deve ser formal, equilibrado e rastreável. A escrita deve evitar ironia, adjetivação excessiva e julgamento pessoal. A credibilidade do parecer decorre da argumentação, não da autoridade retórica do parecerista.

17.13 Erros frequentes na escrita cientométrica

O primeiro erro frequente é confundir descrição com avaliação. Muitos textos afirmam que autores, instituições ou países são "melhores" porque aparecem com maior número de documentos ou citações. Essa passagem é metodologicamente indevida quando não há critérios de qualidade, normalização ou avaliação qualitativa. A redação correta deve nomear o indicador: maior produção, maior número de citações, maior centralidade ou maior presença no *corpus*. O juízo avaliativo, quando necessário, deve ser construído com múltiplas evidências.

O segundo erro é transformar correlação em causalidade. Uma série temporal pode mostrar crescimento após determinada política, mas isso não prova que a política causou o crescimento. Uma rede pode mostrar proximidade entre grupos, mas isso não prova cooperação substantiva além da coautoria. Um *cluster* temático pode emergir após evento social, mas isso não prova relação direta. A escrita deve usar linguagem causal apenas quando o desenho analítico permite tal inferência.

O terceiro erro é ocultar o *corpus*. Relatórios frequentemente apresentam resultados como se fossem universais, mas o leitor não sabe quantos documentos foram analisados, quais bases foram usadas ou quais filtros foram aplicados. Esse erro compromete a interpretação. Toda seção de resultados deve lembrar, quando necessário, que os achados se referem ao *corpus* analisado. Essa prática é especialmente importante em sumários executivos, nos quais a compressão textual aumenta risco de generalização indevida.

O quarto erro é escrever metodologia como lista de softwares. Dizer que foram usados VOSviewer, bibliometrix ou CiteSpace não basta. É preciso informar dados, filtros, unidades de análise, normalização, limiares, algoritmos e decisões de limpeza. Software não é método; software implementa operações dentro de um desenho metodológico. A escrita deve tornar esse desenho explícito.

O quinto erro é sobrecarregar o texto com jargão. Termos técnicos são necessários, mas devem ser usados com função. Em documentos para gestores, é inadequado apresentar modularidade, centralidade, densidade e acoplamento sem explicação. Em artigos especializados, a ausência de termos técnicos também pode empobrecer a precisão. O desafio é calibrar linguagem ao público. A IA pode ajudar a produzir versões em diferentes níveis de tecnicidade, desde que o conteúdo seja preservado.

O sexto erro é apresentar visualizações sem interpretação. Mapas de rede, diagramas Sankey, mapas temáticos e *dashboards* podem impressionar visualmente, mas exigem leitura guiada. O texto deve explicar o que a visualização mostra, quais parâmetros a condicionam e que inferências são aceitáveis. Uma visualização bonita, porém sem legenda analítica, pode ser metodologicamente frágil e de comunicação perigosa.

O sétimo erro é omitir limitações por receio de enfraquecer o documento. Na verdade, limitações bem formuladas fortalecem a credibilidade. Elas mostram que os autores conhecem alcance e fronteiras do método. Uma limitação não precisa invalidar o estudo; precisa orientar uso adequado. A escrita deve apresentar limitações específicas, não apenas frases genéricas como "há limitações inerentes ao método".

O oitavo erro é delegar conclusões à IA. Quando modelos generativos redigem resultados e discussão sem controle, podem suavizar incertezas, inventar conexões e produzir argumentação excessivamente coerente. O autor deve verificar cada frase, especialmente aquelas que explicam causas, atribuem liderança, sugerem recomendações ou citam literatura. A IA pode acelerar a escrita, mas não reduz responsabilidade autoral.

17.14 Governança editorial e versões do documento

A produção de um relatório ou artigo cientométrico raramente ocorre em uma única versão. Há versões de trabalho, versões para revisão interna, versões para gestores, versões para submissão acadêmica e versões finais. Cada versão deve ter finalidade clara. Uma versão técnica pode conter detalhes extensos; uma versão executiva pode condensar achados; uma versão pública pode excluir dados sensíveis; uma versão acadêmica pode aprofundar discussão metodológica. A gestão dessas versões é parte da governança editorial.

Em estudos assistidos por IA, a governança de versões deve registrar mudanças introduzidas por ferramentas generativas. Se a IA reescreveu o sumário executivo, revisou a discussão ou adaptou linguagem para público não técnico, a equipe deve revisar se a alteração preservou dados, incertezas e limitações. Um controle simples consiste em manter uma matriz com versão, seção, tipo de intervenção, responsável humano, ferramenta usada e validação realizada. Essa matriz pode não entrar no documento final, mas deve compor o arquivo do projeto.

A revisão por pares internos é recomendável antes da entrega de relatórios estratégicos. Um revisor metodológico pode verificar coerência dos indicadores; um especialista de domínio pode avaliar interpretação temática; um gestor pode verificar utilidade; um comunicador pode revisar clareza. A IA pode funcionar como camada adicional, mas não substitui essa diversidade de olhares. A qualidade final melhora quando diferentes competências se encontram.

A governança editorial também envolve gestão de sensibilidade. Relatórios podem mencionar desempenho de indivíduos, grupos ou unidades. Nesses casos, a escrita deve evitar exposição indevida, rankings simplistas e conclusões punitivas. Quando possível, recomenda-se apresentar resultados agregados, contextualizados por área e missão. A IA deve ser instruída a não personalizar julgamentos nem formular recomendações que possam ser usadas de forma disciplinar sem análise qualitativa.

A versão final deve preservar anexos metodológicos suficientes. Mesmo quando o relatório principal é enxuto, as evidências de reprodutibilidade devem estar disponíveis: *strings*, datas, bases, filtros, procedimentos de limpeza, parâmetros, *prompts* centrais e validações. Essa documentação permite auditoria futura e facilita atualização do estudo. Sem ela, cada nova análise começa do zero e aumenta risco de inconsistência.

A escrita final deve passar por revisão de consistência terminológica. Termos como impacto, influência, visibilidade, produção, colaboração e excelência não são intercambiáveis. Um documento que usa "impacto" ora como citação, ora como efeito social, ora como qualidade, torna-se ambíguo. A IA pode ajudar a localizar termos e sugerir padronização, mas a decisão conceitual deve ser humana. Um glossário pode ser útil em relatórios extensos.

A publicação ou circulação de relatórios cientométricos deve considerar atualidade dos dados. Bases são atualizadas, citações mudam e documentos recentes acumulam impacto ao longo do tempo. Assim, relatórios devem informar data de validade ou recomendação de atualização. Em temas emergentes, análises podem envelhecer rapidamente. A escrita deve evitar conclusões definitivas quando o *corpus* representa um momento específico.

Por fim, governança editorial significa reconhecer que documentos cientométricos são artefatos sociotécnicos. Eles resultam de dados, métodos, softwares, *prompts*, decisões humanas, interesses institucionais e públicos de leitura. A boa escrita torna essas mediações visíveis. Ao fazer isso, contribui para uma cultura de avaliação mais responsável, transparente e reflexiva.

17.15 Ética da comunicação, responsabilização e usos posteriores

A ética da comunicação cientométrica começa antes da redação final, mas se materializa no texto. O modo como resultados são apresentados pode orientar decisões de financiamento, contratação, promoção, cooperação e reputação institucional. Por isso, a escrita não é neutra. Ela pode ampliar compreensão, mas também pode reforçar hierarquias, invisibilizar áreas menos indexadas, induzir competição inadequada e naturalizar desigualdades. Um capítulo sobre escrita deve, portanto, reconhecer que relatórios e pareceres não apenas descrevem a ciência; eles participam dos processos pelos quais a ciência é governada.

A responsabilização exige autoria clara. Quando um documento é produzido com apoio de IA, os autores humanos devem assumir responsabilidade por conteúdo, referências, dados, interpretações e recomendações. Ferramentas generativas não devem ser usadas como escudo para erros. Se uma tabela foi interpretada de modo equivocado, se uma limitação foi omitida ou se uma recomendação foi exagerada, a responsabilidade é da equipe que validou e entregou o texto. Essa posição deve orientar práticas editoriais e institucionais.

A ética da comunicação também exige cuidado com comparações. Comparar instituições de áreas diferentes, pesquisadores em estágios distintos de carreira ou programas com missões diversas pode produzir injustiça avaliativa. A redação deve indicar quando comparações são exploratórias, quando exigem normalização e quando não são recomendáveis. Em certos casos, a conclusão mais responsável é não comparar diretamente. A IA, quando solicitada a gerar rankings, pode ignorar esses limites; por isso, *prompts* de comunicação devem incluir critérios de justiça comparativa.

Outro aspecto ético é a comunicação de dados incompletos. Bases bibliográficas podem sub-representar produção em português, livros, capítulos, relatórios técnicos, inovação social, dados de pesquisa e atividades de extensão. Quando o relatório é destinado a uma instituição com forte missão regional ou social, essa sub-representação pode ter efeito político. A escrita deve alertar que ausência em base internacional não equivale a ausência de contribuição científica ou social. Esse cuidado é fundamental para evitar decisões baseadas em visibilidade indexada como se fosse totalidade da produção.

A transparência sobre limitações deve ser proporcional ao risco de uso. Se o relatório será usado apenas para planejamento exploratório, limitações podem ser apresentadas de modo sintético. Se será usado para avaliação de desempenho, distribuição de recursos ou formulação de política, as limitações devem ser mais explícitas e próximas das recomendações. Não basta colocá-las em apêndice. O leitor decisor deve encontrá-las no mesmo espaço em que encontra os achados que embasam decisões.

A escrita deve também orientar usos posteriores do documento. Um relatório pode recomendar atualização anual, validação por especialistas, integração com dados qualitativos, ampliação de fontes ou criação de painel de acompanhamento. Essas orientações ajudam a impedir que o relatório seja tratado como fotografia definitiva. Em cientometria, dados mudam, áreas evoluem e políticas alteram comportamentos. O documento deve explicitar que sua validade é situada no tempo e no método.

A IA pode apoiar a ética comunicacional quando usada para auditoria. Um *prompt* pode pedir que o modelo identifique frases que possam ser lidas como ranking punitivo, conclusões fortes demais, apagamento de limitações ou uso ambíguo de indicadores. Pode também pedir adaptação para linguagem mais prudente. Contudo, a IA não possui responsabilidade moral nem compreensão contextual plena. Sua auditoria deve ser uma ferramenta auxiliar, não a decisão final sobre ética do texto.

Relatórios públicos exigem cuidados adicionais. Quando resultados serão divulgados fora da instituição, a escrita deve evitar dados sensíveis, preservar privacidade quando necessário, explicar metodologia de modo acessível e prevenir interpretações midiáticas simplistas. Títulos e subtítulos merecem atenção, pois muitas leituras serão baseadas neles. Um título como "Universidade X lidera área Y" pode ser inadequado se o estudo mediu apenas produção em uma base. Melhor seria "Universidade X apresenta maior produção no *corpus* analisado sobre Y entre 2014 e 2024".

A comunicação ética também envolve reconhecer incerteza sem paralisar decisão. Gestores precisam decidir mesmo com dados incompletos. O papel do relatório não é

eliminar incerteza, mas qualificá-la. Ao explicar quais achados são robustos, quais são sinais preliminares e quais exigem investigação adicional, o documento oferece base mais honesta para ação. A IA pode ajudar a classificar níveis de evidência, mas a definição dos níveis deve ser estabelecida pela equipe.

Em pareceres, a ética exige independência intelectual. O parecerista deve resistir a pressões para validar conclusões desejadas quando os dados não sustentam. Deve também evitar postura excessivamente destrutiva quando problemas são corrigíveis. A escrita do parecer deve ser firme e respeitosa, orientada à melhoria do documento e ao uso responsável dos resultados. A IA pode ajudar a ajustar tom, mas não deve suavizar críticas metodológicas relevantes nem intensificar críticas além do necessário.

A escrita cientométrica responsável produz documentos que podem ser reutilizados, auditados e contestados. Essa abertura é sinal de qualidade. Um relatório que não permite contestação porque esconde método não é forte; é opaco. Um relatório que registra decisões, limitações e incertezas permite debate e melhoria. A ciência da ciência precisa dessa cultura de transparência, especialmente quando ferramentas de IA aceleram a produção de textos e tornam mais fácil gerar narrativas convincentes.

Conclui-se, portanto, que a ética da comunicação deve atravessar todas as seções: metodologia, resultados, discussão, sumário executivo, recomendações e anexos. Não é uma seção isolada, mas um modo de escrever. A IA pode ser in*corpora*da a esse processo como apoio à clareza, à revisão e à adaptação, desde que submetida a protocolos de validação e responsabilidade humana. O texto final deve permitir ao leitor compreender não apenas o que os dados indicam, mas também como chegaram a indicar, sob quais limites e para quais usos legítimos.

17.16 Referências, citações e integridade bibliográfica

A integridade bibliográfica é componente essencial da escrita cientométrica. Relatórios, artigos e pareceres devem citar tanto a literatura substantiva do tema analisado quanto a literatura metodológica que sustenta indicadores, mapas, estratégias de busca, visualizações e princípios de avaliação responsável. A ausência de referências metodológicas pode fazer parecer que procedimentos complexos são escolhas meramente instrumentais. Por outro lado, excesso de referências sem relação com o argumento produz erudição decorativa. A bibliografia deve ser funcional, verificável e coerente com as decisões do estudo. Em estudos cientométricos, essa integridade inclui tanto a fundamentação metodológica quanto a justificativa dos indicadores e procedimentos de análise (DONTHU *et al.*, 2021; HICKS *et al.*, 2015).

Em documentos assistidos por IA, o controle de referências torna-se ainda mais importante. Modelos generativos podem sugerir artigos inexistentes, misturar autores, alterar títulos, produzir DOIs incorretos ou atribuir conclusões a fontes que não as sustentam. Por isso, nenhuma referência sugerida por IA deve entrar no documento sem verificação em base confiável, página do periódico, *Crossref*, PubMed, Scopus, Web *of Science*, *Google Scholar, Lens* ou repositório institucional. A fluência da referência não é evidência de existência. Esse cuidado é necessário porque modelos generativos podem produzir saídas plausíveis, mas factualmente incorretas, inclusive em referências e atribuições (JI *et al.*, 2023).

A escrita deve evitar citações vagas como "a literatura mostra" sem identificar autores ou tradições. Em metodologia cientométrica, é preferível citar trabalhos específicos sobre bibliometria, mapeamento científico, indicadores, redes, visualização ou diretrizes de relato. Isso permite ao leitor avaliar a fundamentação. Quando o capítulo menciona estrutura *IMRAD*, diretrizes *PRISMA*, bibliometrix, VOSviewer ou Leiden Manifesto, deve associar cada elemento

a fonte verificável. Essa prática fortalece a rastreabilidade conceitual. Para estrutura de artigo, relato de revisões, estratégias de busca e ferramentas bibliométricas, as citações devem remeter a fontes específicas e verificáveis (SOLLACI; PEREIRA, 2004; PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021; ARIA; CUCCURULLO, 2017; VAN ECK; WALTMAN, 2010).

Links de referências devem ser estáveis sempre que possível. DOIs são preferíveis a páginas comerciais instáveis, embora links para textos abertos, repositórios ou páginas oficiais também sejam úteis. Em relatórios institucionais, recomenda-se manter lista bibliográfica com DOI ou URL e, quando necessário, data de acesso. Em artigos, deve-se seguir o estilo bibliográfico do periódico. A IA pode converter estilos, mas o resultado precisa ser conferido, pois erros de capitalização, acentuação e autoria são comuns.

A integridade bibliográfica também envolve evitar citação instrumental. Não se deve citar uma referência apenas porque é famosa se ela não sustenta a afirmação. O Leiden Manifesto, por exemplo, apoia princípios de avaliação responsável, mas não serve para justificar qualquer indicador. *PRISMA-S* orienta relato de buscas em revisões sistemáticas, mas deve ser adaptada com cuidado a estudos cientométricos. A escrita científica exige correspondência entre afirmação e referência. Assim, princípios como os do Leiden Manifesto devem ser mobilizados apenas quando sustentarem diretamente a afirmação interpretativa ou normativa apresentada (HICKS *et al.*, 2015).

Por fim, a lista de referências deve refletir pluralidade metodológica. Um capítulo sobre escrita de relatórios cientométricos precisa dialogar com bibliometria, comunicação científica, diretrizes de relato, avaliação responsável e IA generativa. Essa composição mostra que a redação não é etapa isolada, mas ponto de convergência entre método, evidência, interpretação e comunicação. A IA pode ajudar a identificar áreas bibliográficas ausentes, mas a curadoria final deve permanecer com o pesquisador.

A revisão final de referências deve ser tratada como etapa de controle de qualidade. É recomendável conferir se todos os autores citados aparecem na bibliografia, se todos os itens da bibliografia foram citados no texto, se links estão ativos, se os DOIs correspondem aos títulos e se não há duplicatas. Em relatórios longos, essa conferência evita inconsistências que reduzem confiança no documento. Quando a IA é usada para formatar referências, recomenda-se manter uma planilha de controle com título, autores, ano, periódico, DOI, URL e observação sobre a função da referência no argumento.

Outro cuidado envolve distinguir referência metodológica de fonte de dados. A base Scopus, a API OpenAlex ou o software VOSviewer podem ser descritos metodologicamente, mas a fundamentação de suas operações exige artigos, documentação ou referências técnicas. O relatório deve deixar claro quando uma fonte foi usada para coletar dados, quando foi usada para sustentar uma afirmação metodológica e quando foi usada apenas como ferramenta operacional. Essa distinção melhora a clareza e reduz confusão entre evidência empírica, literatura de apoio e infraestrutura de pesquisa.

17.17 Ampliação analítica: notas complementares para redação responsável

A redação cientométrica deve reconhecer que a ciência não é capturada integralmente por bases bibliográficas. Dados indexados representam parte da atividade científica, com maior ou menor cobertura conforme área, idioma, país, tipo documental e política editorial. Assim, uma frase aparentemente simples, como "a produção da área cresceu", deve ser qualificada: cresceu no *corpus* analisado, na base selecionada, no período definido e segundo a estratégia de busca utilizada. Essa qualificação não é excesso de prudência; é condição de honestidade metodológica.

O uso de IA na escrita pode induzir uniformização estilística. Muitos modelos tendem a produzir textos com estruturas previsíveis, vocabulário genérico e conclusões equilibradas demais. Em capítulos de livro, relatórios técnicos e artigos, cabe ao autor preservar especificidade do caso, densidade conceitual e conexão com literatura. A IA pode sugerir formulações, mas o texto final deve expressar o raciocínio analítico construído durante o estudo, não apenas uma síntese plausível de lugares-comuns acadêmicos.

Um documento cientométrico deve evitar a ilusão de completude. Mesmo quando apresenta muitas tabelas e mapas, ele responde a perguntas delimitadas. Parte importante da escrita consiste em dizer o que o estudo não cobre. Não cobrir livros, teses, dados de pesquisa, patentes ou produção técnica pode ser aceitável, desde que informado. O problema não é ter escopo limitado; o problema é escrever como se o escopo limitado fosse totalidade.

A relação entre texto e visualização deve ser complementar. O gráfico mostra padrões; o texto indica o que observar, como interpretar e que limites considerar. Quando o texto apenas repete o gráfico, desperdiça espaço. Quando o texto afirma algo que o gráfico não mostra, rompe rastreabilidade. A redação ideal constrói uma ponte entre evidência visual e argumento interpretativo, preservando a capacidade do leitor de verificar a afirmação.

Pareceres cientométricos diferem de relatórios porque avaliam a qualidade, a suficiência ou o uso potencial de uma análise. Um parecer pode revisar um relatório já produzido, avaliar proposta metodológica, emitir recomendação sobre indicadores ou apoiar decisão institucional. Sua escrita deve ser equilibrada: reconhecer méritos, apontar fragilidades, distinguir problemas corrigíveis de limitações estruturais e formular recomendações objetivas. A IA pode ajudar a organizar pareceres, mas não deve substituir juízo especializado.

A redação de artigos cientométricos exige atenção ao público disciplinar do periódico. Periódicos de ciência da informação valorizam detalhamento metodológico em bibliometria; periódicos substantivos podem valorizar mais implicações para a área analisada. O autor deve ajustar ênfase sem suprimir rigor. Um artigo sobre produção em saúde coletiva, por exemplo, precisa dialogar tanto com métodos cientométricos quanto com debates da saúde coletiva. A IA pode auxiliar na adaptação do enquadramento, mas referências e interpretações devem ser verificadas.

A transparência sobre uso de IA não precisa ser defensiva. Pode ser apresentada como parte do protocolo: "Ferramentas de IA generativa foram utilizadas para apoio à revisão de linguagem, organização preliminar de seções e sugestão de rótulos temáticos; todas as saídas foram revisadas por autores humanos e conferidas contra os dados". Essa formulação comunica utilidade e controle. O que deve ser evitado é ocultar uso relevante ou atribuir à IA responsabilidade interpretativa.

A adaptação para público não técnico deve evitar metáforas simplificadoras que alterem o fenômeno. Dizer que um indicador é uma "nota" da ciência é inadequado, pois sugere avaliação direta de qualidade. Melhor dizer que indicadores são sinais parciais que ajudam a observar dimensões específicas, como volume de publicações, citações ou colaboração. Essa diferença linguística protege contra leitura normativa excessiva.

Em ambientes institucionais, relatórios cientométricos podem ter consequências sobre financiamento, reputação e avaliação de pessoas. Por isso, a escrita deve ser sensível a efeitos de uso. Comparações entre departamentos, grupos ou pesquisadores devem ser contextualizadas por missão, área, tamanho, carreira, tipo de produção e recursos. A IA, se instruída apenas a "destacar melhores desempenhos", pode reforçar rankings simplistas. O *prompt* deve orientar interpretação responsável.

A bibliografia de um relatório ou artigo cientométrico cumpre dupla função. Ela sustenta a área analisada e sustenta o método. Muitos estudos citam apenas literatura substantiva e esquecem referências sobre bibliometria, indicadores, visualização e avaliação responsável. Outros citam apenas metodologia e ignoram debates do domínio. Um documento robusto combina ambos. A IA pode sugerir lacunas bibliográficas, mas as referências devem ser conferidas em bases confiáveis.

A escrita deve preservar memória das decisões. Em projetos longos, escolhas sobre exclusão de termos, agrupamento de instituições, correção de nomes e seleção de indicadores podem se perder. O relatório final deve recuperar essas decisões de forma organizada. Isso é especialmente relevante quando diferentes membros da equipe participaram do processo ou quando o relatório será atualizado periodicamente.

A produção de documentos cientométricos deve considerar acessibilidade. Tabelas extensas devem ser legíveis, figuras devem ter legendas claras, siglas devem ser definidas, cores devem ser usadas com cautela e informações essenciais não devem depender apenas de distinções cromáticas. A IA pode ajudar a revisar legibilidade textual, mas a acessibilidade visual exige inspeção humana e, quando possível, testes com usuários.

17.18 Considerações finais

A escrita de relatórios, artigos e pareceres cientométricos é uma prática metodológica e ética. Ela transforma dados bibliográficos em argumentos, indicadores em evidências, mapas em interpretações e achados em possíveis decisões. Em estudos assistidos por IA, essa transformação deve ser ainda mais cuidadosa, pois modelos de linguagem aumentam a velocidade de produção textual e, simultaneamente, podem amplificar riscos de extrapolação, padronização acrítica e autoridade aparente.

O capítulo defendeu que a escrita cientométrica deve separar descrição, interpretação, hipótese e recomendação. Deve também documentar fontes, métodos, parâmetros, *prompts* e validações. A IA pode apoiar estruturação, redação, revisão, síntese e adaptação de linguagem, mas precisa ser governada por *prompts* restritivos e por revisão humana. A qualidade do documento final depende menos da fluência do texto e mais da transparência do percurso analítico.

Relatórios cientométricos são frequentemente usados em contextos de gestão e política científica. Por isso, sua escrita deve evitar tanto o jargão inacessível quanto a simplificação que distorce. A comunicação deve ser clara, mas prudente; objetiva, mas contextualizada; útil, mas não prescritiva além dos dados. A boa escrita cientométrica não apenas informa: ela ajuda leitores a decidir melhor, compreender limites e formular novas perguntas.

Capítulo 18

Questões epistemológicas, sociológicas e éticas

18.1 Introdução

Os capítulos anteriores estabeleceram bases conceituais, metodológicas e operacionais para o uso de inteligência artificial em estudos cientométricos. Foram discutidas perguntas de pesquisa, estratégias de busca, coleta de dados, pré-processamento, indicadores, redes, mapas temáticos, análise temporal, leitura assistida, visualização e redação de resultados. O presente capítulo desloca o foco da operacionalização para a crítica. Seu objetivo é explicitar que toda análise cientométrica, com ou sem IA, está assentada sobre escolhas epistemológicas, pressupostos sociológicos e responsabilidades éticas.

A cientometria opera sobre rastros formais da ciência. Títulos, resumos, palavras-chave, referências, citações, afiliações, periódicos, áreas de indexação e identificadores persistentes são sinais documentais de atividades científicas. Esses sinais permitem medir, mapear e comparar. Entretanto, não devem ser confundidos com a totalidade da ciência. Pesquisa também envolve experimentos malsucedidos, dados não publicados, revisão por pares, trabalho técnico, orientação, circulação informal de ideias, construção de instrumentos, manutenção de infraestrutura, debates em seminários, produção local, tradução de conhecimento, engajamento com comunidades e contribuições que raramente aparecem com a mesma nitidez nos metadados bibliográficos.

A introdução de modelos de linguagem e outros sistemas de IA intensifica essa ambiguidade. Por um lado, a IA amplia a capacidade do pesquisador de lidar com *corpora* extensos, identificar padrões, comparar interpretações, formular hipóteses e redigir relatórios. Por outro, ela pode produzir enunciados fluentes sem lastro empírico, reforçar assimetrias dos dados de treinamento, ocultar decisões algorítmicas, inventar referências, simplificar controvérsias e emprestar aparência de objetividade a interpretações frágeis. Por isso, a questão ética não se limita a perguntar se a IA é útil. A questão é perguntar sob quais condições seu uso preserva responsabilidade científica, rastreabilidade e justiça avaliativa.

Este capítulo parte da premissa de que mapas, indicadores e rankings não são espelhos neutros da ciência. Eles são representações construídas. Essa afirmação não implica relativismo metodológico nem rejeição das métricas. Implica reconhecer que a validade de uma análise depende da adequação entre pergunta, dados, método, contexto e uso pretendido. Um indicador pode ser excelente para descrever volume de publicações e inadequado para inferir qualidade. Um mapa pode ser útil para revelar comunidades temáticas e inadequado para afirmar causalidade intelectual. Um ranking pode orientar uma decisão exploratória e, ao mesmo tempo, produzir efeitos de hierarquização quando usado de modo performativo em avaliação institucional.

A literatura sobre métricas responsáveis recomenda que a avaliação quantitativa apoie, e não substitua, a avaliação qualitativa especializada. O Manifesto de Leiden e a Declaração de São Francisco sobre Avaliação da Pesquisa (DORA) convergem na crítica ao uso de métricas de periódicos como substitutos para julgar artigos, pesquisadores ou programas. Estudos sobre comportamento citacional e cobertura de bases reforçam que idioma, país, área, tipo de publicação e finalidade avaliativa condicionam o que aparece como ciência mensurável (AMERICAN SOCIETY FOR CELL BIOLOGY, 2012; HICKS *et al.*, 2015).

No contexto da trilha deste livro, *prompts* assumem papel central. Eles não são apenas instruções operacionais. São dispositivos de enquadramento. Um *prompt* define que dados serão considerados relevantes, que perspectiva interpretativa será adotada, que formato de resposta será privilegiado, que incertezas serão explicitadas e que limites serão reconhecidos. Um *prompt* pode induzir prudência, separando dado, inferência e hipótese. Mas também pode induzir excesso de confiança, pedindo conclusões fortes a partir de evidências fracas. Assim, a ética do uso de IA em cientometria passa também pela ética do desenho, da documentação e da auditoria dos *prompts*.

Para desenvolver esse argumento, o capítulo avança das limitações dos indicadores e das fontes documentais para os problemas de interpretação de citações, vieses de cobertura, visualizações, opacidade algorítmica e transparência dos *prompts*. Em seguida, integra essas dimensões em instrumentos operacionais: *prompts* críticos, estudo de caso, checklist, protocolo de auditoria, matriz decisória, diretrizes de governança e critérios mínimos para publicação.

18.2 O que os indicadores mostram e o que escondem

Indicadores cientométricos são instrumentos de redução da complexidade. Eles mostram padrões que dificilmente seriam percebidos por leitura manual isolada, como crescimento de produção, distribuição de citações, colaboração entre instituições, internacionalização, concentração temática e circulação de resultados. Entretanto, todo indicador é também uma escolha de visibilidade: ao selecionar uma dimensão mensurável, deixa outras dimensões fora do quadro analítico.

Do ponto de vista epistemológico, a questão não se limita à correção técnica do indicador: é necessário compreender que realidade ele torna visível, que pressupostos incorpora e quais dimensões permanecem fora do campo de observação. Por isso, métricas de produção, citação ou colaboração devem ser lidas em contexto, com atenção ao campo, à escala e à finalidade avaliativa (BORNMANN; DANIEL, 2008; HICKS *et al.*, 2015; SEGLEN, 1997; WALTMAN, 2016).

Um indicador de produção, por exemplo, mostra volume de publicações, mas não mostra necessariamente originalidade, robustez metodológica, contribuição social, qualidade dos dados, relevância local ou dificuldade de pesquisa. O número de artigos pode indicar vitalidade de uma área, mas também pode refletir pressão por publicação, fragmentação de resultados, diferenças de cultura editorial ou estratégias de produtividade. Quando usado sem contexto, o indicador transforma uma medida de volume em um juízo indevido sobre mérito.

Indicadores de citação tornam visíveis relações de uso, reconhecimento e circulação textual. Contudo, o ato de citar não possui significado único. Uma citação pode indicar influência positiva, crítica, contextualização, uso metodológico, deferência a autores canônicos, autocitação, obrigação editorial ou mera inclusão retórica. Por isso, contagens de citação devem ser interpretadas como sinais de atenção acadêmica, não como equivalentes diretos de qualidade científica.

Indicadores de colaboração mostram padrões de coautoria, coprodução institucional e internacionalização. Ainda assim, a coautoria formal não captura integralmente a divisão real do trabalho. Técnicos, estudantes, analistas, comunidades participantes, revisores, curadores de dados e equipes de infraestrutura podem contribuir decisivamente sem aparecer como autores. Além disso, áreas com tradição de autoria coletiva tendem a produzir indicadores de colaboração muito diferentes de áreas nas quais a autoria individual ou de pequenos grupos permanece predominante.

Indicadores alternativos e de atenção social ampliam a observação para além das citações acadêmicas, incluindo menções em redes sociais, documentos de política, notícias, gestores de referência e plataformas digitais. Eles podem revelar circulação pública ou profissional do conhecimento, mas também são sensíveis a marketing, controvérsia, celebridade, disponibilidade online e desigualdades de acesso digital. Portanto, atenção social não deve ser confundida automaticamente com impacto social positivo.

Na interpretação desses indicadores, a IA pode apoiar a leitura crítica, desde que o *prompt* solicite limitações, vieses, alternativas de interpretação e consequências do uso da métrica. Comandos que pedem apenas para "explicar o desempenho" de um pesquisador a partir de uma tabela tendem a apagar a cobertura da fonte, os padrões disciplinares, o estágio da carreira, as práticas de coautoria e a incerteza.

O primeiro risco ético consiste em reificar indicadores. Reificação ocorre quando uma medida construída passa a ser tratada como atributo natural do objeto medido. Dizer que um artigo tem trinta citações é uma descrição. Dizer que ele é três vezes melhor que outro artigo com dez citações é uma extrapolação sem fundamento, salvo em contextos muito específicos e explicitamente justificados.

O segundo risco consiste em deslocar a finalidade da métrica. Um indicador criado para análise agregada de campos pode ser usado indevidamente para avaliação individual. Um indicador desenhado para comparar periódicos pode ser aplicado a pesquisadores. Uma medida útil para diagnóstico institucional pode ser transformada em regra de financiamento. Tais deslocamentos alteram o significado do indicador e podem produzir incentivos perversos.

O terceiro risco é a substituição da pergunta substantiva pela disponibilidade do dado. Estudos cientométricos podem passar a perguntar apenas aquilo que as bases permitem medir. Nesse caso, a infraestrutura de dados deixa de ser meio e passa a definir o horizonte da investigação. A IA pode agravar esse problema ao sugerir indicadores padronizados sem questionar sua adequação ao objetivo analítico.

A responsabilidade analítica permanece com o pesquisador. Cabe-lhe explicitar o que cada indicador mede e não mede, em que escala faz sentido, quais comparações são legítimas, que normalizações foram aplicadas e quais inferências permanecem proibidas. Essa responsabilidade aumenta quando os resultados subsidiam a avaliação de pessoas, programas, departamentos, países ou políticas.

Considere um programa de pós-graduação que apresenta forte crescimento de publicações em cinco anos. Uma leitura apressada poderia interpretar o crescimento como melhoria de qualidade. Uma leitura responsável perguntaria se houve aumento de docentes credenciados, mudança de base de indexação, alteração da política de publicação, maior participação em artigos multicêntricos, inclusão de artigos de conferência, expansão de coautorias internacionais ou mudança de estratégia de busca. A IA pode ajudar a formular essas hipóteses, mas não pode confirmá-las sem dados adicionais.

Quadro 18.1 - Potencial e limites de indicadores cientométricos

Indicador	Mostra	Pode esconder	Pergunta ética
Número de publicações	Volume formal de produção indexada	Qualidade, contribuição local, trabalho não publicado, pressão produtivista	A métrica está sendo usada como *proxy* indevido de qualidade?

Indicador	Mostra	Pode esconder	Pergunta ética
Número de citações	Atenção acadêmica registrada	Motivo da citação, críticas, autocitações, desigualdade de campo	A interpretação distingue visibilidade de mérito?
Índice h	Combinação cumulativa de produção e citação	Tempo de carreira, área, autoria coletiva, desigualdade de oportunidade	A comparação é justa entre trajetórias diferentes?
Colaboração internacional	Coprodução formal entre países	Assimetria de liderança, dependência, autoria honorária	A métrica examina poder ou apenas presença de países?
Altmetrics	Atenção em ambientes digitais	Marketing, controvérsia, bots, desigualdade de acesso	A atenção social é interpretada como impacto sem evidência?

Fonte: Elaborado pelo autor (2026).

Outro exemplo é a comparação entre duas áreas. Uma área biomédica pode apresentar maior média de citações que uma área de humanidades. Isso não autoriza afirmar maior impacto intelectual em sentido geral. Diferenças de tamanho da comunidade, velocidade de obsolescência, densidade de referências, idioma, tipo de publicação e cultura de citação tornam comparações diretas metodologicamente frágeis.

18.3 Ciência publicada versus ciência em andamento

A cientometria trabalha predominantemente com ciência publicada, indexada e documentada. Esse recorte é poderoso, pois artigos, livros, anais e registros bibliográficos constituem evidências formais da comunicação científica. Entretanto, a ciência em andamento é mais ampla do que aquilo que chega à publicação. Entre a formulação de uma pergunta e o registro bibliográfico final há processos invisíveis, negociações, falhas, ensaios, exclusões, atrasos, censuras, recursos escassos e decisões institucionais.

Dados publicados e indexados constituem amostras estruturadas da atividade científica, e não totalidades neutras. Vieses de publicação, cobertura e comunicação disciplinar condicionam a passagem da pesquisa em andamento ao registro bibliográfico (FANELLI, 2009; IOANNIDIS, 2005; MONGEON; PAUL-HUS, 2016; VISSER *et al.*, 2021).

A ciência publicada é uma versão estabilizada da pesquisa. Ela passou por seleção editorial, revisão por pares, adequação a gêneros textuais, tradução, normalização, indexação e, muitas vezes, reescrita estratégica. A ciência em andamento, por sua vez, envolve incerteza, tentativa, erro, debate informal, dados parciais, instrumentos em desenvolvimento e problemas ainda não estabilizados. A diferença entre esses dois planos é central para evitar que mapas da ciência sejam interpretados como retratos completos da atividade científica.

A publicação também depende de recursos desigualmente distribuídos. Grupos com financiamento, domínio do inglês acadêmico, redes internacionais, apoio estatístico, acesso a periódicos de prestígio e experiência editorial tendem a ter maior probabilidade de transformar pesquisa em publicação indexada. Isso não significa que grupos menos visíveis produzam menos conhecimento substantivo. Significa que seus resultados podem circular por canais menos capturados por bases bibliográficas dominantes.

Há ainda o problema do viés de publicação. Resultados positivos, estatisticamente significativos, inovadores ou alinhados a agendas dominantes têm maior probabilidade de

publicação em alguns campos. Resultados negativos, replicações, estudos locais, dados inconclusivos e trabalhos de infraestrutura podem permanecer sub-representados. Uma análise cientométrica baseada apenas em publicações pode reproduzir esses filtros, apresentando como estrutura da ciência aquilo que é, em parte, estrutura da publicação científica.

Nos estudos assistidos por IA, esse problema exige atenção especial. Modelos de linguagem tendem a interpretar o *corpus* fornecido como se fosse o universo relevante, salvo quando explicitamente instruídos a considerar limites de cobertura. Se o conjunto de dados contém apenas artigos publicados em bases internacionais, a IA pode produzir uma narrativa sobre "a área" sem mencionar que literatura cinzenta, teses, relatórios técnicos, *preprints*, documentos de política, produção local e conhecimento em línguas não hegemônicas ficaram de fora.

A IA pode ajudar a listar tipos de contribuição que provavelmente não aparecem no *corpus*, sugerir fontes complementares, identificar lacunas de cobertura e revisar a seção de limitações. Essas sugestões, contudo, devem ser tratadas como hipóteses de auditoria e verificadas em fontes e documentos adicionais.

O primeiro risco é confundir ausência de evidência indexada com ausência de atividade científica. Essa confusão é particularmente grave em áreas aplicadas, saúde pública, educação, políticas territoriais, humanidades e ciências sociais, nas quais parte da contribuição pode assumir forma de relatório, intervenção, acervo, norma, parecer, base de dados, software ou produção em língua local.

O segundo risco é premiar apenas a ciência que se ajusta aos formatos mais visíveis. Quando métricas de publicação orientam financiamento e reputação, pesquisadores podem adaptar agendas a temas mais publicáveis, reduzir investimentos em pesquisa de longo prazo ou evitar problemas locais que não circulam bem em periódicos centrais. A cientometria pode diagnosticar tais tendências, mas também pode reforçá-las se seus resultados forem usados de modo acrítico.

O terceiro risco é atribuir à IA uma capacidade de síntese que ela não possui. Um modelo pode produzir uma narrativa convincente sobre uma área com base em resumos indexados, mas essa narrativa será tão limitada quanto o *corpus* e as categorias fornecidas. O modelo não sabe, por si, o que ficou fora da amostra.

Cabe ao pesquisador reconstruir o contexto de produção dos dados: canais de comunicação incluídos e excluídos, idiomas predominantes, tipos documentais considerados e práticas científicas que permanecem invisíveis. A escrita do estudo deve explicitar essas fronteiras e evitar que ausência de registro seja confundida com ausência de atividade científica.

Em um estudo sobre inovação em saúde indígena, uma busca em Scopus ou Web of Science pode recuperar artigos internacionais, mas deixar de fora relatórios de organizações comunitárias, documentos do sistema de saúde, dissertações, protocolos locais e materiais bilíngues. Uma análise que ignore essa diferença pode afirmar baixa produção científica quando, na verdade, está observando baixa indexação de determinados formatos de conhecimento.

Em estudos sobre tecnologias educacionais, *preprints*, repositórios institucionais, anais de conferência, softwares e materiais pedagógicos podem circular antes ou fora dos periódicos. A ausência desses registros altera a percepção de emergência de temas e de liderança institucional.

18.4 Citações como reconhecimento, retórica, convenção ou estratégia

A citação é um dos pilares da cientometria, mas também um de seus objetos mais ambíguos. Contar citações parece simples; interpretar seu significado é complexo. Desde os estudos clássicos de sociologia da ciência e análise de citações, sabe-se que citar pode expressar reconhecimento intelectual, dependência metodológica, filiação teórica, contraste, crítica, ritual disciplinar, convenção editorial, estratégia de visibilidade ou resposta a expectativas de pares.

Desde a criação dos índices de citação, a contagem tornou-se instrumento central para observar circulação e reconhecimento (Garfield, 1955). Entretanto, seu significado não é único: citações podem cumprir funções intelectuais, retóricas, críticas, convencionais e estratégicas (BORNMANN; DANIEL, 2008; COZZENS, 1989; MACROBERTS; MACROBERTS, 1989; MERTON, 1968; TAHAMTAN; BORNMANN, 2019).

Quando uma citação é tomada como reconhecimento, supõe-se que o documento citado contribuiu positivamente para o documento citante. Essa suposição sustenta grande parte dos indicadores de impacto. Contudo, nem toda citação é elogiosa ou substantiva. Um artigo pode ser citado como exemplo de erro, como objeto de crítica, como literatura obrigatória, como contexto histórico ou como item marginal em uma longa lista de referências.

A citação também possui função retórica. Autores constroem credibilidade situando seu argumento em relação a tradições, autores canônicos, metodologias reconhecidas e debates estabelecidos. Nesse sentido, a bibliografia de um artigo não é apenas registro de dívida intelectual; é também dispositivo de posicionamento discursivo. Citar pode ser uma forma de declarar pertencimento a uma comunidade epistêmica.

Além disso, citações seguem convenções disciplinares. Em alguns campos, artigos possuem listas extensas de referências e ciclos rápidos de citação. Em outros, livros, textos clássicos e debates de longa duração reduzem a densidade citacional de artigos. As áreas também variam quanto à idade da literatura citada, ao peso de revisões, à relevância de conferências e à centralidade de documentos técnicos. Indicadores que ignoram essas diferenças podem penalizar campos inteiros.

Há ainda estratégias explícitas ou implícitas de citação. Autocitações, citações cruzadas entre grupos, citações a periódicos do mesmo núcleo editorial, inclusão de referências sugeridas por revisores e escolhas orientadas por visibilidade podem afetar padrões de impacto. Nem toda estratégia é indevida, mas todas mostram que a citação é um ato social situado, não um reflexo mecânico de qualidade.

A IA pode ser útil para examinar contextos de citação, resumir padrões de uso, separar citações metodológicas de conceituais em amostras e identificar ambiguidades. Modelos generativos, porém, podem atribuir intenção a autores sem evidência textual; por isso, os *prompts* devem proibir inferência psicológica e exigir justificativas baseadas em trechos verificáveis.

O risco fundamental é interpretar contagens de citação como medida direta de valor científico. Isso ignora que citações são distribuídas de forma assimétrica, acumulativa e dependente de visibilidade prévia. O efeito Mateus, formulado por Merton (1968) e observado em indicadores de impacto por Larivière e Gingras (2010), indica que o reconhecimento tende a se acumular em torno de autores, instituições, periódicos e países já prestigiados, ampliando desigualdades de atenção.

Outro risco é desconsiderar citações negativas ou ambivalentes. Sistemas de contagem raramente distinguem se a citação apoia, critica, corrige ou apenas menciona o trabalho citado. Técnicas de análise de contexto citacional e processamento de linguagem

podem ajudar, mas ainda exigem validação cuidadosa, especialmente quando aplicadas em larga escala e em múltiplos idiomas.

Um terceiro risco ocorre quando a IA transforma padrões citacionais em narrativas causais. Por exemplo, se um autor aparece como muito citado em um *cluster*, o modelo pode afirmar que ele "fundou" a área ou "determinou" sua evolução. Tal inferência pode ser plausível, mas precisa de evidência histórica, leitura dos textos e análise de conteúdo. Centralidade citacional não equivale automaticamente a autoria intelectual de um campo.

Em análises de cocitação, é legítimo afirmar que dois autores são frequentemente citados em conjunto e que isso sugere associação intelectual percebida pela literatura. Sem análise adicional, não se deve concluir que colaboraram, concordavam teoricamente ou pertenciam à mesma escola. A responsabilidade humana consiste em definir e respeitar esse nível de inferência.

Uma rede de cocitação pode mostrar proximidade entre autores de tradições diferentes porque ambos são usados como polos de comparação. Se a IA rotula o *cluster* como "escola unificada", pode apagar controvérsias. Um rótulo mais prudente seria "literatura mobilizada em debates sobre determinado problema", seguido da indicação de que a natureza da relação exige leitura qualitativa.

Em avaliações institucionais, alto número de citações de um grupo pode indicar influência, mas também pode decorrer de participação em grandes consórcios, trabalhos de revisão, bases de dados amplamente utilizadas ou temas de alta visibilidade. A interpretação deve distinguir tipo de contribuição.

18.5 Viés linguístico, geográfico, disciplinar e de base de dados

Nenhuma base bibliográfica oferece cobertura completa, homogênea e neutra da ciência mundial. Bases diferentes têm políticas distintas de seleção, indexação, normalização, atualização, classificação disciplinar e tratamento de documentos. Como consequência, toda análise cientométrica é atravessada por vieses de idioma, região, disciplina, tipo documental e infraestrutura de dados.

Estudos comparativos demonstram que políticas de seleção, idioma, região, disciplina, tipo documental e qualidade dos metadados afetam os resultados bibliométricos. A inclusão em bases seletivas tampouco deve ser tomada isoladamente como sinônimo de qualidade editorial (ARCHAMBAULT *et al.*, 2009; CHAVARRO *et al.*, 2018; MONGEON; PAUL-HUS, 2016; SINGH *et al.*, 2021; VISSER *et al.*, 2021).

O viés linguístico é um dos mais evidentes. Bases internacionais tendem a privilegiar periódicos em inglês, especialmente nos campos de maior circulação global. Pesquisas relevantes publicadas em português, espanhol, francês, chinês, árabe ou outras línguas podem aparecer de modo parcial ou permanecer fora do *corpus*. Esse viés afeta particularmente temas locais, políticas públicas, educação, saúde coletiva, direito, cultura, agricultura regional e estudos territoriais.

O viés geográfico decorre da concentração de periódicos, editoras, indexadores e redes de colaboração em determinados países. Países centrais no sistema científico internacional tendem a ter maior visibilidade e maior capacidade de influenciar categorias de indexação, agendas de pesquisa e padrões editoriais. Países periféricos ou semiperiféricos podem aparecer como menos produtivos ou menos influentes, mesmo quando produzem conhecimento relevante para seus contextos.

O viés disciplinar manifesta-se na diferença de cobertura entre áreas. Ciências biomédicas e naturais geralmente são melhor representadas em bases de artigos de

periódicos, enquanto humanidades, artes, direito, ciências sociais aplicadas e engenharia podem depender de livros, capítulos, anais, normas, patentes, relatórios ou produção técnica. Uma mesma estratégia de coleta pode ser adequada em uma área e insuficiente em outra.

O viés de base de dados surge porque Web of Science, Scopus, Dimensions, OpenAlex, *Crossref,* PubMed*, Lens*, SciELO, DOAJ, Google Scholar e repositórios institucionais diferem em cobertura, qualidade de metadados, links de citação, deduplicação, disponibilidade de resumos e transparência. A escolha da base não é detalhe técnico; é decisão metodológica que molda os resultados.

A IA pode auxiliar na identificação de potenciais vieses se for instruída a procurar lacunas de idioma, país, área, tipo documental e fonte. Também pode comparar estratégias de busca, sugerir descritores em diferentes línguas e revisar afirmações para adequar o escopo. Suas sugestões precisam ser verificadas nas bases e nos documentos recuperados.

O risco mais comum é apresentar resultados como se fossem universais quando são produtos de uma fonte específica. Frases como "a produção mundial sobre o tema" devem ser substituídas, quando apropriado, por "a produção indexada na base X segundo a estratégia Y". Essa precisão evita que limites de dados sejam convertidos em afirmações globais.

Outro risco é naturalizar desigualdades de visibilidade. Quando rankings e mapas mostram centros dominantes, isso pode refletir excelência, mas também língua, recursos, redes históricas, políticas editoriais e cobertura de base. Sem contextualização, a análise reforça hierarquias existentes e transforma vantagem acumulada em prova de superioridade intrínseca.

A IA pode amplificar esse problema ao produzir narrativas globais com base em dados parciais. Se o *prompt* não exigir análise de cobertura, o modelo pode descrever o *corpus* como se fosse a área inteira. Além disso, modelos treinados majoritariamente em textos em inglês podem reproduzir vieses de idioma na interpretação de relevância.

A responsabilidade metodológica consiste em documentar fontes, justificar a seleção de bases, comparar coberturas quando possível, reconhecer ausências e evitar generalizações indevidas. Em áreas sensíveis ao idioma e ao contexto local, bases regionais, repositórios institucionais e literatura cinzenta podem ser complementos indispensáveis.

Um estudo sobre políticas de alfabetização publicado apenas com dados de Web of Science pode favorecer literatura em inglês e excluir produção nacional em periódicos educacionais locais. A conclusão de que determinado país produz pouco sobre o tema seria injustificada sem comparação com bases regionais e repositórios.

Em ciência aberta, bases como OpenAlex e Crossref podem ampliar cobertura, mas apresentam desafios de qualidade e completude de metadados. Já bases comerciais podem oferecer curadoria mais rígida, porém com seletividade que exclui periódicos locais. A escolha depende da pergunta.

18.6 A falsa neutralidade dos mapas e rankings

Mapas e rankings possuem forte poder retórico. Um mapa de rede, um mapa temático ou uma lista ordenada de instituições parece oferecer visão objetiva da estrutura científica. Contudo, tais visualizações dependem de escolhas de base, recorte temporal, unidade de análise, tipo de relação, normalização, algoritmo de *cluster*ização, limiar de inclusão, layout gráfico, rótulo e escala. A aparente neutralidade visual pode ocultar uma cadeia extensa de decisões.

A crítica à neutralidade de mapas e rankings decorre do fato de que toda visualização resulta de escolhas de dados, normalização, limiares, algoritmos, rótulos e escalas. Estudos

de mapeamento científico e avaliação responsável mostram que essas escolhas afetam tanto a forma visual quanto os usos institucionais dos resultados (CALLON; COURTIAL; LAVILLE, 1991; COBO *et al.*, 2011; DE RIJCKE *et al.*, 2016; HICKS *et al.*, 2015; RAFOLS *et al.*, 2012; VAN ECK; WALTMAN, 2010).

Mapas de ciência são representações. Eles tornam visíveis proximidades, centralidades, comunidades e trajetórias, mas não devem ser confundidos com o território da ciência. Dois mapas construídos com os mesmos dados podem diferir se usarem cocitação, acoplamento bibliográfico ou coocorrência de palavras-chave. Podem diferir ainda mais se alterarem limiares, algoritmos ou métodos de normalização. A visualização é produto de modelagem.

Rankings são ainda mais sensíveis a decisões normativas. Ordenar universidades, autores, periódicos ou países exige escolher critérios, pesos, janelas temporais, normalizações, categorias e fontes. O resultado transmite hierarquia, ainda que a métrica seja parcial. O perigo ético surge quando a ordenação é usada para premiar, punir ou orientar recursos sem considerar missão institucional, contexto local e diversidade de contribuições.

O layout visual também cria sentido. Nós centrais parecem importantes, cores sugerem grupos naturais, tamanhos de bolha indicam relevância, distâncias insinuam proximidade intelectual. Esses elementos podem ser úteis, mas também induzem leitura excessiva. Em redes densas, pequenos ajustes de layout podem alterar a percepção de centralidade. Em mapas temáticos, rótulos atribuídos a *clusters* podem estabilizar interpretações que ainda são provisórias.

Quando IA é usada para rotular *clusters*, escrever narrativas de mapas ou resumir rankings, o risco de naturalização aumenta. O modelo tende a transformar padrões visuais em frases interpretativas. Se não for instruído a mencionar limitações, pode afirmar que um *cluster* representa "a principal escola" ou que uma instituição "lidera a área" sem discutir cobertura, método ou incerteza.

A IA pode auxiliar na auditoria de mapas e rankings ao formular perguntas críticas sobre decisões de visualização, identificar interpretações indevidas e propor legendas analíticas com cautelas. Não deve, contudo, decidir sozinha o rótulo final de *clusters* nem recomendar o uso avaliativo de rankings sem análise normativa e validação substantiva.

O primeiro risco é a estetização da evidência. Visualizações bonitas podem parecer convincentes mesmo quando baseadas em dados incompletos, escolhas arbitrárias ou algoritmos pouco compreendidos. Beleza gráfica não é validade metodológica.

O segundo risco é a performatividade. Rankings e mapas não apenas descrevem; podem influenciar comportamento. Pesquisadores passam a buscar posições de visibilidade, instituições ajustam estratégias para melhorar indicadores, periódicos adotam políticas para aumentar métricas e gestores simplificam decisões complexas em painéis. A métrica passa a intervir no sistema que mede.

O terceiro risco é o apagamento do contexto. Um mapa que destaca países centrais pode ocultar redes locais relevantes. Um ranking de periódicos pode desconsiderar missões editoriais regionais. Uma rede de coautoria pode mostrar baixa colaboração internacional sem considerar barreiras de financiamento, idioma e infraestrutura.

Cabe ao pesquisador apresentar mapas e rankings como construções metodológicas. A legenda deve explicitar fonte, período, unidade, limiar, algoritmo, normalização e interpretação permitida. Sempre que possível, a robustez deve ser testada com parâmetros alternativos, e mudanças relevantes precisam ser relatadas.

Um mapa de coocorrência de palavras-chave pode mostrar "inteligência artificial", "aprendizagem de máquina" e "educação" no mesmo *cluster*. A IA poderia rotulá-lo como "IA na educação". Mas, se os resumos indicarem que os trabalhos tratam majoritariamente de avaliação automatizada, o rótulo mais preciso seria "avaliação educacional mediada por aprendizagem de máquina". A validação textual altera o significado do mapa.

Um ranking de universidades por produção em um tema pode favorecer instituições grandes. Sem normalização por tamanho, missão ou corpo docente, o ranking mede volume absoluto, não intensidade, qualidade ou especialização.

18.7 IA, opacidade e responsabilidade analítica

A adoção de IA em estudos cientométricos introduz novas camadas de opacidade. Mesmo quando os dados bibliográficos são documentados, o processo pelo qual um modelo generativo produz uma classificação, uma síntese, um rótulo de *cluster* ou uma interpretação pode não ser plenamente transparente. Modelos de linguagem são sistemas probabilísticos treinados em grandes *corpora*, sensíveis a *prompt*, contexto, parâmetros e versões de modelo. Sua fluência pode mascarar incerteza.

Riscos de alucinação, opacidade e reprodução de vieses exigem prudência na incorporação de modelos generativos à pesquisa científica. A literatura sobre ética algorítmica e sistemas sociotécnicos reforça a necessidade de explicabilidade, justiça, validação e responsabilização (BENDER *et al.*, 2021; FLORIDI; COWLS, 2019; JI *et al.*, 2023; JOBIN; IENCA; VAYENA, 2019; SELBST *et al.*, 2019).

A opacidade manifesta-se de várias formas. Há opacidade técnica, relacionada à arquitetura, aos dados de treinamento, aos pesos e aos mecanismos internos do modelo. Há opacidade operacional, relacionada a versões, parâmetros, filtros, limites de contexto e mudanças de comportamento ao longo do tempo. Há opacidade interpretativa, relacionada à dificuldade de saber por que determinada resposta foi gerada. Em cientometria, todas essas camadas importam porque o modelo pode influenciar classificação, rotulagem, síntese e redação.

Um problema central é a alucinação. Modelos generativos podem produzir afirmações plausíveis, mas não verificadas; podem inventar referências, autores, periódicos, indicadores ou explicações históricas; podem completar lacunas com padrões linguísticos aprendidos, não com evidências do *corpus*. Quando usados em estudos científicos, esses erros ameaçam rastreabilidade e confiabilidade.

Outro problema é a autoridade retórica. A resposta da IA pode soar organizada, equilibrada e tecnicamente sofisticada. Essa forma de apresentação pode reduzir a vigilância do pesquisador, especialmente quando confirma expectativas prévias. O risco não é apenas erro factual, mas excesso de confiança em inferências não auditadas.

A IA também pode reproduzir vieses presentes em dados de treinamento e no *corpus* analisado. Se literatura em inglês é dominante, se certas regiões são mais representadas, se autores canônicos aparecem com maior frequência, o modelo pode reforçar hierarquias de visibilidade. A neutralidade algorítmica é, portanto, uma suposição perigosa.

A IA deve ser tratada como instrumento sociotécnico. Seu uso envolve pessoas, dados, instituições, plataformas, contratos, políticas de privacidade, versões de modelo e normas de autoria. O pesquisador precisa avaliar não apenas a utilidade das respostas, mas também a compatibilidade do processo com confidencialidade, integridade científica e responsabilidade pública.

O primeiro risco é delegar julgamento. A IA pode apoiar o julgamento, mas não deve decidir se uma conclusão é válida, se uma métrica é justa ou se um ranking deve orientar política institucional. Tais decisões envolvem valores, consequências e conhecimento contextual.

O segundo risco é perder rastreabilidade. Se o pesquisador usa IA para classificar documentos, rotular *clusters* ou redigir interpretações sem registrar *prompt*, versão do modelo, dados de entrada e critérios de validação, a análise torna-se difícil de reproduzir ou auditar.

O terceiro risco é confundir coerência textual com evidência. Uma explicação bem escrita pode ser especulativa. Em estudos cientométricos, afirmações devem ser vinculadas a tabelas, mapas, amostras textuais, parâmetros de análise ou literatura.

Governar o uso de IA inclui definir tarefas permitidas e proibidas, níveis de validação, critérios de revisão humana, documentação dos *prompts* e preservação de saídas intermediárias relevantes. A decisão final sobre validade, justiça ou uso institucional não pode ser delegada ao modelo.

Em uma análise de 5.000 resumos, a IA pode classificar documentos em temas. Entretanto, deve-se validar uma amostra, medir concordância com codificadores humanos quando possível, revisar casos ambíguos e documentar critérios. A classificação não deve ser aceita apenas porque a distribuição final parece plausível.

Na redação de uma discussão, a IA pode sugerir interpretações alternativas e limites metodológicos. O pesquisador deve verificar se cada afirmação está apoiada nos resultados, se não há extrapolação e se não foram introduzidos autores ou fatos ausentes.

18.8 Transparência no uso de prompts

Se *prompts* orientam a forma como a IA lê, classifica, resume e interpreta dados, então eles devem ser tratados como parte do método. Em estudos cientométricos assistidos por IA, transparência não se limita a declarar que a IA foi usada. É necessário informar para que foi usada, com quais dados, com quais *prompts*, em que versão de modelo, sob quais parâmetros, com que validação e com que intervenção humana.

Tratar *prompts* como parte do método aproxima a cientometria assistida por IA das práticas de ciência aberta, relato metodológico e métricas responsáveis. A transparência deve permitir que leitores e pares compreendam o enquadramento adotado, os limites impostos ao modelo e a intervenção humana na produção da análise.

A documentação de *prompts* cumpre várias funções. Primeiro, permite reprodutibilidade aproximada, ainda que modelos generativos possam variar. Segundo, permite auditoria por pares, que podem avaliar se o *prompt* induziu respostas enviesadas ou conclusões fortes demais. Terceiro, protege a responsabilidade autoral, pois evidencia que a IA foi instrumento de apoio e que a decisão final permaneceu humana. Quarto, contribui para o aprendizado metodológico da comunidade, permitindo comparar estratégias de *prompting*.

A transparência deve incluir o texto integral dos *prompts* principais, a descrição dos dados de entrada, os critérios de seleção de exemplos, as instruções de limitação, o formato de saída solicitado, as regras de validação e o modo como as respostas foram incorporadas ao estudo. Quando o *prompt* contém dados sensíveis, pode-se apresentar versão anonimizada ou estrutural, preservando confidencialidade sem apagar o método.

Prompts também devem registrar proibições. Por exemplo: não inventar referências; não afirmar causalidade; não inferir intenção de autores; não generalizar além da base; não usar linguagem avaliativa sem evidência; separar dado, inferência e hipótese; indicar grau

de confiança; marcar ambiguidades para revisão humana. Essas restrições são parte do desenho metodológico.

Além disso, a transparência exige registrar mudanças. Um estudo pode começar com *prompts* exploratórios e evoluir para *prompts* analíticos mais controlados. Essa trajetória deve ser preservada quando influenciar decisões substantivas. O registro não precisa incluir cada interação irrelevante, mas deve documentar *prompts* que afetaram coleta, limpeza, classificação, interpretação, visualização ou redação.

A IA pode apoiar a própria documentação, por exemplo, transformando uma sequência de comandos em protocolo, identificando lacunas de rastreabilidade e sugerindo campos de auditoria. O registro final, entretanto, deve ser revisado pelo pesquisador e vinculado aos arquivos efetivamente usados no estudo.

O risco de não documentar *prompts* é transformar a IA em caixa-preta metodológica. O leitor vê resultados, mas não sabe como classificações foram produzidas ou como interpretações foram orientadas. Isso fragiliza confiança e dificulta reprodução.

Outro risco é o uso ornamental de transparência. Declarar genericamente "utilizou-se ChatGPT para apoio" é insuficiente. Essa frase não informa se a IA foi usada para revisão gramatical, geração de hipóteses, classificação de documentos ou redação de conclusões. Cada uso possui implicações distintas.

Um terceiro risco é ocultar a intervenção humana. Se o pesquisador edita, corrige, combina e descarta saídas da IA, o relatório deve deixar claro que houve curadoria. A transparência não significa atribuir autoria à máquina, mas esclarecer o processo sociotécnico de produção do texto analítico.

Em termos de documentação, recomenda-se construir um apêndice metodológico de IA com tarefa, ferramenta, versão, *prompt*, dados de entrada, saída esperada, validação e decisão final. Em estudos avaliativos, esse registro é especialmente importante, pois os resultados podem afetar reputações e recursos.

Em uma classificação temática, o apêndice pode apresentar o *prompt* central, o esquema de categorias, a amostra usada para validação, a taxa de concordância humana, os critérios de revisão de casos ambíguos e exemplos de documentos classificados com justificativa textual.

Em uma interpretação de mapas, o apêndice pode apresentar o *prompt* usado para nomear *clusters*, as alternativas geradas, o critério de escolha do rótulo final e a evidência textual utilizada para validar o nome escolhido.

18.9 Síntese integradora: da crítica das métricas à governança dos prompts

As questões discutidas neste capítulo convergem para uma ideia central: cientometria é uma prática de mediação. Ela mede por meio de indicadores, representa por meio de mapas, ordena por meio de rankings, interpreta por meio de narrativas e, agora, também dialoga por meio de *prompts*. Em cada uma dessas operações há escolhas que podem ampliar a compreensão da ciência ou empobrecê-la. A diferença entre uso responsável e uso inadequado não está na presença ou ausência de números, mas na relação entre números, perguntas, contexto, método e consequências.

A crítica epistemológica não deve ser confundida com rejeição da mensuração. Sem indicadores, seria difícil observar a expansão de áreas, a distribuição de colaborações, a circulação de temas, a emergência de frentes de pesquisa e a concentração de recursos simbólicos. O problema surge quando a medida deixa de ser ferramenta e passa a ser substituto

do julgamento. Em termos práticos, uma boa análise cientométrica apresenta indicadores como evidências parciais e articuladas, não como verdades finais.

A crítica sociológica lembra que a ciência é produzida por comunidades, instituições, carreiras, hierarquias, recursos e normas. Métricas não apenas observam essa realidade: podem interagir com ela. Quando pesquisadores sabem que serão avaliados por número de artigos, citações ou posição em rankings, podem modificar estratégias de publicação, colaboração e escolha de temas; essa performatividade deve ser considerada em qualquer sistema avaliativo (de Rijcke *et al.*, 2016; Hicks *et al.*, 2015; Muller; de Rijcke, 2017).

A crítica ética introduz a pergunta pelas consequências. Quem será beneficiado ou prejudicado por determinado uso de métricas? Quais formas de contribuição serão reconhecidas? Quais serão invisibilizadas? A análise reforçará desigualdades de idioma, país, gênero, instituição ou disciplina? O uso de IA aumentará transparência ou criará nova opacidade? Essas perguntas não são externas à metodologia; elas fazem parte da validade social de um estudo cientométrico.

A governança dos *prompts* pode ser compreendida como resposta prática a esse desafio. Um *prompt* bem desenhado incorpora cautela, solicita limites, exige separação entre evidência e hipótese, demanda rastreabilidade e convida à revisão crítica. Um *prompt* mal desenhado pede síntese conclusiva sem contexto, transforma tabelas em julgamentos fortes, solicita rankings sem normalização ou induz o modelo a preencher lacunas com conhecimento aparente. Assim, a qualidade ética da análise começa antes da saída do modelo: começa no desenho da pergunta feita à IA.

Além disso, *prompts* devem ser avaliados por sua capacidade de produzir boas perguntas, e não apenas boas respostas. Em cientometria, muitas vezes a contribuição mais relevante da IA é ajudar o pesquisador a perceber o que ainda não foi controlado: cobertura da base, limiar de inclusão, mudança terminológica, viés de idioma, diferença disciplinar, efeitos de carreira, papel de revisões, autocitação, literatura local e fragilidade de inferências. O modelo deve ser usado para ampliar a vigilância metodológica, não para reduzir o esforço crítico.

A adoção de IA também exige alfabetização estatística, bibliométrica e sociológica. Um pesquisador que não compreende o significado de normalização por área, janela de citação, cobertura de base, modularidade, autocitação ou análise de cocitação terá dificuldade de avaliar se a resposta da IA é apropriada. A fluência da resposta não compensa ausência de conhecimento substantivo. Portanto, a formação em cientometria com IA deve combinar competência técnica, repertório teórico e ética da avaliação.

Em relatórios de política científica, esses cuidados tornam-se ainda mais importantes. Resultados cientométricos podem influenciar financiamento, reconhecimento, contratação, encerramento de programas, internacionalização e desenho de agendas. Uma conclusão exagerada pode ter efeitos reais. Por isso, recomenda-se que estudos usados para decisão institucional incluam revisão metodológica independente, validação de dados, consulta a especialistas de área e declaração explícita de limites.

Na prática editorial do livro, este capítulo cumpre papel de freio crítico. Depois de ensinar a formular perguntas, buscar dados, limpar bases, calcular indicadores, construir redes, mapear temas, detectar tendências, classificar literatura, visualizar resultados e redigir relatórios, é necessário lembrar que cada uma dessas etapas pode produzir injustiça cognitiva se não for acompanhada de reflexividade. A IA não elimina esse risco; ela pode tanto reduzi-lo quanto ampliá-lo, dependendo do modo como é utilizada.

A expressão “responsabilidade analítica” sintetiza essa posição. Responsabilidade analítica significa responder pelo recorte, pelos dados, pelos métodos, pelos *prompts*, pelas validações, pelas interpretações e pelas consequências previsíveis do uso dos resultados. Significa também resistir à tentação de transformar uma análise parcial em diagnóstico absoluto. Em estudos cientométricos, a honestidade metodológica é parte do resultado.

18.10 Prompts do capítulo

Os *prompts* a seguir foram elaborados para incorporar os princípios discutidos ao longo do capítulo. Eles devem ser adaptados ao *corpus*, à base de dados, à área científica, à escala de análise e ao objetivo do estudo. Recomenda-se registrar a versão final de cada *prompt* no protocolo ou apêndice metodológico.

Prompt 18.10.1 Prompt para identificar vieses

Atue como especialista em cientometria crítica e métricas responsáveis. Analise o desenho do estudo descrito abaixo e identifique possíveis vieses linguísticos, geográficos, disciplinares, temporais, de base de dados, de tipo documental, de autoria, de citação e de visualização.

Para cada viés, apresente:

(1) descrição do risco;

(2) etapa em que pode ocorrer;

(3) consequência interpretativa;

(4) estratégia de mitigação;

(5) informação adicional necessária para avaliar sua gravidade.

Não afirme que um viés existe sem evidência; diferencie risco plausível, evidência observada e hipótese a verificar. Dados do estudo: [INSERIR DESCRIÇÃO].

Prompt 18.10.2 Prompt para revisar implicações éticas

Atue como revisor ético de um estudo cientométrico assistido por IA.

Examine os resultados e conclusões abaixo considerando possíveis efeitos sobre pesquisadores, grupos, instituições, países, áreas do conhecimento e políticas científicas.

Indique se há risco de uso indevido de métricas, comparação injusta, reforço de desigualdades, inferência causal não justificada, apagamento de produção local, exposição indevida de indivíduos ou excesso de confiança em saídas de IA.

Para cada risco, proponha uma redação mais prudente e uma ação metodológica de mitigação.

Resultados e conclusões: [INSERIR TEXTO].

Prompt 18.10.3 Prompt para construir seção de limitações

A partir da metodologia e dos resultados abaixo, redija uma seção de limitações para artigo ou relatório cientométrico.

A seção deve abordar:

- base de dados;
- estratégia de busca;
- recorte temporal;
- idioma;
- tipo documental;
- cobertura disciplinar;
- qualidade de metadados;
- tratamento de duplicatas;
- indicadores usados;
- parâmetros de rede;
- visualizações;
- uso de IA;
- validação humana.

Separe limitações que afetam cobertura, limitações que afetam interpretação e limitações que afetam generalização. Não use linguagem defensiva; escreva em tom científico e transparente.

Metodologia e resultados: [INSERIR CONTEÚDO].

Prompt 18.10.4 Prompt para avaliar se uma conclusão é forte demais

Atue como auditor de inferências cientométricas.

Examine as conclusões abaixo e classifique cada uma como:

(A) diretamente sustentada pelos dados;

(B) inferência plausível, mas dependente de contexto;

(C) hipótese que exige evidência adicional;

(D) extrapolação indevida;

(E) afirmação normativa que requer justificativa ética ou política.

Para cada conclusão, explique a classificação, indique que dado seria necessário para fortalecê-la e proponha uma versão mais prudente.

Conclusões: [INSERIR LISTA].

Prompt 18.10.5 Prompt para declarar uso de IA e prompts

Com base nas informações abaixo, redija uma declaração metodológica transparente sobre o uso de IA no estudo cientométrico.

Inclua:

- ferramenta e versão quando disponível;
- tarefas realizadas; *prompts* principais;
- tipo de dados fornecidos ao modelo;
- procedimentos de validação humana;
- decisões tomadas pelos pesquisadores;
- limitações do uso de IA;
- afirmação de responsabilidade autoral.

Não atribua à IA autoria científica. Informações: [INSERIR REGISTRO].

Prompt 18.10.6 Prompt para separar dado, inferência, hipótese e recomendação

Analise o trecho abaixo e reestruture a interpretação em quatro categorias: dado observado, inferência interpretativa, hipótese que exige evidência adicional e recomendação prática.

Indique também quais afirmações podem ser mantidas, quais devem ser enfraquecidas e quais devem ser removidas por falta de suporte.

Trecho: [INSERIR TEXTO].

Quadro 18.2 - Transparência e validação no uso de inteligência artificial

Uso da IA	Exigência mínima de transparência	Validação recomendada	Risco se não houver controle
Expansão de vocabulário	Registrar *prompt* e termos aceitos/rejeitados	Revisão por especialista e teste em base	Inclusão de termos inexistentes ou ruído excessivo
Classificação temática	Registrar categorias, *prompt*, amostra e critérios	Amostragem dupla e revisão de ambíguos	Categorias artificiais ou enviesadas
Rotulagem de clusters	Registrar dados do *cluster* e alternativas de rótulo	Confronto com títulos, resumos e especialistas	Rótulos sedutores, mas imprecisos
Interpretação de indicadores	Registrar tabela de entrada e regras de inferência	Checagem contra literatura e contexto disciplinar	Extrapolação de qualidade ou causalidade
Redação de discussão	Registrar uso como apoio redacional ou analítico	Verificação de todas as afirmações substantivas	Introdução de fatos ou referências inventadas

Fonte: Elaborado pelo autor (2026).

18.11 Estudo de caso ilustrativo: avaliando a ética de um mapa cientométrico sobre IA na saúde

Imagine-se um estudo destinado a mapear a produção científica sobre inteligência artificial aplicada à saúde no período de 2015 a 2025. A busca foi realizada em Scopus e Web of Science, com termos em inglês, e os resultados foram usados para gerar indicadores de produção, redes de coautoria, mapas de coocorrência de palavras-chave e ranking de instituições. Em seguida, um modelo de linguagem foi usado para nomear *clusters*, redigir a interpretação e produzir um sumário executivo para gestores.

À primeira vista, o desenho parece metodologicamente aceitável. As bases são reconhecidas, o período cobre uma fase de crescimento da área, os indicadores são comuns e a IA foi usada para tarefas compatíveis com apoio interpretativo. Contudo, uma auditoria epistemológica e ética revela questões importantes. O uso exclusivo de termos em inglês pode excluir literatura relevante em saúde pública local. A escolha de bases internacionais pode sub-representar periódicos regionais e documentos de política. Rankings institucionais podem favorecer universidades grandes e países com maior financiamento. Redes de coautoria podem não revelar assimetrias de liderança em consórcios internacionais. Mapas de palavras-chave podem refletir terminologia de indexação mais do que conteúdo substantivo dos estudos.

A IA, ao nomear *clusters*, poderia identificar um grupo como "diagnóstico automatizado de doenças", outro como "medicina personalizada" e outro como "ética e governança". Esses rótulos podem ser úteis, mas devem ser validados. É necessário verificar títulos, resumos e, se possível, amostras de textos completos. Um *cluster* com termos como "*bias*", "*fairness*", "*ethics*" e "*explainability*" pode incluir tanto estudos normativos quanto estudos técnicos de mitigação de viés. Um rótulo genérico apagaria diferenças importantes.

Na redação para gestores, o risco aumenta. A frase "o país X lidera a pesquisa mundial em IA na saúde" é forte demais se baseada apenas em contagem de publicações indexadas. Uma versão mais prudente seria: "No *corpus* composto por registros da Scopus e da Web of Science, o país X apresentou o maior volume de publicações indexadas no período analisado; essa posição deve ser interpretada à luz da cobertura das bases, do uso de termos em inglês e da ausência de normalização por tamanho do sistema científico". A diferença entre as frases é ética e metodológica.

O estudo de caso mostra que a IA pode ajudar a revelar riscos se for usada como auditora crítica. Um *prompt* poderia pedir ao modelo que identificasse vieses de cobertura, inferências fortes demais, termos ambíguos, limites de ranking e lacunas de documentação. Nesse uso, a IA atua como ferramenta de reflexividade. Mas, se solicitada apenas a "escrever uma interpretação convincente", ela pode aumentar o risco de simplificação.

18.12 Checklist de qualidade epistemológica, sociológica e ética

Quadro 18.3 - Checklist de qualidade epistemológica, sociológica e ética

CHECKLIST DE QUALIDADE *(antes de fechar o estudo relacional assistido por IA)*
☐ Pergunta de pesquisa clara e alinhada ao método.
☐ *Corpus* bem definido, reproduzível e documentado.
☐ Construção da matriz relacional transparente.
☐ Parâmetros, limiares e algoritmos informados.

CHECKLIST DE QUALIDADE *(antes de fechar o estudo relacional assistido por IA)*
☐ Análises de sensibilidade realizadas.
☐ Interpretações separadas de evidências.
☐ Limitações e vieses explicitados.
☐ Documentação dos *prompts* e das saídas da IA.
☐ Revisão humana e validação qualitativa.
☐ Conclusões proporcionais às evidências.

Registro da verificação

Estudo ou etapa verificada	Responsável	Data

Fonte: Elaborado pelo autor (2026).

18.13 Orientações para redação ética de resultados cientométricos

A redação de resultados cientométricos deve evitar verbos e adjetivos que atribuam valor além do que os dados sustentam. Expressões como "melhor", "superior", "mais relevante" e "líder incontestável" exigem critérios robustos. Em muitos casos, é mais adequado usar "maior volume de publicações indexadas", "maior número de citações no *corpus*", "maior centralidade na rede analisada" ou "maior presença nos documentos recuperados". Essa precisão protege o estudo contra extrapolação.

Também se recomenda explicitar a escala da inferência. Uma afirmação pode ser válida no nível agregado e inválida no nível individual. Um departamento pode apresentar alta média de citação sem que todos os seus pesquisadores tenham desempenho semelhante. Um país pode aparecer central em coautorias internacionais sem liderar intelectualmente os projetos. Um periódico pode ser central em uma rede de cocitação sem que todos os seus artigos tenham impacto equivalente.

A redação deve ainda diferenciar resultados descritivos de interpretações. Resultados descritivos apresentam tabelas, frequências, mapas e padrões observados. Interpretações propõem sentidos possíveis para esses padrões. Discussões relacionam interpretações com literatura, contexto e limitações. Recomendações traduzem achados em possibilidades de ação. Misturar esses níveis torna o relatório menos transparente.

Quando IA for usada para redigir, convém solicitar ao modelo que marque afirmações que dependem de evidência externa. Essa técnica transforma o modelo em dispositivo de autocontrole textual. Por exemplo, após gerar uma discussão, o pesquisador pode pedir: "liste todas as frases que afirmam causalidade, liderança, qualidade ou impacto social e indique que evidência as sustenta". Esse procedimento frequentemente revela pontos que precisam de enfraquecimento.

A seção de limitações não deve ser tratada como formalidade. Em cientometria, limitações são parte essencial da interpretação. Elas dizem ao leitor em que universo a análise opera, quais dimensões ficaram fora, que comparações são frágeis e que conclusões devem ser lidas com cautela. Uma limitação bem escrita aumenta, e não diminui, a credibilidade do estudo.

18.14 Protocolo de auditoria crítica de indicadores e mapas

Uma forma prática de operacionalizar as discussões deste capítulo é adotar um protocolo de auditoria crítica antes da divulgação de qualquer resultado cientométrico. Esse protocolo deve ser aplicado a indicadores, mapas, rankings, *dashboards* e textos interpretativos. A auditoria não tem a função de paralisar a análise, mas de tornar explícitas as condições sob as quais os resultados podem ser usados. Em vez de perguntar apenas se o cálculo está correto, pergunta-se se o cálculo é adequado à pergunta, se o dado é suficiente, se a interpretação respeita a escala e se as consequências potenciais foram consideradas.

Quadro 18.4 - Protocolo de auditoria crítica de indicadores e mapas

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
01 Cadeia de produção do dado	Reconstruir todo o percurso que gerou os dados e os outputs apresentados.	• De onde veio o dado? • Quando foi coletado? • Como foi selecionado e tratado?	• Base e cobertura • Data de coleta e extração • Estratégia de busca (string) • Campos utilizados • Tipos documentais e idiomas • Critérios de inclusão e exclusão • Deduplicação e normalizações • Transformações e agregações • Ferramentas, versões e parâmetros • *Prompts* da IA (tarefa, versão, data) • Saídas da IA aceitas e rejeitadas • Validação das saídas

Riscos comuns	Decisão / ações	OK?
• Dados incompletos ou não rastreáveis • Filtros não declarados • Deduplicação incorreta • Métodos não registrados • Uso da IA sem registro • Saídas aceitas sem validação	• Solicitar informações faltantes • Corrigir ou complementar o registro • Padronizar e documentar • Registrar *prompts* e versões • Validar saídas da IA • Reexecutar se necessário	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
02 Congruência entre indicador e pergunta	Verificar se o indicador ou método responde à pergunta formulada.	• O indicador mede o que a pergunta exige? • Quais são os limites do indicador? • O texto respeita esses limites?	• Definição do indicador e do unidade de análise • Tipo de relação (citação, coautoria, cocitação, coocorrência, acoplamento etc.) • Escala de interpretação (descritiva, comparativa, avaliativa) • Pressupostos do método • Fronteiras de inferência

Riscos comuns	Decisão / ações	OK?
• Confundir produção com qualidade • Tratar citação como prova de qualidade • Inferir causalidade de associações • Extrapolar funcionalidade do método	• Ajustar indicador ou método • Explicitar limites no texto • Reformular interpretação • Readequar visualização	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
03 Interpretações alternativas	Examinar explicações alternativas para os padrões observados.	• Que outras explicações são plausíveis? • Que dados poderiam distingui-las?	• Lista de hipóteses alternativas • Evidências a favor e contra cada hipótese • Verificações em fontes comparativas (textos, dados adicionais) • Limitações que impedem discriminação

Riscos comuns	Decisão / ações	OK?
• Interpretar única hipótese como causa única • Ignorar mudanças de base ou de cobertura • Viés de confirmação	• Coletar dados adicionais • Realizar testes comparativos • Temperar conclusões • Declarar incerteza	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
04 Sensibilidade a parâmetros	Avaliar a estabilidade dos resultados frente a escolhas metodológicas.	• O resultado muda com diferentes parâmetros? • Quais escolhas são mais sensíveis?	• Testes com limiares diferentes • Algoritmos e medidas de normalização alternativas • Janelas e amostras (quando aplicável) • Cenários comparados e publicados ou descritos • Relatório de *clusters* por parâmetro e limiares

Riscos comuns	Decisão / ações	OK?
• Conclusões dependentes de um único parâmetro • *Cluster* instável • Ranking frágil • Layout interpretado como evidência	• Reportar testes de sensibilidade • Escolher cenário mais robusto ou justificar a escolha • Explicitar instabilidade	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
05 Linguagem avaliativa	Garantir precisão e prudência na redação interpretativa.	• A linguagem é neutra e precisa? • Há termos avaliativos não justificados?	• Termos usados para descrever achados • Correspondência entre termo e evidência • Substituição por linguagem descritiva quando apropriado • Indicação do nível de incerteza

Riscos comuns	Decisão / ações	OK?
• Uso de rótulos normativos (melhor, líder, excelência, relevância) • Exageros de impacto • Julgamentos não sustentados	• Ajustar linguagem • Substituir termos fortes por descrições metodológicas • Declarar grau de certeza	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
06 Justiça comparativa	Verificar equidade e adequação das comparações feitas.	• Comparações são justas? • Há ajustes de contexto suficientes?	• Normalizações aplicadas (por tamanho, área, período etc.) • Diferenças disciplinares (cultura de citação) • Diferenças institucionais (missão, recursos) • Diferenças de contexto geográfico e linguístico • Finalidade da comparação (descritiva, diagnóstica, avaliativa)

Riscos comuns	Decisão / ações	OK?
• Favorecer grandes instituições ou países • Penalizar áreas com menor cultura de citação • Reforçar desigualdades estruturais	• Aplicar normalizações adequadas • Contextualizar interpretações • Evitar rankings absolutos • Declarar limites da comparação	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
07 Validação por especialistas	Incorporar julgamento especializado e conhecimento situado.	• Especialistas validariam interpretações e rótulos? • Houve consulta ao contexto local?	• Lista de especialistas consultados • Método de consulta (entrevistas, survey, reunião) • Principais contribuições e ajustes acordados • Rótulos ou temas validados ou contestados • Evidências contextuais não captadas nos metadados

Riscos comuns	Decisão / ações	OK?
• Rótulos artificiais ou distantes • Ignorar contexto institucional • Perder nuance conceitual • Assumir erros interpretativos	• Revisar rótulos e interpretações • Incluir evidências contextuais • Documentar contribuições • Ajustar narrativa	☐

Etapa	Objetivo da auditoria	Questões-chave	Evidências / o que verificar
08 Decisão final e documentação	Registrar decisões e garantir auditabilidade do estudo.	• Decisão final está clara? • Riscos foram considerados? • Registro é completo?	• Decisão por adotar, manter, reformular, remover, coletar mais dados ou modificar • Justificativa da decisão • Riscos residuais e limites comunicados • Versão final do texto e das figuras • Log de mudanças

Riscos comuns	Decisão / ações	OK?
• Decisões implícitas • Falta de rastreabilidade • Remoção de achados sem justificativa • Uso de achados para fins não indicados	• Documentar decisão e justificativa • Atualizar texto e figuras • Registrar riscos residuais • Arquivar evidências	☐

Fonte: Elaborado pelo autor (2026).

Esse protocolo pode ser aplicado como checklist interno, apêndice metodológico ou etapa formal de revisão por pares. Sua adoção aproxima a cientometria assistida por IA de práticas de governança de dados, ciência aberta e avaliação responsável. Mais do que acrescentar burocracia, ele cria memória analítica: permite saber por que um indicador foi escolhido, por que um *prompt* foi usado, porque uma conclusão foi mantida e por que certos limites foram reconhecidos.

18.15 Justiça cognitiva, diversidade epistêmica e invisibilidade bibliométrica

A discussão ética da cientometria precisa ir além da correção estatística. Mesmo um indicador calculado corretamente pode produzir injustiça se comparar realidades incomparáveis, desconsiderar missões institucionais ou tornar invisíveis formas legítimas de conhecimento. Vieses de indexação e rankings podem transformar assimetrias de visibilidade em hierarquias de valor (CHAVARRO *et al.*, 2018; MONGEON; PAUL-HUS, 2016; RAFOLS *et al.*, 2012).

Uma dimensão importante é a diversidade linguística. A hegemonia do inglês facilita circulação internacional, mas também cria barreiras para pesquisadores que produzem em outras línguas ou que investigam problemas locais. Em algumas áreas, publicar em língua local é condição para dialogar com profissionais, comunidades, professores, gestores públicos ou movimentos sociais. Se a avaliação valoriza apenas a publicação internacional em inglês, pode desestimular a comunicação com públicos que mais se beneficiariam da pesquisa.

Outra dimensão é a diversidade de formatos. Livros, capítulos, relatórios técnicos, pareceres, softwares, bases de dados, protocolos, materiais educativos, normas, exposições, curadorias e intervenções públicas podem ser produtos relevantes. Muitas dessas formas são centrais em humanidades, ciências sociais aplicadas, educação, saúde coletiva, arquitetura, direito, artes e tecnologias sociais. Um estudo cientométrico que as exclui deve reconhecer que está analisando apenas parte da contribuição científica.

A invisibilidade bibliométrica também pode afetar grupos historicamente marginalizados. Pesquisadores de instituições periféricas, mulheres, grupos racializados, pesquisadores em início de carreira, comunidades tradicionais e colaboradores técnicos podem aparecer menos nos indicadores por razões estruturais. O problema não se resolve apenas com normalização estatística, pois envolve distribuição de oportunidades, redes de reconhecimento, acesso a financiamento, domínio de idiomas e legitimidade institucional.

A IA pode reproduzir essas invisibilidades. Modelos treinados em *corpora* desbalanceados tendem a reconhecer melhor autores, instituições, periódicos e conceitos dominantes. Ao sugerir referências, rótulos ou interpretações, podem reforçar cânones já estabelecidos e ignorar contribuições locais. Por isso, *prompts* críticos devem solicitar explicitamente atenção a literatura periférica, termos em outras línguas, produção regional, diversidade de formatos e ausências do *corpus*.

No entanto, também é possível usar IA para ampliar a sensibilidade à diversidade epistêmica. Um modelo pode ajudar a traduzir termos entre tradições disciplinares, identificar sinônimos regionais, sugerir bases complementares, comparar vocabulários e revisar linguagem avaliativa. Pode ainda ajudar a formular perguntas como: que comunidades não aparecem neste *corpus*? que tipos de documento foram excluídos? que idiomas estão ausentes? que públicos são afetados pela interpretação proposta? Essas perguntas transformam a IA em instrumento de crítica, não de autoridade.

A justiça cognitiva exige que o relatório cientométrico explique sua relação com o território analisado. Em um estudo sobre pesquisa brasileira, por exemplo, o uso exclusivo

de bases internacionais deve ser justificado e complementado quando necessário. Em um estudo sobre políticas públicas, documentos governamentais e literatura técnica podem ser indispensáveis. Em um estudo sobre inovação social, publicações acadêmicas podem ser apenas um dos sinais possíveis de contribuição.

A ética da cientometria também envolve evitar a desqualificação automática do que não é indexado. Ausência em uma base pode significar baixa visibilidade internacional, mas também pode refletir orientação local, barreira de idioma, exclusão editorial, juventude do periódico, formato documental inadequado à base ou escolha estratégica de comunicação. A redação deve evitar frases que convertam invisibilidade bibliográfica em inexistência científica.

A noção de diversidade epistêmica não implica que todos os documentos tenham a mesma função ou que qualquer fonte deva ser aceita sem critério. Implica que os critérios devem ser alinhados à pergunta. Se a pergunta é sobre circulação em periódicos internacionais, bases seletivas podem ser adequadas. Se a pergunta é sobre contribuição para políticas locais, o desenho precisa incluir fontes capazes de capturar essa contribuição. Justiça metodológica é adequação entre objeto, evidência e finalidade.

Em síntese, uma cientometria com IA comprometida com justiça cognitiva deve perguntar não apenas "o que aparece nos dados?", mas também "por que aparece?", "quem não aparece?", "que canais foram reconhecidos?", "que formas de conhecimento foram traduzidas em indicadores?" e "que consequências essa representação pode produzir?". Essas perguntas tornam a análise mais complexa, porém também mais honesta e socialmente responsável.

18.16 Matriz decisória para conclusões cientométricas assistidas por IA

Ao final de um estudo cientométrico, a equipe costuma enfrentar uma passagem delicada: converter resultados em conclusões. Essa passagem é onde muitos erros éticos e epistemológicos se concentram. Dados descritivos tornam-se narrativas de liderança; associações tornam-se causalidades; lacunas de cobertura tornam-se lacunas de produção; diferenças de campo tornam-se diferenças de qualidade. Para reduzir esse risco, recomenda-se usar uma matriz decisória antes de finalizar conclusões assistidas por IA.

A primeira pergunta da matriz é: a conclusão é diretamente observável nos dados? Se a tabela mostra que uma instituição possui maior número de artigos no *corpus*, é possível afirmar isso. Não é possível, sem outras evidências, afirmar que a instituição é a mais inovadora, mais qualificada ou mais influente socialmente. O grau de força da conclusão deve acompanhar o grau de evidência.

A segunda pergunta é: a conclusão depende de normalização? Comparações por país, instituição, área ou autor frequentemente dependem de tamanho, tempo de carreira, campo, tipo de documento e cobertura. Se a análise usa números absolutos, a conclusão deve mencionar números absolutos. Se pretende falar de desempenho relativo, precisa justificar normalização. A IA deve ser instruída a verificar esse ponto antes de escrever frases avaliativas.

A terceira pergunta é: a conclusão é sensível à base escolhida? Se estudos anteriores mostram que Scopus, Web *of Science,* Dimensions, *Lens* e OpenAlex possuem coberturas diferentes, conclusões sobre produção global precisam declarar a fonte. Quando possível, a matriz recomenda comparar resultados em mais de uma base ou, ao menos, discutir como a escolha pode afetar o achado.

A quarta pergunta é: a conclusão transforma correlação em causalidade? Uma coocorrência de palavras não prova relação causal entre conceitos. Uma cocitação não prova concordância. Uma coautoria não prova colaboração simétrica. Crescimento temporal não

prova impacto de política específica sem desenho adicional. A matriz deve marcar qualquer frase causal e exigir evidência independente.

A quinta pergunta é: a conclusão tem implicações avaliativas para pessoas ou instituições? Quanto maior o potencial de consequência, maior deve ser a exigência de validação. Um mapa exploratório para orientar leitura pode admitir mais incerteza do que um parecer usado para distribuir recursos. A ética da evidência é proporcional ao efeito previsto da decisão.

A sexta pergunta é: a IA introduziu conteúdo não presente nos dados? Modelos de linguagem podem enriquecer texto com conhecimentos gerais, exemplos e explicações. Isso pode ser útil, mas também pode introduzir afirmações não verificadas. Antes de publicar, deve-se comparar a interpretação com os resultados efetivamente gerados. Toda referência, dado histórico, nome de autor e explicação causal sugerida pela IA deve ser verificada.

A sétima pergunta é: há uma versão mais prudente da conclusão? Em muitos casos, a solução não é remover o achado, mas ajustar sua redação. Em vez de "o tema emergiu devido ao avanço de políticas de financiamento", pode-se escrever "o crescimento do tema coincide temporalmente com políticas de financiamento, mas o estudo não permite inferir causalidade". A conclusão permanece informativa e ganha honestidade.

A oitava pergunta é: a seção de limitações conversa com as conclusões? Limitações não devem aparecer isoladas no fim do texto. Elas devem modular a força das conclusões. Se a limitação é cobertura restrita a artigos em inglês, conclusões sobre literatura mundial devem ser evitadas. Se a limitação é ausência de análise de conteúdo, conclusões sobre escolas teóricas devem ser prudentes. Se a limitação é uso de IA para classificação, deve-se informar validação.

A matriz decisória pode ser aplicada com o apoio de IA. O pesquisador pode fornecer a lista de conclusões e solicitar que o modelo classifique cada uma por nível de evidência. O ganho está em tornar visíveis os saltos inferenciais. Contudo, a decisão final deve permanecer humana, pois envolve conhecimento da área, leitura dos dados e avaliação das consequências.

Quadro 18.5 - Níveis de força das afirmações cientométricas

Nível da afirmação	Exemplo de redação	Condição mínima para uso	Ação recomendada
Descrição direta	A produção indexada cresceu de 120 para 310 documentos.	Dado apresentado em tabela ou gráfico.	Manter com fonte e período.
Inferência contextual	O crescimento sugere maior consolidação do tema no *corpus* analisado.	Padrão temporal consistente e discussão de cobertura.	Manter com linguagem prudente.
Hipótese explicativa	O aumento pode estar associado a políticas de financiamento.	Evidência externa ou análise complementar ainda ausente.	Marcar como hipótese e indicar dados necessários.
Causalidade forte	A política X causou o crescimento da área.	Desenho causal ou evidência documental robusta.	Evitar se o estudo for apenas cientométrico descritivo.
Juízo avaliativo	A instituição Y é a melhor do campo.	Critérios múltiplos, normalização e avaliação qualitativa.	Substituir por descrição limitada ao indicador.

Fonte: Elaborado pelo autor (2026).

18.17 Governança institucional do uso de métricas e IA

A reflexão ética não se completa no nível individual do pesquisador. Estudos cientométricos frequentemente circulam em universidades, agências de fomento, editoras, hospitais, institutos de pesquisa, programas de pós-graduação e órgãos governamentais. Nesses ambientes, indicadores e análises assistidas por IA podem adquirir função administrativa. Por isso, é necessário pensar em governança institucional do uso de métricas e de sistemas generativos.

Governança institucional significa estabelecer regras, responsabilidades, limites e procedimentos de revisão antes que métricas sejam incorporadas a decisões. Uma instituição que utiliza cientometria para planejamento deve definir se os resultados terão finalidade exploratória, diagnóstica, avaliativa ou distributiva. Cada finalidade exige nível distinto de evidência. Um painel exploratório pode orientar perguntas; um parecer avaliativo pode afetar carreiras e, portanto, requer maior transparência, revisão e possibilidade de contestação.

A governança também deve impedir a monocultura métrica. Quando uma organização adota poucos indicadores como eixo principal de avaliação, pesquisadores tendem a ajustar comportamentos a esses sinais. Isso pode reduzir a diversidade de agendas, favorecer publicações de curto prazo, estimular fragmentação de resultados e desvalorizar ensino, extensão, curadoria, reprodução de estudos, infraestrutura e diálogo social. Métricas devem compor um ecossistema de evidências, e não um mecanismo automático de julgamento (AMERICAN SOCIETY FOR CELL BIOLOGY, 2012; HICKS *et al.*, 2015).

No caso da IA, a governança deve definir tarefas permitidas, tarefas de uso restrito e tarefas proibidas. Pode ser aceitável usar IA para revisar clareza de texto, sugerir limitações ou organizar categorias preliminares. Pode ser restrito usar IA para classificar produção de pesquisadores individuais sem validação. Deve ser proibido usar IA para inventar pareceres, atribuir intenção a autores, gerar referências não verificadas ou tomar decisões avaliativas automáticas. Essas regras precisam ser públicas para os envolvidos.

Outro aspecto é a contestabilidade. Pessoas e unidades avaliadas devem ter condições de compreender os dados usados, corrigir erros de afiliação, apontar lacunas de cobertura e questionar interpretações. A cientometria responsável não se encerra no cálculo; ela inclui diálogo sobre o significado dos resultados. Quando IA participa da análise, a contestabilidade requer acesso ao registro de *prompts*, critérios de validação e versões de saída que sustentaram o parecer.

Instituições também devem investir em formação. Gestores que recebem *dashboards* precisam entender diferença entre produção, impacto, colaboração, visibilidade e qualidade. Pesquisadores que utilizam IA precisam saber documentar *prompts*, verificar referências, reconhecer vieses e evitar extrapolações. Sem formação, ferramentas sofisticadas podem gerar decisões simplistas. A alfabetização em métricas e IA é condição de governança democrática.

Quadro 18.6 - Elementos de governança institucional

Elemento de governança	Pergunta orientadora	Procedimento recomendado
Finalidade do estudo	O resultado será exploratório, diagnóstico, avaliativo ou distributivo?	Declarar finalidade antes da análise e ajustar rigor metodológico.
Transparência	Dados, métodos, *prompts* e validações podem ser auditados?	Manter apêndice metodológico e registro de versões.
Contestabilidade	A unidade analisada pode corrigir dados e questionar interpretações?	Prever etapa de devolutiva e revisão contextual.

Elemento de governança	Pergunta orientadora	Procedimento recomendado
Pluralidade de evidências	A decisão depende de uma métrica isolada?	Combinar indicadores, parecer qualitativo e missão institucional.
Limites da IA	Quais tarefas não podem ser delegadas ao modelo?	Definir política de uso permitido, restrito e proibido.

Fonte: Elaborado pelo autor (2026).

A política institucional deve incluir proteção contra usos punitivos de indicadores frágeis. Resultados cientométricos podem revelar problemas reais, mas não devem ser convertidos automaticamente em sanções. Baixa visibilidade em uma base pode decorrer de missão local, produção técnica, mudança de área, barreiras linguísticas ou desigualdade de recursos. Antes de qualquer decisão, é necessário triangulação com avaliação qualitativa, análise contextual e escuta dos envolvidos.

Por fim, governança implica revisão periódica. Bases de dados mudam, modelos de IA mudam, políticas editoriais mudam e práticas científicas mudam. Um protocolo considerado adequado em determinado momento pode tornar-se insuficiente. Assim, recomenda-se que instituições revisem seus critérios de uso de métricas e IA, publiquem orientações atualizadas, monitorem efeitos indesejados e abram canais de crítica. A ética da cientometria é processo contínuo, não declaração estática.

18.18 Critérios mínimos para publicação de estudos cientométricos com IA

Antes de submeter artigo, relatório ou parecer cientométrico que tenha utilizado IA, recomenda-se verificar critérios mínimos de publicação. O primeiro é explicitar a finalidade analítica: descrever um campo, apoiar planejamento, comparar instituições, identificar lacunas, avaliar desempenho ou orientar política. Sem essa definição, leitores não conseguem julgar se os indicadores e o nível de validação são adequados.

O segundo critério é a rastreabilidade dos dados. Devem estar descritos base, data de coleta, estratégia de busca, campos exportados, filtros, formatos, critérios de limpeza e procedimentos de deduplicação. Quando restrições contratuais impedirem disponibilização integral dos dados, ainda assim é possível disponibilizar scripts, listas de campos, estratégias de busca e amostras anonimizadas. Transparência parcial é preferível à ausência de transparência.

O terceiro critério é a rastreabilidade das intervenções de IA. O texto deve informar se a IA foi usada para busca, classificação, sumarização, rotulagem, interpretação, visualização ou redação. Para tarefas analíticas, é recomendável apresentar *prompts* centrais e procedimentos de validação. Para tarefas meramente redacionais, deve-se declarar que a responsabilidade final pelo conteúdo permanece com os autores.

O quarto critério é a coerência inferencial. Conclusões devem derivar dos resultados e devem respeitar seus limites. Se o estudo é descritivo, não deve afirmar causalidade. Se o *corpus* é restrito a uma base, não deve falar da totalidade da ciência mundial. Se a classificação foi feita por IA, deve indicar validação humana. Se o mapa é exploratório, deve evitar recomendações institucionais fortes sem triangulação adicional.

O quinto critério é a presença de uma seção de limitações substantiva. Essa seção deve mencionar cobertura de bases, idioma, tipo documental, qualidade de metadados, escolhas

de indicadores, parâmetros de visualização, uso de IA, incerteza de classificação e riscos de interpretação. Uma limitação bem escrita não enfraquece o estudo; ela define as condições sob as quais o conhecimento produzido é confiável.

O sexto critério é a prudência ética. Quando resultados envolverem pessoas, grupos ou instituições identificáveis, o texto deve evitar julgamentos simplificadores. A análise deve favorecer aprendizado, diagnóstico e melhoria, e não exposição injusta. A cientometria com IA será mais legítima quando combinar poder computacional com responsabilidade interpretativa, respeito à diversidade científica e compromisso com a transparência pública.

Em termos pedagógicos, esses critérios também funcionam como guia de formação. Estudantes e equipes iniciantes podem usá-los para revisar projetos antes da coleta de dados; avaliadores podem usá-los para examinar manuscritos; gestores podem usá-los para decidir se um relatório é suficientemente robusto para orientar políticas. A padronização desses critérios fortalece a confiança na cientometria assistida por IA sem reduzir sua flexibilidade metodológica.

Assim, a qualidade do capítulo não se mede pela extensão do texto, mas pela capacidade de converter crítica epistemológica em rotinas práticas de pesquisa, documentação, validação e comunicação responsável.

18.19 Conclusão

Este capítulo argumentou que a cientometria com inteligência artificial exige dupla competência: competência técnica para medir, mapear, visualizar e interpretar dados científicos; e competência crítica para reconhecer os limites epistemológicos, sociológicos e éticos dessas operações. Indicadores mostram padrões, mas escondem contextos. Publicações registram parte da ciência, mas não toda a ciência em andamento. Citações sinalizam circulação e reconhecimento, mas também retórica, convenção e estratégia. Bases bibliográficas ampliam a observação, mas introduzem vieses de idioma, geografia, disciplina e cobertura. Mapas e rankings ajudam a comunicar estruturas, mas podem produzir falsa neutralidade. Modelos de IA apoiam análise e redação, mas trazem opacidade, alucinação e risco de autoridade artificial.

A resposta a esses desafios não é abandonar métricas nem rejeitar IA. A resposta é usar ambos com responsabilidade. Isso significa documentar dados, métodos, parâmetros, *prompts* e validações. Significa separar dado, inferência, hipótese e recomendação. Significa evitar generalizações indevidas. Significa reconhecer que avaliação científica envolve valores, missões, contextos e consequências. Significa, finalmente, compreender que a inteligência artificial não substitui julgamento científico; ela amplia a necessidade de julgamento.

Na trilha deste livro, o capítulo funciona como eixo de reflexividade. Todas as técnicas apresentadas nos capítulos anteriores podem ser fortalecidas quando incorporam os *prompts* críticos aqui propostos. Uma cientometria assistida por IA será tanto mais científica quanto mais transparente forem suas mediações, quanto mais explícitos forem seus limites e quanto mais responsável for sua interpretação. O objetivo não é produzir mapas mais impressionantes ou relatórios mais persuasivos, mas análises mais justas, rastreáveis e epistemologicamente honestas.

Apêndice A.18 - Modelo de registro ético do uso de IA em estudo cientométrico

O registro ético do uso de IA deve ser preenchido antes da redação final do relatório. Ele não substitui o protocolo metodológico, mas complementa a documentação da pesquisa. Sua finalidade é permitir que leitores, avaliadores e gestores compreendam em que momentos a IA interferiu no processo analítico e quais salvaguardas foram adotadas.

- Campo 1 - Tarefa: descrever se a IA foi usada para expansão de termos, limpeza, classificação, rotulagem, interpretação, visualização, redação, revisão ou auditoria.
- Campo 2 - Dados de entrada: indicar se foram fornecidos títulos, resumos, tabelas, mapas, palavras-chave, resultados estatísticos ou trechos de relatório.
- Campo 3 - *Prompt*: registrar a versão final do comando.
- Campo 4 - Saída: descrever o tipo de resultado produzido.
- Campo 5 - Validação: indicar como a saída foi verificada.
- Campo 6 - Decisão humana: registrar o que foi aceito, modificado ou rejeitado.
- Campo 7 - Limites: indicar o que a IA não avaliou. Por exemplo, o modelo pode ter classificado resumos, mas não leu textos completos; pode ter sugerido rótulos de *clusters*, mas não calculou a rede; pode ter redigido uma síntese, mas não verificou DOI de referências.
- Campo 8 - Risco residual: registrar riscos que permanecem mesmo após validação.
- Campo 9 - Responsável: indicar pesquisador ou equipe responsável pela decisão final.

Quadro 18.7 - Registro ético do uso de IA em estudo cientométrico

Campo	Descrição a registrar	Exemplo
Tarefa da IA	Função desempenhada no estudo	Rotulagem preliminar de *clusters* temáticos
Dados de entrada	Material fornecido ao modelo	Lista de 40 palavras-chave por *cluster* e 20 títulos representativos
Prompt utilizado	Comando integral ou versão anonimizada	Nomeie o *cluster* e apresente 3 alternativas com justificativa
Saída incorporada	Parte da resposta usada no relatório	Rótulo final e justificativa resumida
Validação humana	Procedimento de revisão	Dois especialistas revisaram amostra de documentos
Decisão final	Aceite, modificação ou rejeição	Rótulo gerado foi modificado para termo mais específico
Limitação	Restrição do procedimento	A IA não leu textos completos
Risco residual	Risco que permanece	Possível simplificação de subtemas minoritários

Fonte: Elaborado pelo autor (2026).

Esse registro também serve como instrumento pedagógico. Ao preencher cada campo, a equipe percebe lacunas de documentação e pode corrigir o protocolo antes da publicação. Em estudos sensíveis, especialmente aqueles com consequências avaliativas, recomenda-se anexar esse registro ao relatório final ou mantê-lo disponível para auditoria.

Capítulo 19

Aplicações da cientometria com inteligência artificial

19.1 Introdução

A cientometria com inteligência artificial constitui, simultaneamente, um campo de aplicação e uma questão metodológica. Como campo de aplicação, indicadores, mapas, redes, séries temporais e análises textuais podem apoiar a gestão da pesquisa, a formulação de políticas científicas, a organização do conhecimento, a prospecção tecnológica e o desenho de estratégias de colaboração. Como questão metodológica, a incorporação de modelos generativos altera a forma de traduzir perguntas em procedimentos, descrever dados, comunicar evidências e explicitar incertezas. O uso de IA não deve, portanto, ser compreendido como mera automação, mas como mediação analítica que precisa ser controlada, documentada e criticamente interpretada.

Este capítulo organiza as aplicações em oito domínios: diagnóstico de áreas de pesquisa, identificação de temas emergentes, apoio à política científica, avaliação de programas de pós-graduação, estratégias de colaboração, inteligência competitiva acadêmica, prospecção tecnológica e organização do conhecimento e curadoria de acervos. Em todos eles, a IA pode auxiliar na leitura de *corpora* extensos, na rotulagem de *clusters*, na comparação de indicadores e na elaboração de relatórios. Essa capacidade não elimina a necessidade de delimitar a pergunta, qualificar as fontes, registrar escolhas metodológicas e preservar a responsabilidade humana sobre conclusões e recomendações.

A literatura de mapeamento científico demonstra que cocitação, acoplamento bibliográfico, coocorrência de palavras-chave, mapas de sobreposição e técnicas de visualização revelam dimensões distintas da estrutura e da dinâmica da ciência. Esses instrumentos oferecem representações parciais da produção publicada, condicionadas por bases de dados, estratégias de busca, normalizações, janelas temporais e escolhas algorítmicas. A IA pode tornar mais explícita a passagem entre evidência, interpretação e ação, desde que não seja tratada como fonte autônoma de autoridade (Small, 1973; Kessler, 1963; White e Griffith, 1981; Callon, Courtial e Laville, 1991; Porter e Rafols, 2009; Rafols, Porter e Leydesdorff, 2010; Cobo *et al.*, 2012).

Aplicações cientométricas são sensíveis a efeitos performativos: quando indicadores se convertem em metas, podem induzir comportamentos estratégicos de publicação, colaboração e citação. A avaliação responsável recomenda combinar métricas com julgamento especializado, considerar missões institucionais distintas e evitar comparações descontextualizadas. A automação não neutraliza esses riscos e pode ampliá-los se transformar associações bibliométricas em recomendações prescritivas (MERTON, 1968; HICKS *et al.*, 2015; FIRE; GUESTRIN, 2019).

Quadro 19.1 - Aplicações, produtos e cautelas gerais da cientometria com IA

Aplicação	Produto típico	Cautela metodológica
Diagnóstico de área	Relatório diagnóstico, mapa temático e painel de indicadores	Não confundir visibilidade bibliográfica com qualidade intrínseca.

Aplicação	Produto típico	Cautela metodológica
Temas emergentes	Radar de tendências e relatório de sinais fracos	Distinguir tendência substantiva de modismo terminológico ou artefato de busca.
Política científica	Nota técnica, monitoramento e mapa de competências	Evitar metas rígidas e rankings simplificadores.
Pós-graduação	Autoavaliação, análise de linhas e internacionalização	Considerar missão, cultura disciplinar e produtos não indexados.
Colaboração	Mapa de parceiros e matriz de complementaridade	Validar relações sociais, reciprocidade e viabilidade de cooperação.
Inteligência competitiva	Radar de posicionamento e observatório temático	Evitar vigilância individual e competição baseada apenas em métricas.
Prospecção tecnológica	Roadmap, mapa ciência-tecnologia e sinais fracos	Tratar resultados como cenários, não como previsões determinísticas.
Curadoria de acervos	Taxonomia, tesauro preliminar e mapa de coleção	Validar termos e preservar diversidade linguística e conceitual.

Fonte: Elaborado pelo autor (2026).

Neste livro, *prompts* são tratados como instrumentos metodológicos. Devem explicitar tarefa, dados fornecidos, restrições, incertezas e formato de saída. Em um diagnóstico institucional, por exemplo, cada afirmação precisa ser vinculada à fonte, ao indicador, ao período e ao nível de confiança. Em uma análise de colaboração, é necessário distinguir proximidade temática, complementaridade metodológica, conexão existente e viabilidade estratégica. A documentação do *prompt* permite auditoria, revisão e reprodução aproximada do processo.

As aplicações discutidas não são independentes. O diagnóstico de áreas pode revelar lacunas que alimentam políticas científicas; a detecção de temas emergentes pode orientar prospecção; a análise de colaboração subsidia internacionalização; e a organização do conhecimento melhora recuperação e curadoria. A visão integrada conecta evidências, interpretação contextual e ação institucional responsável, sem converter instrumentos analíticos em mecanismos automáticos de decisão.

19.2 Diagnóstico de áreas de pesquisa

O diagnóstico de uma área de pesquisa busca responder quais temas estruturam o campo, quem são os principais grupos e instituições, quais canais concentram a produção, quais métodos predominam e que subáreas crescem, estabilizam ou declinam. A IA pode apoiar a tradução dessas perguntas em protocolo analítico, sugerindo unidades de análise, métricas, mapas e formas de comunicação. A validade do diagnóstico, porém, depende da qualidade da busca, da adequação da base e da contextualização histórica e sociológica do campo (ARIA; CUCCURULLO, 2017; DONTHU *et al.*, 2021; ZUPIC; ČATER, 2015).

Um diagnóstico cientométrico não deve ser reduzido a contagens de publicações ou listas de autores mais citados. Ele combina dimensões produtivas, relacionais, temáticas e temporais. A produção anual indica expansão ou retração; a distribuição por periódicos revela canais de comunicação; a coautoria mostra formas de organização social; a cocitação ajuda

a reconhecer bases intelectuais; o acoplamento bibliográfico aproxima frentes recentes; e a coocorrência de palavras-chave sugere tópicos de pesquisa. Diferentes técnicas respondem a perguntas distintas e não devem ser tratadas como equivalentes.

Em diagnósticos institucionais, a IA pode apoiar a construção de painéis interpretativos. Um programa de pós-graduação ou centro de pesquisa pode comparar linhas, colaboração internacional, impacto relativo, temas emergentes e aderência a agendas estratégicas. O modelo de linguagem pode redigir uma síntese inicial, apontar inconsistências e sugerir perguntas complementares. A definição do que constitui força, fragilidade, oportunidade ou ameaça, entretanto, depende da missão acadêmica, da política científica e do conhecimento especializado.

A principal contribuição dos *prompts* está em tornar o diagnóstico replicável. Em vez de solicitar genericamente que a IA "analise uma área", o pesquisador deve especificar escopo, período, fonte, unidade de análise, indicadores, nível de agregação e formato de saída. A resposta deve diferenciar evidências observáveis, inferências plausíveis, hipóteses a verificar e limitações. Esse desenho reduz narrativas impressionistas e aproxima a interação com IA de um protocolo científico.

O diagnóstico pode apoiar decisões editoriais, de financiamento e de capacitação. Ao identificar subáreas pouco representadas, redes desconectadas ou dependência excessiva de poucos canais, gestores podem planejar chamadas, formação e internacionalização. A cientometria, contudo, deve iluminar possibilidades, não prescrever automaticamente prioridades. A IA organiza o raciocínio, mas não substitui a deliberação institucional.

Em áreas novas ou difusas, a IA pode comparar rótulos de *clusters*, resumir documentos representativos e propor denominações provisórias. Esses rótulos precisam ser validados por especialistas, porque agrupamentos terminologicamente próximos podem reunir tradições, métodos e objetos distintos. A categoria gerada por IA deve permanecer auditável e revisável.

19.3 Identificação de temas emergentes

A identificação de temas emergentes permite observar movimentos iniciais de reorganização da literatura. Crescimento acelerado, aumento de citações recentes, surgimento de termos, entrada de novos autores e presença em conferências são sinais possíveis, mas não suficientes. Emergência não equivale automaticamente a relevância, maturidade ou impacto futuro; exige combinação de novidade, crescimento, coerência e potencial de persistência (CHEN, 2006; BORNMANN; MUTZ, 2015; ROTOLO; HICKS; MARTIN, 2015).

A análise deve combinar sinais quantitativos e qualitativos. A frequência de uma palavra-chave pode refletir mudança de indexação, nomenclatura de moda ou alteração na estratégia de busca. É necessário observar continuidade temporal, presença em diferentes bases, diversidade de atores, formação de literatura conceitual e metodológica própria e conexão com problemas científicos, sociais ou tecnológicos persistentes.

Em ambientes institucionais, a detecção de tendências apoia decisões sobre editais, contratação, disciplinas, redes e infraestrutura. A resposta cientométrica precisa mostrar crescimento, distribuição geográfica, principais atores, relação com capacidades locais e incertezas. A IA pode organizar essas dimensões em quadros comparativos e cenários, mas não deve recomendar investimento sem critérios explícitos.

Prompts bem construídos ajudam a diferenciar tendência substantiva de artefato de busca. Devem solicitar verificação da dependência de um único termo, país, periódico, evento ou base; pedir sinônimos, evidências contrárias e necessidade de validação manual; e separar sinais fortes, sinais fracos e ruído potencial.

Mapas longitudinais permitem observar aproximação, separação e mudança de rótulos entre *clusters*. Um tema inicialmente metodológico pode migrar para aplicações em saúde, educação ou sustentabilidade. Essa transição deve ser descrita como hipótese sustentada pelos documentos observados, e não como trajetória inevitável, pois mapas dependem de limiares, normalizações e escolhas de visualização.

Temas emergentes podem aparecer primeiro em *preprints*, anais, patentes, projetos financiados ou documentos técnicos. A integração de fontes amplia a sensibilidade, mas aumenta heterogeneidade e ruído. O analista deve definir o que conta como evidência de emergência e quais fontes são adequadas ao domínio.

19.4 Apoio à política científica

A política científica utiliza evidências para orientar prioridades, distribuir recursos, acompanhar programas e avaliar resultados. A cientometria com IA pode organizar informações sobre capacidades instaladas, especialização temática, redes, lacunas regionais, circulação internacional e relação entre produção científica e agendas públicas. Como esses resultados podem afetar financiamento e reputação, seu uso exige transparência, pluralidade de métricas e avaliação contextual (LEYDESDORFF; RAFOLS, 2009; RAFOLS; PORTER; LEYDESDORFF, 2010; HICKS *et al.*, 2015).

Não existe uma métrica única capaz de representar qualidade, relevância ou impacto social. Políticas de internacionalização, fortalecimento regional, ciência aberta e inovação exigem evidências diferentes. A IA pode ajudar a selecionar indicadores coerentes com cada objetivo, desde que o *prompt* explicite a finalidade, a população afetada e os critérios normativos.

A análise cientométrica pode monitorar diversificação temática, consolidação de grupos, colaboração e presença em áreas-alvo. A comparação entre períodos, contudo, não autoriza atribuir causalidade automática à política. Mudanças podem decorrer de tendências globais, alterações da base, incentivos editoriais ou outros fatores externos. A avaliação de impacto requer desenho adicional.

Prompts para política científica devem solicitar cenários, condições de implementação, lacunas de informação e riscos distributivos. Diante de concentração em poucas instituições, por exemplo, a IA pode formular perguntas sobre descentralização, cooperação regional e formação de redes. A decisão deve envolver gestores, pesquisadores e comunidades afetadas.

Mapas de competências podem indicar grupos associados a missões nacionais ou regionais, mas produção bibliográfica não equivale a capacidade de implementação. Infraestrutura, experiência tácita, extensão, ensino e transferência tecnológica precisam ser verificados por outras fontes.

A governança de dados é central: bases comerciais e abertas possuem coberturas distintas, algoritmos de afiliação podem errar e publicações locais podem ser sub-representadas. A comparação entre fontes e a documentação da busca são indispensáveis (MONGEON; PAUL-HUS, 2016; VISSER; VAN ECK; WALTMAN, 2021; PRIEM; PIWOWAR; ORR, 2022; ALPERIN *et al.*, 2024).

19.5 Avaliação de programas de pós-graduação

A avaliação de programas de pós-graduação envolve desempenho acadêmico, formação, consolidação de linhas, internacionalização e contribuição social. Indicadores podem oferecer evidências sobre produção, colaboração e circulação, mas não capturam

integralmente qualidade formativa, orientação, inovação pedagógica, inserção regional ou relevância social. A IA pode organizar dados e redigir sínteses, desde que opere em modelo avaliativo multidimensional e contextual (ABRAMO; D'ANGELO, 2011; MOED; HALEVI, 2015).

A análise deve distinguir indivíduos, linhas de pesquisa, grupos, programas e instituições. Agregações podem esconder desigualdades internas; comparações entre programas podem ser injustas quando ignoram idade, área, perfil regional, infraestrutura e tipos documentais. A IA pode decompor dimensões e apontar coerência entre projetos, produção e palavras-chave, mas não deve atribuir mérito sem validação institucional.

A coautoria docente-discente pode indicar integração entre formação e pesquisa, mas varia segundo culturas disciplinares. Em algumas áreas predominam artigos; em outras, livros, capítulos, ensaios e produtos técnicos. O resultado deve ser apresentado como evidência parcial, e não como julgamento final.

A internacionalização exige indicadores múltiplos: coautoria, mobilidade, projetos conjuntos, citações externas e participação em redes representam fenômenos diferentes. A IA pode mostrar se a inserção está concentrada ou distribuída e alertar para o risco de confundir publicação em inglês com internacionalização substantiva.

Prompts para avaliação devem equilibrar evidências quantitativas e narrativa qualitativa. Uma estrutura útil separa dados observáveis, interpretação contextual e perguntas para validação. O modelo deve assinalar lacunas, erros de afiliação e indicadores inadequados.

A ética é central porque dados avaliativos afetam carreiras, recursos e reputações. Devem ser preservadas privacidade, contestabilidade e revisão humana. Indicadores de citação e métricas alternativas precisam ser usados segundo sua finalidade e seus limites, sem substituição automática do julgamento especializado (WALTMAN, 2016; BORNMANN; HAUNSCHILD, 2018).

19.6 Estratégias de colaboração

Redes de coautoria, acoplamento bibliográfico e coocorrência temática podem revelar atores próximos, atores-ponte, comunidades pouco conectadas e oportunidades de cooperação. A IA pode apoiar a leitura dessas redes segundo proximidade temática, complementaridade metodológica, localização e histórico de relação. Colaboração, entretanto, é fenômeno social: afinidade bibliométrica não garante confiança, interesse, recursos ou viabilidade política (VAN ECK; WALTMAN, 2010; ZUPIC; ČATER, 2015).

Um mapa deve diferenciar colaboração existente, potencial e desejável. A existente aparece em coautorias e projetos; a potencial pode ser inferida por temas ou métodos complementares; a desejável depende de objetivos estratégicos. A IA pode organizar essas categorias, mas precisa justificar sugestões e indicar evidências.

Centralidade e intermediação oferecem subsídios, mas não devem ser interpretadas de forma simplista. Atores centrais podem ser influentes ou sobrecarregados; atores periféricos podem representar diversidade regional e nichos especializados; e dependência de poucos mediadores pode revelar fragilidade da rede.

Redes por país, instituição e tema podem revelar dependência de poucos parceiros, ausência em regiões estratégicas e oportunidades de cooperação Sul-Sul. Cenários devem combinar fortalecimento de laços, diversificação, complementaridade e reciprocidade, e não apenas maximização de citações.

Prompts para oportunidades de colaboração devem exigir parceiro potencial, evidência bibliométrica, complementaridade, modalidade de cooperação, riscos e informações a verificar. Expressões como "parceiro ideal" ou "melhor instituição" devem ser evitadas sem critérios explícitos.

A análise também precisa examinar assimetrias de liderança, autoria correspondente, controle de dados e distribuição de benefícios. Estratégias cientometricamente informadas devem promover qualidade científica e equidade relacional.

19.7 Inteligência competitiva acadêmica

A inteligência competitiva acadêmica utiliza informações sobre produção, temas, atores e redes para compreender posicionamento e oportunidades no ambiente de pesquisa. Diferentemente de uma lógica empresarial estrita, deve respeitar cooperação, abertura, integridade e missão pública. A ciência da ciência oferece instrumentos para monitorar portfólios e dinâmicas sem reduzir o sistema científico a uma competição de curto prazo (FORTUNATO *et al.*, 2018).

Uma instituição pode investigar em quais temas possui especialização relativa, onde perde presença, que áreas crescem em organizações comparáveis e quais lacunas podem ser exploradas. Vantagem bibliométrica, contudo, indica presença mensurável em determinadas fontes e períodos; não significa superioridade científica.

A inteligência acadêmica também pode apoiar decisões editoriais e de comunicação. A IA pode organizar periódicos, conferências e redes relevantes, mas não deve inventar índices, rankings ou fatores de impacto. Toda recomendação precisa ser baseada nos dados fornecidos ou em fonte verificável.

Em áreas de rápida transformação, o monitoramento pode identificar grupos que ingressam em determinado tema, métodos adotados, financiadores e reorganização de comunidades. Esse acompanhamento deve evitar vigilância hostil: concorrentes também podem ser colaboradores.

Prompts devem separar posicionamento, diferenciação, risco de imitação, lacunas de dados e possíveis alianças. Em vez de prescrever "invista neste tema", a saída deve apresentar crescimento, compatibilidade com capacidades internas, intensidade competitiva, maturidade e alternativas estratégicas.

Comparações individuais são especialmente frágeis e potencialmente invasivas. Sempre que possível, a análise deve priorizar agregados institucionais, grupos e temas, com regras claras de privacidade, finalidade e contestação.

19.8 Prospecção tecnológica

A prospecção tecnológica conecta cientometria, estudos de inovação e análise de patentes. Seu objetivo não é prever o futuro com certeza, mas identificar sinais, trajetórias, convergências e incertezas. Publicações podem indicar bases de conhecimento; patentes, apropriação tecnológica; projetos financiados, apostas estratégicas; e conferências, temas iniciais. A IA pode integrar essas fontes em narrativas prospectivas, desde que registre o grau de evidência (PORTER; DETAMPEL, 1995; KAYSER; BLIND, 2017).

A relação entre ciência e tecnologia é complexa. Nem toda literatura gera inovação aplicada, e nem toda inovação depende de artigos altamente citados. A prospecção deve diferenciar pesquisa básica, prova de conceito, desenvolvimento experimental, aplicação comercial e adoção social.

Modelos de linguagem podem analisar títulos, resumos e descrições de patentes para sugerir categorias, aplicações e relações. Sinais fracos, por definição, são ambíguos: podem representar ruído, moda, promessa de mercado ou início de trajetória relevante. O *prompt* deve classificar evidências por força, fonte, período e necessidade de validação.

Roadmaps tecnológicos podem combinar literatura, patentes, documentos regulatórios e dados de financiamento em horizontes de curto, médio e longo prazo. Esses horizontes devem ser apresentados como cenários, não como previsões, porque infraestrutura, aceitação social, custos e políticas públicas condicionam a trajetória.

Prompts prospectivos devem substituir perguntas determinísticas por questões sobre trajetórias plausíveis, premissas e sinais de confirmação ou refutação. Essa formulação transforma a IA em ferramenta de estruturação de cenários, e não em oráculo tecnológico.

Patentes variam por setor e país; empresas podem publicar pouco; tecnologias sociais podem não aparecer em bases tradicionais. A validade prospectiva depende de triangulação entre dados, teoria, experiência especializada e contexto (ROTOLO; HICKS; MARTIN, 2015).

19.9 Organização do conhecimento e curadoria de acervos

A organização do conhecimento e a curadoria de acervos são aplicações fundamentais da cientometria com IA. Bibliotecas, repositórios e grupos de pesquisa precisam classificar, descrever, recuperar e atualizar coleções. A cientometria permite mapear temas, lacunas, vocabulários e comunidades documentais; a IA pode apoiar descritores, resumos, taxonomias e revisão de metadados, sempre sob critérios documentários e validação especializada.

Mapas de coocorrência revelam associações conceituais, mas similaridade textual não equivale a equivalência semântica. Termos próximos podem pertencer a tradições distintas; termos diferentes podem designar o mesmo fenômeno. A IA deve propor hipóteses de organização, não impor classificações definitivas (CALLON; COURTIAL; LAVILLE, 1991).

Em repositórios institucionais, a IA pode identificar metadados incompletos, afiliações inconsistentes, palavras-chave genéricas e lacunas de cobertura. Qualquer alteração precisa ser rastreável e, quando necessário, validada por bibliotecários, autores ou curadores.

A análise temporal permite comparar vocabulários, identificar termos em declínio e sugerir equivalências. A revisão deve incorporar vocabulário contemporâneo sem apagar termos históricos necessários à memória e à recuperação.

Prompts para curadoria devem apresentar agrupamentos com justificativa, grau de confiança e ambiguidades. A saída deve separar termo preferencial, variantes, sinônimos, termos relacionados e casos para revisão. Análises de proximidade citacional também podem apoiar a identificação de documentos relacionados, desde que sejam tratadas como evidência complementar (GIPP; BEEL, 2009).

Sistemas de classificação podem reproduzir hierarquias linguísticas e invisibilizar produção regional. A resposta responsável combina tecnologia, *expertise* biblioteconômica, participação das comunidades e documentação das escolhas.

19.10 Matriz operacional das aplicações

Os blocos operacionais das oito aplicações compartilham a mesma lógica metodológica: formular a decisão a apoiar, selecionar unidades e indicadores compatíveis, definir um produto e controlar riscos específicos. Para evitar repetição e facilitar a comparação, o Quadro 19.2 consolida esses elementos.

Quadro 19.2 - Síntese operacional das aplicações cientométricas com IA

Aplicação	Unidades e indicadores	Produtos	Risco central
Diagnóstico de área	Artigos, autores, instituições, países, termos e referências; produção, citações normalizadas e redes.	Relatório diagnóstico, mapa temático, matriz SWOT e painel.	Cobertura parcial, viés linguístico e confusão entre visibilidade e qualidade.
Temas emergentes	Termos, títulos, resumos, citações recentes, conferências e patentes; crescimento, *burst* e novidade.	Radar de tendências, mapa temporal e hipóteses prospectivas.	Modismo terminológico, ruído de base e extrapolação de séries curtas.
Política científica	Instituições, programas, áreas estratégicas, chamadas, colaborações e patentes; especialização e alinhamento.	Nota técnica, painel de monitoramento e mapa de competências.	Subordinação a rankings, metas rígidas e marginalização de missões locais.
Pós-graduação	Docentes, discentes, egressos, linhas, projetos e produtos; formação, colaboração e impacto multidimensional.	Relatório de autoavaliação, matriz de evidências e diagnóstico de linhas.	Redução da qualidade formativa a publicação e comparação entre campos incomparáveis.
Colaboração	Autores, grupos, instituições, países, temas e métodos; centralidade, intermediação e complementaridade.	Mapa de parceiros, matriz de complementaridade e plano de aproximação.	Inferência de afinidade sem validação social e reprodução de assimetrias.
Inteligência competitiva	Instituições, grupos, temas, conferências e portfólios; especialização, crescimento e redes.	Observatório temático, radar de posicionamento e briefing estratégico.	Ranqueamento agressivo, vigilância individual e imitação de agendas.
Prospecção tecnológica	Publicações, patentes, projetos, classes tecnológicas, atores e aplicações; convergência e sinais fracos.	Roadmap, radar de oportunidades e mapa ciência-tecnologia.	Confusão entre promessa e maturidade; determinismo tecnológico.
Curadoria de acervos	Documentos, descritores, conceitos, citações e classificações; redes semânticas, cobertura e atualização.	Tesauro preliminar, taxonomia, mapa de acervo e guia de recuperação.	Automação sem validação, apagamento terminológico e viés de indexação.

Fonte: Elaborado pelo autor (2026).

19.11 Integração entre aplicações: da evidência à decisão responsável

As oito aplicações compartilham uma estrutura: pergunta situada, seleção de fontes, limpeza, escolha de indicadores, interpretação e comunicação. A IA pode intervir em todas as etapas, mas sua utilidade aumenta quando os *prompts* indicam objetivo, público, dados disponíveis, limites de inferência, formato e obrigação de explicitar incertezas.

Produtos distintos exigem linguagens distintas. Um mapa destinado a pesquisadores pode enfatizar método; um briefing para gestores deve evidenciar implicações e incertezas; uma nota de política precisa apresentar opções e consequências distributivas; e uma taxonomia requer relações conceituais. A adaptação comunicacional deve preservar fidelidade aos dados.

A integração entre IA e cientometria amplia a capacidade analítica e também o risco de automatização de decisões. Diagnósticos, rankings e recomendações podem adquirir

autoridade aparente. Por isso, cada aplicação deve incluir limitações, matriz de evidências, registro dos *prompts* e indicação das conclusões validadas por especialistas.

O princípio da proporcionalidade é transversal. Nem toda decisão exige análise complexa; nem toda base justifica modelos sofisticados; e mais indicadores não compensam dados inadequados. Em contextos políticos, a cientometria deve integrar um conjunto mais amplo de evidências.

A aplicação responsável combina competência técnica, interpretação substantiva e ética. *Prompts* eficazes não substituem essas competências; tornam suas relações mais explícitas, revisáveis e documentáveis.

19.12 Prompts do capítulo

Os *prompts* a seguir transformam aplicações cientométricas em procedimentos auditáveis. Devem ser adaptados ao domínio, à base, ao público e à decisão. Recomenda-se anexar tabelas, mapas ou descrições dos dados e proibir o preenchimento de lacunas com informações não verificadas.

Prompt 19.12.1 Diagnóstico institucional

Atue como especialista em cientometria aplicada e inteligência artificial.

Com base nos dados abaixo sobre [INSTITUIÇÃO/ÁREA/PROGRAMA/TEMA], produza um diagnóstico em cinco partes:

(1) descrição objetiva;

(2) pontos fortes apoiados por evidências;

(3) fragilidades ou lacunas apoiadas por evidências;

(4) oportunidades estratégicas;

(5) limitações metodológicas.

Para cada afirmação, indique fonte, indicador, período e nível de confiança.

Não transforme indicadores em julgamento de qualidade sem contextualização.

Separe dados observados, inferências e recomendações.

Prompt 19.12.2 Mapeamento de oportunidades de colaboração

Atue como analista de redes científicas.

A partir da lista de autores, instituições, temas, coautorias e indicadores abaixo, identifique oportunidades de colaboração.

Para cada oportunidade, apresente: parceiro potencial, evidência bibliométrica, proximidade temática, complementaridade metodológica, histórico de relação, modalidade possível, riscos e informações que precisam ser verificadas manualmente.

Não afirme que um parceiro é ideal sem critérios explícitos e não use prestígio genérico como justificativa.

Prompt 19.12.3 Detecção de lacunas estratégicas

Atue como especialista em estratégia científica.

Analise os resultados abaixo e classifique lacunas em: temática, metodológica, de colaboração, geográfica, de internacionalização, de acervo ou de impacto social.

Para cada lacuna, indique evidência, explicação possível, risco de interpretação indevida, ação exploratória e dado adicional necessário.

Não recomende ações irreversíveis com base em evidência fraca.

Prompt 19.12.4 Briefing para tomada de decisão

Atue como assessor técnico em ciência, tecnologia e inovação.

Transforme os achados abaixo em um briefing executivo de até duas páginas para gestores não especialistas.

Inclua: problema, evidências principais, interpretações prudentes, opções de ação, riscos, incertezas, implicações éticas e perguntas para deliberação.

Explique métricas técnicas quando necessário e evite afirmações causais quando houver apenas associação.

Prompt 19.12.5 Avaliação de tema emergente

Atue como especialista em tendências científicas.

Avalie se [TEMA] pode ser considerado emergente com base nos dados abaixo.

Considere crescimento temporal, diversidade de autores, entrada de periódicos, *burst* de termos, citações recentes, conferências, patentes e consistência entre bases.

Separe sinais fortes, sinais fracos, artefatos de busca e evidências ausentes.

Formule hipóteses prospectivas com cautela e indique como testá-las.

Prompt 19.12.6 Organização do conhecimento e curadoria

Atue como especialista em organização do conhecimento.

Receba termos, títulos, resumos e descritores e proponha uma taxonomia preliminar.

Para cada categoria, informe: nome, termos incluídos, termos relacionados, justificativa conceitual, documentos representativos, grau de confiança e casos ambíguos para revisão humana.

Preserve termos históricos relevantes e indique vieses linguísticos ou disciplinares.

19.13 Checklist de qualidade para aplicações com IA

O checklist sintetiza pontos de controle para o planejamento, a revisão dos resultados e a aprovação de relatórios, pareceres e briefings.

Quadro 19.3 - Checklist de qualidade para aplicações cientométricas assistidas por IA

Dimensão	Pergunta de controle
Pergunta aplicada	A decisão a apoiar e o público destinatário foram definidos?
Fonte de dados	Bases, período, filtros e critérios de inclusão foram documentados?
Unidade de análise	Está claro se a análise se refere a artigos, autores, instituições, países, temas ou conceitos?
Indicadores	Os indicadores são adequados ao objetivo e foram interpretados em contexto?
IA e prompts	Os *prompts* foram preservados, versionados e associados às saídas?
Validação	Houve revisão humana por especialistas, gestores, bibliotecários ou curadores?
Limitações	Foram explicitados vieses de base, língua, disciplina, cobertura e período?
Uso decisório	As recomendações são proporcionais à força das evidências e aos riscos?

Fonte: Elaborado pelo autor (2026).

19.14 Estudo de caso ilustrativo: saúde digital em uma universidade pública

Considere uma universidade pública que deseja compreender sua posição em pesquisa sobre saúde digital. A equipe reúne registros de Web of Science, Scopus e OpenAlex, complementa-os com o repositório institucional e inclui patentes quando disponíveis. A estratégia de busca, a deduplicação, a normalização de afiliações e os períodos são documentados. A IA sugere termos, classifica resumos e produz uma matriz inicial de atores; todos os agrupamentos são revisados por especialistas.

O diagnóstico revela crescimento de publicações, concentração em poucos grupos, colaboração internacional moderada e presença de telemedicina, prontuário eletrônico, aprendizagem de máquina e vigilância epidemiológica. A síntese distingue dados observáveis, interpretações e hipóteses. A baixa integração entre computação, saúde coletiva e hospitais universitários resulta da triangulação entre coautoria, palavras-chave, projetos e conhecimento institucional, e não de uma única métrica.

A universidade solicita um briefing para deliberar sobre a criação de um programa estratégico. A IA organiza opções: fortalecer a rede interna, buscar cooperação, criar disciplina transversal, submeter projetos e melhorar a curadoria de dados clínicos. Cada opção inclui evidências, incertezas e riscos. A decisão permanece com o comitê, que considera prioridades sociais, orçamento, infraestrutura ética e formação discente.

O caso também mostra o risco de extrapolação. Apenas citações internacionais poderiam sugerir baixa competitividade; apenas produção nacional poderia ocultar lacunas de internacionalização; e rotulagem automática poderia apagar diferenças clínicas,

computacionais e organizacionais. O valor da abordagem está em integrar fontes, explicitar incertezas e qualificar a deliberação.

19.15 Riscos transversais e estratégias de mitigação

O primeiro risco é a naturalização dos dados. Bases bibliográficas são construções sociotécnicas que diferem em cobertura, indexação, metadados, idioma e tipos documentais. A IA facilita a exploração, mas não corrige automaticamente vieses. Comparação entre fontes, amostragem manual, documentação da busca e preservação dos registros brutos reduzem esse problema (Mongeon e Paul-Hus, 2016; Visser, Van Eck e Waltman, 2021; Alperin *et al.*, 2024).

O segundo risco é a autoridade retórica. Textos fluentes podem conferir aparência de certeza a inferências frágeis. *Prompts* devem exigir evidências, contraevidências, níveis de confiança, hipóteses alternativas e limitações; a revisão humana precisa buscar excesso de assertividade, não apenas erros factuais.

O terceiro risco é o deslocamento da responsabilidade. A responsabilidade permanece com autores, gestores e instituições. Documentos devem declarar ferramenta, finalidade, *prompts*, etapas revisadas e limites das saídas.

O quarto risco é transformar instrumentos analíticos em mecanismos de controle. *Dashboards* e rankings podem orientar aprendizagem, mas também vigilância e incentivos perversos. A mitigação requer pluralidade de indicadores, alinhamento à missão, participação dos avaliados e revisão dos efeitos.

O quinto risco é reduzir a diversidade científica. Aplicações estratégicas podem favorecer temas visíveis, língua inglesa, redes centrais e retornos mensuráveis. *Prompts* críticos devem procurar ausências, línguas, formatos não indexados, temas locais e dimensões sociais não capturadas por citações.

19.16 Considerações finais

As aplicações da cientometria com IA podem apoiar decisões que vão da curadoria de acervos à política científica. Sua potência depende de rigor metodológico e responsabilidade interpretativa. A IA amplia leitura, comparação, síntese e comunicação, mas não confere neutralidade aos indicadores nem elimina vieses, diferenças disciplinares e disputas de valor.

Prompts são peças centrais dessa governança porque definem papel, tarefa, dados, restrições, formato, cautelas e validação. Seu valor não está apenas na obtenção de respostas, mas na estruturação documentável do raciocínio analítico.

A cientometria com IA é prática sociotécnica: combina bases, algoritmos, modelos de linguagem, *expertise* humana, valores institucionais e contextos decisórios. A aplicação responsável equilibra análise quantitativa, compreensão qualitativa e deliberação coletiva.

Em síntese, a IA pode tornar a cientometria mais acessível e comunicável, mas também mais arriscada quando usada de modo opaco. O sucesso dependerá menos da fluência textual e mais da capacidade de produzir análises rastreáveis, contestáveis, contextualizadas e socialmente responsáveis.

Apêndice A.19 - Protocolo operacional para aplicações da cientometria com IA

O protocolo a seguir substitui a repetição das mesmas etapas para cada domínio. A sequência é comum às oito aplicações; os elementos específicos de *corpus*, indicadores, produtos e riscos são apresentados no Quadro A.1.

Etapa 1 – Formular a demanda decisória: Registrar quem solicita a análise, qual decisão será apoiada, quem será afetado, quais consequências podem decorrer do relatório e quais critérios normativos orientarão a interpretação. A IA pode propor perguntas, mas a formulação final deve ser validada por responsáveis humanos.

Etapa 2 – Construir o *corpus:* Selecionar fontes compatíveis com o objetivo: bases bibliográficas, repositórios, patentes, documentos institucionais, registros de projetos ou outras evidências. Documentar data de coleta, estratégia de busca, filtros, versões, critérios de inclusão e exclusão e complementações manuais.

Etapa 3 – Limpar e normalizar os dados: Padronizar autores, instituições, países, periódicos e termos. A IA pode sugerir agrupamentos e marcar ambiguidades, mas a decisão final precisa ser documentada. Erros de desambiguação alteram redes, rankings e interpretações.

Etapa 4 – Selecionar indicadores e visualizações: Escolher indicadores em função da pergunta, e não da disponibilidade da ferramenta. Justificar combinações, parâmetros, limiares e normalizações. Visualizações devem tornar relações compreensíveis sem esconder exclusões.

Etapa 5 – Interpretar com prudência: Separar descrição, inferência, hipótese e recomendação. Para cada afirmação estratégica, perguntar: qual dado a sustenta? Há explicações alternativas? A conclusão depende de uma única fonte? O resultado pode ser artefato de busca?

Etapa 6 – Comunicar, validar e revisar: Apresentar metodologia, achados, limites, recomendações proporcionais e perguntas pendentes. Adaptar a linguagem ao público sem remover incertezas. Submeter o relatório a usuários e especialistas, permitindo correção de dados e ajuste das interpretações.

Quadro 19.4 - Especificidades do protocolo por aplicação

Aplicação	*Corpus* e unidades	Indicadores prioritários	Produto e controle principal
Diagnóstico de área	Bases bibliográficas, repositórios e produção institucional	Produção, impacto normalizado, cocitação, acoplamento, coocorrência e distribuição geográfica	Relatório, mapa e painel; controlar cobertura e viés de visibilidade.
Temas emergentes	Artigos, *preprints*, conferências, patentes e projetos	Crescimento, *burst*, novidade, citações recentes e entrada de atores	Radar e hipóteses; controlar modismo e ruído.
Política científica	Programas, instituições, chamadas, publicações, patentes e resultados sociais	Especialização, colaboração, interdisciplinaridade, abertura e alinhamento	Nota técnica e monitoramento; controlar metas rígidas e desigualdades.

Aplicação	*Corpus* e unidades	Indicadores prioritários	Produto e controle principal
Pós-graduação	Docentes, discentes, egressos, linhas, projetos e produtos	Formação, produção por linha, coautoria, colaboração e impacto multidimensional	Autoavaliação; controlar comparações injustas e lacunas de produtos.
Colaboração	Autores, grupos, instituições, países, projetos, temas e métodos	Centralidade, intermediação, densidade, complementaridade e histórico	Mapa de parceiros; validar relações e assimetrias.
Inteligência competitiva	Instituições, grupos, temas, conferências, chamadas e portfólios	Especialização, crescimento, impacto, redes e produção tecnológica	Radar e briefing; evitar vigilância e competição simplificadora.
Prospecção tecnológica	Publicações, patentes, projetos, atores, classes e regulações	Crescimento, convergência, citações de patentes e colaboração universidade-empresa	Roadmap e cenários; evitar determinismo e extrapolação.
Curadoria de acervos	Documentos, descritores, classificações, autores, citações e conceitos	Frequência, redes semânticas, cobertura, uso e atualização terminológica	Taxonomia e tesauro; controlar automatização e apagamento de termos.

Fonte: Elaborado pelo autor (2026).

Em todas as aplicações, devem ser preservados os *prompts*, as versões das bases, os parâmetros, as tabelas intermediárias, as saídas aceitas e rejeitadas e as alterações realizadas por revisão humana. Esse rastro de auditoria torna a incerteza discutível e permite contestação fundamentada.

Capítulo 20

Protocolos completos de estudo

20.1 Introdução: protocolos como infraestrutura metodológica da cientometria com IA

Um protocolo completo de estudo é mais do que uma lista de procedimentos. Em cientometria, ele funciona como infraestrutura metodológica para tornar explícitas as decisões que ligam pergunta, *corpus*, indicadores, modelos, visualizações, interpretação e comunicação. Quando a inteligência artificial é incorporada ao processo, essa função se torna ainda mais importante, porque parte das decisões passa a ser mediada por respostas probabilísticas, sugestões terminológicas, classificações automáticas, sumarizações e interpretações geradas por modelos de linguagem. O protocolo, nesse contexto, impede que o uso de IA permaneça invisível, improvisado ou retoricamente naturalizado. Ele transforma uma sequência de interações com sistemas computacionais em um procedimento passível de revisão, repetição, crítica e aperfeiçoamento.

A cientometria tradicional já depende de escolhas metodológicas sensíveis: que base de dados usar, qual período analisar, como construir a estratégia de busca, como tratar duplicatas, como normalizar nomes de autores e instituições, quais indicadores calcular, que limiares aplicar a redes, como interpretar *clusters* e como comunicar resultados a públicos diferentes. A inteligência artificial não elimina essas escolhas. Ao contrário, ela amplia o número de decisões possíveis e, por isso, aumenta a necessidade de documentação. Um *prompt* que expande uma lista de palavras-chave pode alterar drasticamente a composição do *corpus*. Um *prompt* que sugere rótulos para *clusters* pode influenciar a narrativa sobre a estrutura intelectual de uma área. Um *prompt* que redige um sumário executivo pode suavizar limitações, exagerar conclusões ou aproximar indevidamente evidências exploratórias de recomendações normativas.

O capítulo assume que protocolos completos devem responder a três perguntas fundamentais. A primeira é substantiva: que fenômeno científico se pretende compreender? A segunda é técnica: que dados, métodos, ferramentas e indicadores podem produzir evidência adequada? A terceira é crítica: que incertezas, vieses, lacunas e responsabilidades acompanham o estudo? A resposta a essas perguntas exige um desenho metodológico que não se limite a executar softwares bibliométricos, mas que integre curadoria de dados, validação humana, interpretação contextual e transparência no uso de IA.

No sumário geral da obra, este capítulo ocupa uma posição de síntese aplicada. Os capítulos anteriores trataram de fundamentos, indicadores, redes, mapas temáticos, análise temporal, visualização, interpretação, escrita e questões éticas. Agora, esses elementos são reorganizados em protocolos operacionais. A ideia não é oferecer receitas inflexíveis, mas modelos adaptáveis. Cada protocolo pode ser usado como ponto de partida para estudos empíricos, relatórios institucionais, diagnósticos estratégicos, artigos científicos, projetos de pesquisa, avaliações de programas, análises prospectivas ou pareceres para gestores e formuladores de políticas.

A padronização dos protocolos não significa padronização mecânica da ciência. Ao contrário, protocolos bem formulados tornam visíveis as particularidades de cada estudo. Uma análise de temas emergentes exige estratégias distintas de uma análise institucional. Um estudo de colaboração internacional demanda tratamento cuidadoso de afiliações, países, coautorias

e redes geográficas. Um relatório de política científica precisa articular evidências quantitativas com missão pública, contexto socioeconômico, prioridades nacionais e limitações do dado bibliográfico. O protocolo, portanto, é uma forma de equilibrar comparabilidade e contextualização.

O uso de *prompts* nesse cenário deve ser compreendido como prática de mediação analítica. *Prompts* podem ajudar a decompor perguntas amplas, revisar estratégias de busca, propor indicadores, estruturar tabelas, redigir metodologias, formular hipóteses e revisar limitações. Contudo, *prompts* não são evidências. Eles são instruções textuais que orientam uma resposta do modelo. A evidência permanece nos dados, nos procedimentos, nas métricas e na interpretação validada. Um protocolo robusto deve registrar não apenas os resultados obtidos, mas também como a IA foi convocada, com que dados, sob quais restrições e com que forma de validação posterior.

Os protocolos apresentados neste capítulo seguem uma arquitetura comum. Para cada tipo de estudo são definidos objetivo, pergunta, fonte de dados, critérios de busca, indicadores, *prompts*, ferramentas, validação, produto final e limitações. Essa estrutura permite comparar desenhos de pesquisa e facilita a formação de equipes interdisciplinares. Bibliotecários, especialistas em cientometria, gestores, pesquisadores de domínio, especialistas em dados e analistas de IA podem trabalhar sobre o mesmo documento, sabendo quais decisões já foram tomadas e quais ainda requerem revisão.

A literatura de mapeamento científico, análise bibliométrica e métricas responsáveis mostra que estudos desse tipo são poderosos, mas sensíveis a escolhas de base de dados, cobertura linguística, normalização de campos, diferenças disciplinares e usos avaliativos indevidos. Ferramentas como bibliometrix, VOSviewer, CiteSpace e SciMAT ampliaram a capacidade de análise e visualização; bases como Scopus, Web of Science, Dimensions, Crossref, OpenAlex, Lens e OpenCitations diversificaram as fontes de dados; e diretrizes como o Leiden Manifesto reforçaram a importância de usar indicadores como apoio ao julgamento qualitativo, e não como substitutos de avaliação substantiva. A IA generativa se insere nesse ecossistema como uma camada adicional, capaz de acelerar sínteses e apoiar decisões, mas também de introduzir opacidade, viés e falsa autoridade se não for controlada (COBO *et al.*, 2011; DONTHU *et al.*, 2021; HICKS *et al.*, 2015; MONGEON; PAUL-HUS, 2016).

Um protocolo completo deve, portanto, cumprir quatro funções. A primeira é prospectiva: orientar o estudo antes da coleta e evitar decisões *ad hoc*. A segunda é operacional: indicar como executar cada etapa. A terceira é documental: registrar dados, *prompts*, versões, filtros e parâmetros. A quarta é crítica: explicitar limitações e riscos de interpretação. A integração dessas funções é o que permite que a cientometria com IA seja tratada como prática científica e não apenas como uso instrumental de ferramentas.

20.2 Arquitetura comum dos protocolos

Todos os protocolos apresentados neste capítulo compartilham uma estrutura mínima. Essa arquitetura não deve ser interpretada como formulário burocrático, mas como um dispositivo de qualidade metodológica. O objetivo define a finalidade substantiva do estudo. A pergunta delimita o fenômeno observável. As fontes de dados indicam a base empírica. Os critérios de busca determinam o perímetro do *corpus*. Os indicadores especificam as medidas utilizadas. Os *prompts* documentam a mediação por IA. As ferramentas registram os ambientes de processamento. A validação explicita como os resultados serão examinados. O produto final indica a forma de comunicação. As limitações previnem extrapolações indevidas.

Quadro 20.1 - Arquitetura comum dos protocolos

Elemento do protocolo	Função metodológica	Risco quando ausente
Objetivo	Define a finalidade substantiva e o uso esperado do estudo.	O estudo acumula indicadores sem direção interpretativa.
Pergunta	Transforma interesse amplo em problema analisável.	Os resultados respondem a uma curiosidade genérica, mas não a uma pergunta científica.
Fonte de dados	Estabelece a base empírica da análise.	A cobertura, os vieses e as lacunas da fonte ficam invisíveis.
Critérios de busca	Delimitam *corpus*, período, idioma, tipo documental e escopo temático.	A amostra torna-se irreprodutível e vulnerável a ruído ou exclusão arbitrária.
Indicadores	Operacionalizam produção, impacto, colaboração, interdisciplinaridade ou emergência.	Métricas são escolhidas por conveniência e não por adequação à pergunta.
Prompts	Registram como a IA foi usada para apoiar decisões e redação.	A mediação algorítmica desaparece do método e não pode ser auditada.
Ferramentas	Identificam softwares, bibliotecas, versões e parâmetros.	Resultados podem variar sem explicação entre ambientes analíticos.
Validação	Define revisão humana, amostragem, triangulação e testes de sensibilidade.	Erros de IA, metadados e interpretação passam despercebidos.
Produto final	Define relatório, artigo, *dashboard*, parecer, mapa, protocolo ou briefing.	A comunicação não se ajusta ao público nem ao uso decisório.
Limitações	Explicita o que o estudo não permite afirmar.	Associações são convertidas em causalidade ou avaliação normativa.

Fonte: Elaborado pelo autor (2026).

A homogeneidade da arquitetura facilita a reprodutibilidade e a comparação, mas cada protocolo requer adaptações. Estudos de autores demandam cuidado com homônimos, carreiras longas, mudanças de afiliação e autocitações. Estudos institucionais exigem desambiguação de nomes, identificação de unidades internas e diferenciação entre produção própria e produção associada. Estudos de temas emergentes dependem de janelas temporais recentes, vocabulário instável e mecanismos de detecção de ruptura. Relatórios de política científica requerem tradução de achados técnicos em opções de ação, com cautela normativa.

A seguir, cada protocolo é desenvolvido como módulo aplicável. A redação privilegia uma abordagem formal e científica, mas os *prompts* são apresentados em linguagem operacional para que possam ser reutilizados, adaptados e documentados em projetos reais.

20.3 Protocolo para análise de uma área científica

Objetivo do protocolo

O objetivo de uma análise de área científica é caracterizar a formação, a estrutura, a dinâmica e a orientação recente de um domínio de pesquisa. Esse domínio pode corresponder a uma disciplina consolidada, a um campo interdisciplinar, a uma tecnologia emergente, a uma linha de investigação institucional ou a um problema social que mobiliza múltiplas áreas.

O protocolo deve permitir responder como a área cresceu, quais autores e instituições são relevantes, quais periódicos ou conferências concentram a produção, quais temas organizam o campo, que redes de colaboração sustentam a produção e quais tendências recentes sugerem deslocamentos de agenda. A análise de área não deve ser confundida com inventário bibliográfico exaustivo. Trata-se de uma investigação sobre padrões de produção, circulação e estrutura cognitiva (DONTHU et al., 2021; CHEN; SONG, 2019).

Pergunta cientométrica

A pergunta central pode assumir a forma: como se configura a produção científica sobre determinado tema em um período definido, considerando crescimento, impacto, colaboração, estrutura temática e evolução temporal? Perguntas secundárias podem examinar quais subtemas dominam a literatura, quais frentes de pesquisa emergem, que países e instituições lideram a produção, quais documentos ou referências funcionam como núcleos intelectuais e como o vocabulário da área se transforma ao longo do tempo. A pergunta precisa declarar a unidade de análise predominante: documentos, autores, instituições, países, termos, referências ou *clusters* temáticos.

Fontes de dados

As fontes de dados devem ser escolhidas conforme cobertura, granularidade, acesso e finalidade. Scopus e Web of Science oferecem ampla padronização e recursos de citação, mas são bases comerciais com critérios próprios de indexação. Dimensions, Crossref, OpenAlex e OpenCitations ampliam alternativas abertas ou parcialmente abertas, embora demandem avaliação cuidadosa de completude, deduplicação e qualidade de metadados. O Google Scholar pode ampliar a cobertura, sobretudo em áreas com livros, capítulos e literatura regional, mas apresenta limitações de exportação, transparência e replicabilidade. Em áreas de saúde, PubMed pode ser essencial; em computação, bases de anais de conferências e repositórios específicos podem ser decisivos. O protocolo deve justificar a escolha da fonte, e não apenas mencioná-la (MONGEON; PAUL-HUS, 2016; VISSER; VAN ECK; WALTMAN, 2021; MARTÍN-MARTÍN *et al.*, 2021).

Critérios de busca

A estratégia de busca deve combinar termos centrais, sinônimos, acrônimos, grafias alternativas, descritores controlados e termos emergentes. Para uma área interdisciplinar, convém construir blocos conceituais: objeto, método, domínio, aplicação e população. Em bases diferentes, a mesma estratégia deve ser adaptada à sintaxe local. O protocolo deve registrar a data da busca, campos pesquisados, operadores booleanos, filtros por ano, idioma, tipo documental e critérios de exclusão. A IA pode apoiar a expansão terminológica, mas a lista final precisa ser revisada por especialistas e testada por amostras de resultados recuperados (LIBERATI *et al.*, 2009; RETHLEFSEN *et al.*, 2021; MACFARLANE; RUSSELL-ROSE; SHOKRANEH, 2022).

Indicadores

Os indicadores básicos incluem número anual de publicações, taxa de crescimento, distribuição por tipo documental, países, instituições, periódicos, autores, citações totais, média de citações por documento, documentos mais citados e redes de coautoria. Para estrutura intelectual, podem ser usadas cocitação, acoplamento bibliográfico, coocorrência

de palavras-chave e análise temática. Para evolução, podem ser construídas séries temporais, mapas temáticos por período, *burst detection*, análise de termos emergentes e comparação de *clusters*. Nenhum indicador deve ser interpretado isoladamente; crescimento pode refletir moda terminológica, ampliação de indexação ou mudança de estratégia de busca (GARFIELD, 1955; KESSLER, 1963; SMALL, 1973; BOYACK; KLAVANS, 2010).

Ferramentas

Ferramentas adequadas incluem bibliometrix/*biblioshiny* para fluxo integrado de análise, VOSviewer para mapas de rede e visualizações bibliométricas, CiteSpace para análise temporal e detecção de *bursts*, SciMAT para evolução temática, Gephi para redes customizadas, R ou Python para tratamento de dados, OpenRefine para limpeza e planilhas para inspeção manual. Modelos de linguagem podem apoiar sumarização de *clusters*, geração de narrativas e revisão crítica, mas não substituem cálculo bibliométrico nem leitura dos documentos centrais (ARIA; CUCCURULLO, 2017; CHEN, 2006; COBO *et al.*, 2012; VAN ECK; WALTMAN, 2010).

Validação

A validação deve ocorrer em diferentes níveis. No *corpus*, recomenda-se inspeção de amostras aleatórias e amostras dirigidas dos documentos mais citados, mais recentes e mais conectados. Na estratégia de busca, recomenda-se verificar se documentos canônicos esperados foram recuperados. Na interpretação, os rótulos de *clusters* devem ser comparados com títulos, resumos, palavras-chave e documentos representativos. Na análise temporal, tendências devem ser comparadas com mudanças na indexação e com eventos externos relevantes. A validação por especialista do domínio é indispensável quando o estudo gerar conclusões sobre estrutura intelectual ou direção futura da área (CHEN; SONG, 2019; DONTHU *et al.*, 2021).

Produto final

O produto pode ser um artigo de revisão cientométrica, um relatório técnico, um mapa da área, um *dashboard*, uma apresentação para gestores ou uma seção de diagnóstico em projeto estratégico. O produto deve conter resumo executivo, método reprodutível, descrição do *corpus*, indicadores, mapas, interpretação, limitações e anexos com *strings* de busca, *prompts*, parâmetros de rede e versões de ferramentas. Quando a IA for usada, o produto deve indicar em que etapas ela apoiou o processo.

Limitações

As limitações típicas incluem viés de cobertura da base, invisibilidade de literatura não indexada, dependência de termos de busca, ambiguidade terminológica, diferenças disciplinares de publicação, atraso de citações, instabilidade de temas recentes e risco de interpretar *clusters* como comunidades substantivas homogêneas. A análise de área oferece evidências sobre padrões bibliográficos e relacionais; ela não demonstra qualidade científica intrínseca, impacto social ou causalidade entre atores sem dados complementares.

Prompt 20.3.1 Planejamento de análise de área científica

Atue como especialista em cientometria e desenho de estudos bibliométricos.

Quero analisar a área [TEMA] no período [ANO INICIAL–ANO FINAL].

Proponha um protocolo contendo: pergunta principal, perguntas secundárias, bases de dados recomendadas, campos de busca, blocos conceituais de termos, critérios de inclusão e exclusão, metadados necessários, indicadores de produção, impacto, colaboração, estrutura intelectual e evolução temporal, ferramentas adequadas, procedimentos de validação e limitações prováveis.

Separe claramente o que é decisão metodológica, o que é sugestão exploratória e o que exige revisão por especialista.

Prompt 20.3.2 Auditoria de corpus de área científica

Receba a descrição da estratégia de busca, o número de registros por base, a lista dos documentos mais citados e uma amostra aleatória de títulos e resumos.

Avalie se o *corpus* parece adequado ao tema [TEMA].

Identifique possíveis ruídos, lacunas, termos ausentes, documentos canônicos não recuperados, vieses de idioma, vieses de base e riscos de exclusão.

Não proponha conclusões sobre a área; concentre-se na qualidade do *corpus*.

A análise de uma área científica é frequentemente usada como porta de entrada para decisões institucionais, seleção de parcerias, identificação de lacunas e construção de agendas. Por isso, sua forma de apresentação deve ser proporcional ao uso pretendido. Um artigo acadêmico pode enfatizar método e discussão teórica; um relatório institucional pode priorizar mapas, indicadores e implicações; um briefing de gestão deve destacar achados acionáveis e incertezas. Em todos os casos, a fronteira entre descrição e recomendação precisa ser explicitada.

20.4 Protocolo para análise de autores e grupos de pesquisa

Objetivo do protocolo

O protocolo para análise de autores e grupos de pesquisa busca caracterizar trajetórias, perfis de produção, redes de colaboração, influência bibliográfica, especialização temática e inserção institucional de pesquisadores ou coletivos. Ele pode ser usado para compreender a evolução de um grupo, mapear sua contribuição a uma área, apoiar planejamento de colaboração, subsidiar processos de memória institucional ou compor diagnósticos de programas e laboratórios. O protocolo deve evitar reduzir pesquisadores a rankings simples. Em vez disso, deve combinar métricas de produção, impacto, colaboração, coerência temática, diversidade de veículos e contribuição para redes de conhecimento.

Pergunta cientométrica

A pergunta central pode ser: como se caracteriza a trajetória científica de determinado autor ou grupo, considerando produção, impacto, colaboração, temas, veículos e evolução

temporal? Perguntas secundárias podem abordar quais linhas temáticas são mais persistentes, como o grupo se conecta a redes externas, em que períodos ocorrem mudanças de agenda, quais documentos funcionam como marcos e como a produção se distribui entre autoria principal, coautoria e colaboração internacional. Para grupos, é necessário definir quem pertence ao grupo e em que período, pois composições variáveis alteram resultados.

Fontes de dados

A fonte de dados deve permitir desambiguação confiável de autores. ORCID, Scopus Author ID, ResearcherID, perfis institucionais, currículos nacionais, repositórios institucionais e bases bibliográficas podem ser combinados. Em contextos brasileiros, a Plataforma Lattes pode ser relevante, embora seu uso demande procedimentos específicos de extração e normalização. A integração entre bases exige cuidado para evitar duplicatas e inconsistências. A ausência de identificadores persistentes aumenta a necessidade de validação manual, especialmente em casos de homônimos, mudanças de nome, sobrenomes compostos e diferentes convenções de citação (LIN; FENNER, 2013).

Critérios de busca

A busca pode ser construída com identificadores de autor, variações de nome, afiliação, títulos conhecidos e listas curriculares. Para grupos, recomenda-se partir de uma lista nominal validada, com identificadores quando disponíveis. O protocolo deve registrar regras para inclusão de publicações anteriores à entrada do pesquisador no grupo, publicações sem afiliação atual, trabalhos em colaboração externa e documentos sem DOI. Se o objetivo for analisar o grupo enquanto unidade coletiva, pode ser necessário distinguir produção total dos membros e produção atribuível ao grupo ou instituição.

Indicadores

Indicadores úteis incluem produção anual, tipos documentais, periódicos e conferências, citações, citações normalizadas por campo quando disponíveis, índice h com cautela, coautoria média, diversidade de colaboradores, centralidade em redes, internacionalização, interdisciplinaridade e evolução temática. Para grupos, podem ser calculadas medidas de coesão interna, colaboração externa, densidade de rede, dependência de poucos autores e distribuição de liderança. Indicadores de impacto devem ser contextualizados por idade acadêmica, área, tipo de produção e práticas disciplinares (HIRSCH, 2005; WALTMAN, 2016; LEYDESDORFF *et al.*, 2016).

Prompts

Prompts são úteis para organizar perfis narrativos, identificar fases de trajetória, comparar temas ao longo do tempo e gerar perguntas de validação. No entanto, a IA não deve inferir mérito individual a partir de métricas brutas nem produzir juízos avaliativos sobre carreira sem contextualização. *Prompts* devem ser configurados para separar dados observáveis, interpretações prováveis e hipóteses que requerem leitura qualitativa dos trabalhos.

Ferramentas

Além de bibliometrix, VOSviewer, CiteSpace e ferramentas de rede, podem ser usadas planilhas, ORCID, ROR, OpenAlex, Scopus Author Profiles, Web of Science Researcher Profiles

e scripts em R ou Python para combinar registros. Ferramentas de desambiguação e limpeza são fundamentais. A IA pode apoiar a conferência de variantes de nome e a redação de sínteses, mas a validação precisa permanecer humana (ARIA; CUCCURULLO, 2017; VAN ECK; WALTMAN, 2010; PRIEM; PIWOWAR; ORR, 2022).

Validação

A validação deve incluir conferência de amostra de publicações por autor, verificação de homônimos, comparação com currículos ou perfis oficiais, inspeção de documentos com maior impacto e revisão dos mapas de colaboração. Para grupos, recomenda-se validação com coordenador ou pesquisador responsável, especialmente para confirmar pertencimento, linhas de pesquisa e mudanças institucionais. A avaliação de autoria deve respeitar normas disciplinares e não presumir contribuição individual com base apenas na posição na lista de autores (NEWMAN, 2004).

Produto final

O produto pode ser um relatório de trajetória, um mapa de colaboração, um perfil de grupo, um memorial analítico, um diagnóstico de fortalecimento de redes ou um anexo institucional. Deve apresentar resultados de modo contextualizado, evitando linguagem competitiva excessiva. Gráficos de produção e colaboração devem ser acompanhados de notas metodológicas sobre cobertura, identificação de autores e interpretação de métricas.

Limitações

As limitações incluem homonímia, incompletude de perfis, inconsistência de afiliações, variação de práticas de autoria, vantagem cumulativa de pesquisadores seniores, atraso de citações, invisibilidade de contribuições não publicadas e risco de transformar métricas em avaliação individual simplista. O protocolo deve explicitar que a análise cientométrica de autores e grupos descreve padrões bibliográficos, mas não substitui avaliação qualitativa por pares nem compreensão das condições de trabalho e colaboração.

Prompt 20.4.1 Análise de trajetória de autor ou grupo

Atue como cientometrista.

A partir dos dados bibliográficos de [AUTOR/GRUPO], produza uma interpretação em três camadas: descrição objetiva da produção, interpretação contextual provável e hipóteses que exigem validação qualitativa.

Considere produção anual, temas, veículos, colaboração, citações, redes e mudanças temporais.

Não use indicadores como prova direta de qualidade. Indique limitações de identificação de autores, cobertura de bases e diferenças disciplinares.

Prompt 20.4.2 Validação de pertencimento e autoria

> Examine a lista de publicações atribuídas ao grupo [NOME DO GRUPO].
>
> Identifique registros que podem estar incorretamente atribuídos, autores homônimos, afiliações inconsistentes, publicações fora do período de interesse e documentos que exigem conferência manual.
>
> Classifique cada caso como provável inclusão, provável exclusão ou caso ambíguo, sempre justificando a decisão com base nos metadados disponíveis.

Quando aplicado com cuidado, esse protocolo apoia a memória científica e a gestão de grupos sem reduzir a atividade acadêmica a produtividade agregada. A IA pode auxiliar a construir narrativas de trajetória, mas essas narrativas devem ser apresentadas como interpretações baseadas em evidências e não como avaliações definitivas de desempenho. A análise de autores e grupos é particularmente sensível a usos avaliativos; por isso, a seção de limitações deve ser robusta e visível.

20.5 Protocolo para análise institucional

Objetivo do protocolo

A análise institucional procura caracterizar a produção, o impacto, as redes, as áreas de especialização e a posição relacional de uma universidade, centro de pesquisa, programa, departamento, laboratório ou sistema organizacional. Seu objetivo pode ser diagnóstico interno, planejamento estratégico, prestação de contas, identificação de fortalezas, avaliação de colaboração, prospecção de parcerias ou apoio à política institucional de pesquisa. O protocolo deve alinhar os indicadores à missão institucional. Uma instituição regional, por exemplo, pode ter impacto social e contribuição local que não aparecem em indicadores internacionais de citação.

Pergunta cientométrica

A pergunta central pode ser: como se caracteriza o perfil científico da instituição [NOME], considerando produção, áreas, impacto, colaboração, internacionalização, interdisciplinaridade, temas emergentes e inserção em redes? Perguntas secundárias podem examinar quais unidades são mais produtivas em determinadas áreas, quais parcerias são recorrentes, onde há oportunidades de colaboração, que temas cresceram recentemente, como a instituição se posiciona em relação a pares e quais lacunas de visibilidade decorrem da cobertura das bases.

Fontes de dados

A análise institucional exige fontes que registrem afiliações com razoável qualidade. Scopus, Web of Science, Dimensions e OpenAlex podem ser usadas, mas cada uma possui mecanismos próprios de desambiguação institucional. Identificadores como ROR e GRID podem auxiliar. Repositórios institucionais, sistemas internos, currículos e bases nacionais podem complementar a cobertura, sobretudo em áreas com produção local, livros, capítulos, relatórios técnicos ou publicações em português. A escolha das fontes deve considerar tanto a comparabilidade externa quanto a completude interna (MONGEON; PAUL-HUS, 2016; VISSER; VAN ECK; WALTMAN, 2021; MARTÍN-MARTÍN *et al.*, 2021).

Critérios de busca

A busca institucional deve incluir nomes oficiais, variações históricas, siglas, traduções, unidades acadêmicas, campi, hospitais universitários, institutos associados e possíveis erros de grafia. O protocolo deve definir se serão incluídas publicações de autores afiliados no momento da publicação ou autores atualmente vinculados à instituição. Também deve indicar como tratar múltiplas afiliações, instituições parceiras, consórcios, autores visitantes e publicações com afiliação incompleta. Uma regra clara de atribuição evita superestimação ou subestimação da produção institucional.

Indicadores

Indicadores institucionais podem incluir produção por ano, área, unidade, periódico e tipo documental; citações totais e normalizadas; colaboração nacional e internacional; redes de instituições parceiras; distribuição por acesso aberto; interdisciplinaridade; participação em temas emergentes; e alinhamento com áreas estratégicas. Para programas ou departamentos, indicadores precisam ser ajustados à escala e à disciplina. A comparação com instituições pares deve usar critérios transparentes de seleção e evitar rankings simplificados (HICKS *et al.*, 2015; WALTMAN, 2016).

Prompts

Prompts podem apoiar a construção de diagnóstico institucional, a identificação de oportunidades de colaboração, a elaboração de sumários executivos e a adaptação de linguagem para gestores. Devem também ser usados para auditar riscos: vieses de base, invisibilidade de produção local, concentração de resultados em poucos autores, distorções por megacolaborações e extrapolação de métricas. O *prompt* não deve autorizar a IA a classificar unidades como melhores ou piores sem critérios explícitos e sem revisão humana.

Ferramentas

Além de softwares bibliométricos, análises institucionais podem exigir integração com sistemas internos, bases administrativas, identificadores ROR, ferramentas de geocodificação, *dashboards* e bancos relacionais. Python e R podem ser usados para padronização de afiliações, combinação de fontes e geração de indicadores. Ferramentas de visualização como Power BI, Tableau, Observable, R Shiny ou *dashboards* customizados podem comunicar resultados a gestores. A IA pode ajudar a redigir notas interpretativas e explicar indicadores.

Validação

A validação institucional deve combinar auditoria de metadados, consulta a unidades internas, amostragem de publicações, comparação entre bases e revisão por especialistas de áreas. É recomendável validar nomes de instituições parceiras, revisar atribuições de afiliação e conferir documentos altamente citados que podem resultar de grandes consórcios. Quando os resultados forem usados em planejamento ou avaliação, a instituição deve permitir que unidades verificadas corrijam dados antes da publicação do relatório.

Produto final

O produto pode ser um relatório institucional, um painel de indicadores, um mapa de parcerias, um diagnóstico de áreas estratégicas, um briefing para reitoria, uma prestação

de contas ou um anexo para avaliação externa. Deve incluir síntese executiva, metodologia, resultados por dimensão, mapas, limitações, recomendações condicionais e anexos técnicos. A linguagem deve ser clara, sem transformar métricas em juízos absolutos.

Limitações

As limitações incluem desambiguação institucional imperfeita, cobertura desigual entre áreas, sub-representação de produção em língua local, dificuldade de atribuir publicações a unidades internas, efeito de grandes colaborações, variação temporal de nomes institucionais e risco de uso gerencial punitivo. Uma análise institucional responsável deve apoiar aprendizagem organizacional e planejamento, não apenas competição ou vigilância produtivista (HICKS *et al.*, 2015; MONGEON; PAUL-HUS, 2016).

Prompt 20.5.1 Diagnóstico institucional

Atue como especialista em cientometria institucional.

A partir dos indicadores e mapas fornecidos, produza um diagnóstico da instituição [NOME] em quatro blocos: perfil de produção, perfil de impacto e visibilidade, perfil de colaboração e áreas estratégicas.

Para cada bloco, indique evidências, possíveis interpretações, riscos de viés, perguntas adicionais e cautelas para comunicação a gestores.

Não gere ranking moral nem conclua qualidade científica apenas a partir de métricas.

Prompt 20.5.2 Oportunidades de colaboração institucional

Com base na rede de colaboração da instituição [NOME], identifique oportunidades potenciais de parceria.

Separe parceiros consolidados, parceiros emergentes, lacunas geográficas, áreas com baixa conectividade e temas que poderiam justificar novas colaborações.

Para cada oportunidade, indique evidência bibliométrica, hipótese estratégica, dados adicionais necessários e riscos de interpretação.

A análise institucional é uma das aplicações mais sensíveis da cientometria com IA, pois seus resultados podem influenciar decisões de financiamento, contratação, avaliação e reputação. Por esse motivo, os *prompts* devem ser desenhados para ampliar a reflexividade, não para acelerar julgamentos. A seção de limitações precisa ser lida como parte integrante do resultado, e não como apêndice defensivo.

20.6 Protocolo para análise de temas emergentes

Objetivo do protocolo

A análise de temas emergentes busca identificar sinais de novidade, crescimento, reorganização terminológica, aumento de atenção, consolidação de comunidades e possível formação de frentes de pesquisa. Em cientometria, emergência não deve ser entendida apenas como aumento recente de publicações. Um tema pode parecer emergente por mudança de

vocabulário, por reclassificação de base, por entrada de periódicos novos, por expansão de uma tecnologia, por pressão de financiamento ou por associação com eventos sociais. O objetivo do protocolo é distinguir sinais robustos de emergência de artefatos de busca, indexação ou moda discursiva (BORNMANN; MUTZ, 2015; CHEN, 2006; CHEN; SONG, 2019).

Pergunta cientométrica

A pergunta central pode ser: quais temas apresentam sinais de emergência em determinado domínio científico, e quais evidências sustentam essa interpretação? Perguntas secundárias devem examinar crescimento temporal, mudança de termos, surgimento de *clusters*, aumento de centralidade, citações recentes, expansão geográfica, entrada de novos atores, vínculos com tecnologias, políticas ou problemas sociais e relação com literatura anterior. A pergunta deve especificar o horizonte temporal, pois a emergência depende da janela analisada.

Fontes de dados

Fontes com boa atualização temporal são especialmente importantes. Scopus, Web of Science, Dimensions e OpenAlex podem oferecer dados recentes, mas diferenças de atualização e cobertura podem afetar a detecção. *preprints*, anais de congressos, repositórios e bases especializadas podem captar sinais precoces em áreas como computação, física, biomedicina e ciências sociais aplicadas. Entretanto, incluir *preprints* altera a comparabilidade com séries históricas de artigos revisados por pares. O protocolo deve explicitar se a intenção é detectar sinais precoces ou tendências consolidadas (PRIEM; PIWOWAR; ORR, 2022; VISSER; VAN ECK; WALTMAN, 2021).

Critérios de busca

A busca deve combinar termos estabelecidos e termos candidatos emergentes. A IA pode auxiliar a identificar vocabulário novo em resumos, títulos, palavras-chave e literatura cinzenta, mas esses termos precisam ser validados. O protocolo deve prever busca iterativa: iniciar com termos conhecidos, extrair termos frequentes e emergentes, revisar amostras, ajustar a estratégia e documentar as versões. Também é importante testar se o crescimento de um termo resulta de mudança de nomenclatura para fenômeno antigo.

Indicadores

Indicadores úteis incluem taxa de crescimento anual, crescimento relativo em janelas móveis, *burst* de termos ou referências, idade média das publicações, citações recentes, centralidade de *clusters* novos, densidade temática, mudança de palavras-chave, surgimento de autores e instituições, participação em conferências recentes e sobreposição com chamadas de financiamento. Para evitar falsos positivos, recomenda-se combinar indicadores de volume, estrutura e semântica. Um tema com poucos documentos, mas alta coesão e rápido crescimento, pode merecer atenção, mas não deve ser apresentado como consolidado (CHEN, 2006; COBO *et al.*, 2012).

Prompts

Prompts podem apoiar a identificação de temas candidatos, a comparação entre nomes alternativos, a distinção entre novidade real e mudança terminológica e a formulação de hipóteses prospectivas. Devem exigir que o modelo diferencie sinal, evidência, interpretação

e incerteza. A IA não deve declarar que um tema é promissor ou estratégico sem explicitar critérios e sem triangulação com dados externos.

Ferramentas

CiteSpace pode apoiar *burst detection* e visualização temporal; SciMAT pode apoiar evolução temática; bibliometrix oferece mapas temáticos e análise de tendências; VOSviewer pode gerar redes por períodos; modelos de tópicos dinâmicos e técnicas de NLP podem complementar a análise. OpenAlex e Dimensions podem ser úteis para atualização recente. A IA generativa pode auxiliar na rotulagem e narrativa, mas os cálculos de emergência devem ser reproduzíveis (ARIA; CUCCURULLO, 2017; CHEN, 2006; COBO *et al.*, 2012).

Validação

A validação deve incluir inspeção de documentos recentes representativos, comparação com revisões ou editoriais da área, consulta a especialistas, verificação de termos em bases diferentes e teste de sensibilidade temporal. É necessário verificar se a emergência observada se mantém quando a estratégia de busca muda, quando se excluem documentos de baixa relevância ou quando se usam bases alternativas. A triangulação com financiamento, patentes, conferências e políticas pode reforçar a interpretação, mas cada fonte tem vieses próprios.

Produto final

O produto pode ser um radar de tendências, um relatório prospectivo, um mapa de temas emergentes, um briefing estratégico, uma seção de agenda futura ou um painel de monitoramento. O produto deve classificar os temas por nível de evidência, não apenas por atratividade. Pode-se usar categorias como sinal fraco, tendência incipiente, tema em consolidação e tema consolidado. Cada categoria deve ter critérios explícitos.

Limitações

As limitações incluem viés de recência, atrasos de indexação, instabilidade terminológica, influência de *hype*, baixa comparabilidade entre tipos documentais, ruído de *preprints*, diferenças disciplinares e risco de profecia autorrealizável. Ao chamar um tema de emergente, o analista pode influenciar prioridades e recursos. Por isso, a linguagem deve ser cuidadosa e probabilística (CHEN; SONG, 2019).

Prompt 20.6.1 Detecção de temas emergentes

Atue como analista cientométrico. A partir da série temporal de termos, *clusters*, documentos e citações, identifique candidatos a temas emergentes.

Para cada candidato, apresente: evidências de crescimento, evidências de novidade semântica, atores envolvidos, relação com temas anteriores, nível de confiança, possíveis artefatos de busca e dados adicionais necessários.

Não confunda aumento de frequência com emergência substantiva.

Prompt 20.6.2 Diferenciação entre tendência real e artefato

> Examine a tendência observada para o termo [TERMO].
>
> Avalie se o crescimento pode ser explicado por mudança de vocabulário, ampliação da base, entrada de novos periódicos, evento externo, erro de busca, polissemia ou real consolidação temática.
>
> Organize a resposta em evidência disponível, explicações alternativas, testes recomendados e conclusão provisória.

A análise de temas emergentes é uma ponte entre cientometria, prospecção e política científica. Seu valor está menos em prever o futuro e mais em organizar sinais presentes de forma transparente. A IA pode ampliar a capacidade de leitura e comparação de termos, mas a prudência metodológica continua sendo indispensável.

20.7 Protocolo para análise de colaboração internacional

Objetivo do protocolo

A análise de colaboração internacional busca compreender como pesquisadores, instituições e países se conectam na produção científica. Ela pode revelar padrões de coprodução, dependências assimétricas, redes regionais, parcerias estratégicas, circulação de conhecimento e inserção internacional de áreas ou instituições. O protocolo deve distinguir colaboração nominal, baseada em coautoria com múltiplas afiliações, de colaboração substantiva, que exigiria dados qualitativos adicionais. A presença de autores de diferentes países em um artigo não explica por si só a natureza da parceria, a distribuição de trabalho ou os benefícios científicos (NEWMAN, 2004).

Pergunta cientométrica

A pergunta central pode ser: como se estrutura a colaboração internacional em determinado domínio, instituição ou país, considerando intensidade, diversidade, centralidade, áreas temáticas e evolução temporal? Perguntas secundárias podem examinar quais países são parceiros principais, quais instituições atuam como pontes, que temas concentram colaboração internacional, como as redes mudaram ao longo do tempo e se há dependência de poucos parceiros. A pergunta deve definir se o foco é país, instituição, área, programa ou tema.

Fontes de dados

A análise depende de metadados de afiliação e país. Bases comerciais e abertas apresentam diferenças na padronização de afiliações. OpenAlex, Scopus e Web of Science podem oferecer dados de instituições e países, mas os resultados exigem limpeza. ROR pode apoiar desambiguação institucional. Para alguns países, bases nacionais podem capturar produção local não indexada internacionalmente. O protocolo deve indicar como tratar documentos com múltiplas afiliações, autores com dupla afiliação e instituições transnacionais.

Critérios de busca

Os critérios de busca devem definir o *corpus* temático ou institucional e o período de análise. Em seguida, devem ser extraídas afiliações por documento e autor. O protocolo deve

explicitar o método de contagem: contagem inteira, fracionária, por documento, por autor ou por endereço. A escolha altera conclusões. Grandes colaborações internacionais podem dominar resultados se a contagem inteira for usada sem controle. Em áreas como física de altas energias e genômica, esse efeito pode ser significativo.

Indicadores

Indicadores incluem número e proporção de publicações internacionais, diversidade de países parceiros, rede de coautoria entre países, centralidade, grau, intermediação, modularidade, densidade, colaboração regional, documentos com múltiplos países, impacto relativo de publicações internacionais e distribuição temática das parcerias. Também podem ser analisadas trajetórias temporais de entrada e saída de parceiros, estabilidade de vínculos e concentração de colaboração (NEWMAN, 2004; WALTMAN; VAN ECK; NOYONS, 2010).

Prompts

Prompts podem apoiar interpretação de redes, identificação de padrões de dependência, elaboração de hipóteses sobre oportunidades e comunicação de resultados a gestores. Devem impedir inferências causais não sustentadas. A IA pode sugerir que parcerias internacionais ampliam visibilidade, mas essa afirmação precisa ser tratada como hipótese ou relação associativa, não como causalidade automática.

Ferramentas

VOSviewer, Gephi, bibliometrix, R/igraph, Python/networkx e ferramentas GIS podem apoiar mapas de rede e geográficos. *Dashboards* podem integrar indicadores por país, instituição e tema. A IA pode apoiar a redação de narrativas e a identificação de perguntas de validação, mas a análise de redes deve ser calculada com ferramentas apropriadas e parâmetros documentados (ARIA; CUCCURULLO, 2017; VAN ECK; WALTMAN, 2010).

Validação

A validação deve verificar a qualidade das afiliações, normalização de países, duplicação de instituições, efeito de megacolaborações e consistência entre bases. É recomendável revisar os documentos que geram conexões mais fortes e verificar se pertencem a projetos multinacionais específicos. Também é importante comparar redes com dados institucionais de convênios, mobilidade, projetos financiados ou programas de cooperação, quando disponíveis.

Produto final

O produto pode ser um mapa de colaboração internacional, um relatório de inserção global, uma análise de oportunidades de parceria, um diagnóstico de internacionalização ou um briefing estratégico. Deve incluir mapas de rede, indicadores de intensidade e diversidade, interpretação por área e limitações de coautoria como *proxy* de colaboração.

Limitações

As limitações incluem uso de coautoria como aproximação imperfeita de colaboração, afiliações incompletas, mudanças geopolíticas, diferenças de idioma, megacolaborações, desigualdades centro-periferia, invisibilidade de colaboração informal e risco de valorizar

internacionalização como fim em si mesmo. O protocolo deve lembrar que colaboração internacional deve ser avaliada em relação à missão científica, social e institucional do estudo.

Prompt 20.7.1 Mapeamento de oportunidades de colaboração internacional

> Atue como analista de redes científicas.
>
> A partir da matriz de colaboração por países e instituições, identifique padrões de colaboração consolidados, conexões frágeis, parceiros emergentes e lacunas geográficas.
>
> Para cada achado, apresente evidência, interpretação provável, hipótese estratégica, dados adicionais necessários e risco de inferência indevida.
>
> Não afirme que colaboração causa impacto sem teste específico.

Prompt 20.7.2 Auditoria de megacolaborações

> Examine os documentos que mais influenciam a rede internacional.
>
> Identifique se há megacolaborações, consórcios, artigos com número excepcional de autores ou múltiplas afiliações que possam distorcer a interpretação.
>
> Sugira análises com e sem esses documentos, contagem inteira e contagem fracionária, e explique como cada escolha pode alterar os resultados.

O protocolo de colaboração internacional é particularmente útil para políticas de internacionalização, mas deve ser usado com cautela. A internacionalização não é automaticamente sinônimo de qualidade, relevância ou equidade. A cientometria pode mostrar padrões de conexão; a interpretação deve considerar objetivos científicos, assimetrias de recursos, língua, mobilidade, financiamento e prioridades locais.

20.8 Protocolo para relatório de política científica

Objetivo do protocolo

O relatório de política científica é um produto de tradução entre análise cientométrica e tomada de decisão. Seu objetivo é organizar evidências sobre áreas, instituições, temas, redes ou programas para subsidiar escolhas públicas ou institucionais. Diferentemente de um artigo acadêmico, o relatório de política científica precisa responder a problemas de decisão: onde investir, que capacidades fortalecer, que lacunas monitorar, que parcerias estimular, que indicadores acompanhar e que riscos evitar. Contudo, evidência cientométrica não determina automaticamente política; ela informa cenários e alternativas (MOED, 2009; HICKS *et al.*, 2015).

Pergunta cientométrica

A pergunta central pode ser: que evidências cientométricas podem apoiar decisões de política científica relativas a determinado tema, área, instituição ou sistema? Perguntas secundárias podem examinar capacidade instalada, vantagens comparativas, dependência externa, temas emergentes, lacunas estratégicas, cooperação internacional, alinhamento com demandas sociais, distribuição regional e impacto potencial. A pergunta deve ser formulada com participação dos interessados, para evitar que indicadores respondam a problemas irrelevantes.

Fontes de dados

Além de bases bibliográficas, relatórios de política podem combinar dados de financiamento, patentes, pós-graduação, repositórios, currículos, projetos, infraestrutura, dados socioeconômicos e indicadores de inovação. A integração de dados bibliográficos com dados administrativos exige governança, documentação e respeito a privacidade. Bases como OpenAlex, Dimensions, Scopus, Web of Science, Crossref, Lens e bases nacionais podem cumprir funções diferentes. A escolha deve ser guiada pelo problema de decisão, não apenas pela disponibilidade (MONGEON; PAUL-HUS, 2016; VISSER; VAN ECK; WALTMAN, 2021).

Critérios de busca

Os critérios de busca devem ser documentados de forma que um leitor externo compreenda o perímetro do diagnóstico. Em política científica, a delimitação do tema pode ter consequências distributivas. Excluir termos locais, publicações em português ou áreas aplicadas pode reduzir a visibilidade de capacidades relevantes. O protocolo deve prever consulta a especialistas, revisão de termos, testes de sensibilidade e exposição transparente das escolhas.

Indicadores

Indicadores podem incluir capacidade de produção, massa crítica, impacto, colaboração, internacionalização, interdisciplinaridade, alinhamento temático, emergência, distribuição regional, relação com patentes, participação em redes e diversidade de atores. Devem ser usados em painel, não isoladamente. O relatório deve destacar incertezas e evitar transformar diferenças pequenas em recomendações fortes. Indicadores de política devem ser contextualizados por missão, recursos, desigualdades e objetivos públicos (MOED, 2009; WALTMAN, 2016).

Prompts

Prompts podem apoiar a transformação de resultados técnicos em briefing, a identificação de riscos, a redação de opções de política e a adaptação de linguagem. Devem ser formulados para separar evidência, interpretação e recomendação. A IA pode ajudar a organizar alternativas, mas não deve decidir prioridades. Um bom *prompt* de política científica exige explicitar valores, critérios e limites do uso dos dados.

Ferramentas

Relatórios de política podem combinar softwares bibliométricos, bancos de dados, painéis de BI, ferramentas GIS, modelos de linguagem, planilhas auditáveis e repositórios de código. A governança do fluxo é tão importante quanto a ferramenta. O protocolo deve indicar onde dados brutos serão armazenados, como versões serão controladas, quem revisará os resultados e como os interessados terão acesso às evidências.

Validação

A validação deve ser técnica, substantiva e deliberativa. Técnica: verificar dados, indicadores e visualizações. Substantiva: consultar especialistas de domínio. Deliberativa: discutir interpretações com gestores e atores afetados. Relatórios de política científica podem orientar recursos; por isso, precisam de transparência e abertura a contestação. Quando possível, recomenda-se publicar anexos metodológicos e *datasets* agregados (HICKS *et al.*, 2015).

Produto final

O produto final deve conter sumário executivo, problema de decisão, método, evidências centrais, cenários, opções de ação, riscos, limitações, indicadores de monitoramento e anexos técnicos. O relatório deve distinguir achados descritivos, interpretações analíticas e recomendações. Recomendações devem ser condicionadas: se o objetivo for fortalecer colaboração internacional, então determinados padrões sugerem oportunidades; se o objetivo for consolidar capacidade local, outros indicadores podem ser mais relevantes.

Limitações

As limitações incluem redução de problemas complexos a métricas, pressão por respostas rápidas, seletividade de bases, influência de agendas políticas, risco de reforçar desigualdades e possibilidade de uso performativo de indicadores. A IA adiciona riscos de opacidade, formulação persuasiva excessiva e falsa segurança. O protocolo deve assegurar que o relatório apresente incertezas de forma clara e que decisões finais permaneçam sob responsabilidade humana e institucional (HICKS *et al.*, 2015; LEYDESDORFF *et al.*, 2016).

Prompt 20.8.1 Briefing para tomada de decisão

Atue como analista de política científica.

A partir dos resultados cientométricos fornecidos, produza um briefing com: problema de decisão, evidências principais, interpretações prudentes, opções de ação, riscos, dados adicionais necessários, indicadores de monitoramento e limitações. Separe claramente achados descritivos, inferências e recomendações condicionais.

Não apresente a recomendação como consequência automática dos indicadores.

Prompt 20.8.2 Revisão ética de relatório de política científica

Revise o relatório abaixo quanto a riscos éticos e metodológicos.

Identifique usos indevidos de indicadores, extrapolações causais, invisibilidade de grupos ou áreas, viés linguístico/geográfico, excesso de certeza, ausência de limitações, linguagem de ranking sem critério e dependência não documentada de IA.

Sugira ajustes de redação e documentação.

O relatório de política científica é a forma mais diretamente normativa da cientometria aplicada. Por isso, sua qualidade não se mede apenas pela precisão técnica, mas também pela responsabilidade comunicacional. Um relatório pode ser tecnicamente sofisticado e politicamente inadequado se esconder limitações, naturalizar desigualdades ou sugerir ações sem explicitar valores e incertezas. A IA deve ser usada para ampliar clareza, consistência e revisão crítica, não para fabricar autoridade.

20.9 Protocolos comparados: quando usar cada um

A escolha do protocolo deve ser guiada pela finalidade do estudo. Muitas pesquisas combinam protocolos. Um diagnóstico institucional pode conter análise de áreas científicas, temas emergentes e colaboração internacional. Um relatório de política científica pode exigir todos os protocolos anteriores em escala reduzida. A combinação, entretanto, deve ser planejada. Não é adequado acumular indicadores de diferentes protocolos sem explicar como se relacionam à pergunta central.

Quadro 20.2 - Comparação entre os protocolos

Protocolo	Uso principal	Unidade de análise dominante	Produto típico
Área científica	Compreender estrutura, evolução e temas de um domínio.	Documentos, termos, referências, autores e *clusters*.	Artigo, relatório de área, mapa temático.
Autores e grupos	Caracterizar trajetória, colaboração e perfil temático.	Autores, grupos, documentos e redes de coautoria.	Perfil analítico, relatório de grupo, mapa de colaboração.
Instituição	Diagnosticar produção, impacto, redes e áreas estratégicas.	Instituições, unidades, autores, países e temas.	Relatório institucional, *dashboard*, briefing.
Temas emergentes	Identificar sinais de novidade e mudança temática.	Termos, *clusters*, documentos recentes e séries temporais.	Radar de tendências, mapa prospectivo.
Colaboração internacional	Mapear redes, parceiros e internacionalização.	Países, instituições, documentos e autores.	Mapa de colaboração, diagnóstico de parcerias.
Política científica	Traduzir evidências em opções de decisão.	Áreas, instituições, temas, indicadores e cenários.	Relatório de política, briefing, painel de monitoramento.

Fonte: Elaborado pelo autor (2026).

20.10 Documentação reprodutível dos protocolos

Um protocolo cientométrico com IA deve gerar documentação suficiente para que outro analista compreenda como os resultados foram produzidos. Essa documentação não precisa garantir replicação perfeita, especialmente quando modelos generativos mudam com o tempo, mas deve permitir auditoria. Para isso, recomenda-se manter um caderno de protocolo com versões de perguntas, *strings* de busca, datas de coleta, bases usadas, filtros, número de registros recuperados, regras de deduplicação, procedimentos de limpeza, parâmetros de rede, *prompts*, respostas relevantes da IA, decisões humanas e justificativas (PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021).

A documentação dos *prompts* é particularmente importante. O registro deve incluir o texto do *prompt*, o modelo usado, a data, os dados fornecidos, as restrições impostas, o formato de saída esperado, a resposta gerada, as correções humanas e a decisão final. Quando a IA for usada apenas para redação, isso deve ser indicado. Quando for usada para classificação,

rotulagem ou apoio interpretativo, o protocolo deve incluir validação por amostra e, quando possível, comparação entre avaliadores humanos.

Protocolos reprodutíveis também exigem preservação dos dados brutos. A base exportada originalmente deve ser armazenada separadamente da base limpa. Cada transformação deve ser documentada. Registros excluídos devem ser rastreáveis. Dicionários de normalização de autores, instituições e palavras-chave devem ser mantidos. Em estudos com APIs, é necessário registrar *endpoint*, parâmetros, versão da API, data de consulta, limites, paginação e eventuais falhas de coleta.

Em relação a ferramentas, o protocolo deve registrar versões de software, pacotes, parâmetros e layouts. Em VOSviewer, por exemplo, escolhas como método de contagem, limiar mínimo de ocorrência, normalização e resolução de *cluster*ing alteram mapas. Em bibliometrix, a seleção de campos e funções influencia resultados. Em redes, medidas de centralidade dependem do tipo de grafo, direção, peso e tratamento de componentes. A IA pode ajudar a redigir a documentação, mas não deve substituir o registro técnico (ARIA; CUCCURULLO, 2017; VAN ECK; WALTMAN, 2010).

Prompt 20.10.1 Transformar saída textual em protocolo reprodutível

Transforme a descrição abaixo em um protocolo reprodutível.

Organize em: objetivo, pergunta, *corpus*, bases, estratégia de busca, critérios de inclusão e exclusão, metadados, limpeza, indicadores, ferramentas, parâmetros, *prompts* usados, validação, produtos, limitações e anexos necessários.

Identifique lacunas de informação que impedem a reprodução do estudo.

Prompt 20.10.2 Auditoria de documentação

Avalie se o protocolo abaixo é suficientemente documentado para auditoria.

Verifique se contém data de busca, bases, *strings*, filtros, número de registros, deduplicação, limpeza, indicadores, parâmetros, versões de ferramentas, *prompts*, validação e limitações.

Classifique cada item como adequado, incompleto ou ausente, e sugira correções.

20.11 Qualidade, validação e triangulação

A qualidade de um estudo cientométrico com IA não depende apenas da sofisticação dos indicadores. Depende da coerência entre pergunta, dado, método, interpretação e uso. Um protocolo pode falhar por excesso de complexidade, quando utiliza redes e métricas avançadas sem necessidade; por subespecificação, quando não documenta busca e limpeza; por viés de fonte, quando trata uma base como espelho completo da ciência; por inferência indevida, quando transforma coocorrência em causalidade; ou por retórica de IA, quando respostas geradas por modelo são apresentadas como análise validada (DONTHU *et al.*, 2021; HICKS *et al.*, 2015).

A validação deve ser planejada desde o início. Recomenda-se usar validação de *corpus*, validação de metadados, validação de indicadores, validação de rótulos, validação interpretativa e validação comunicacional. A validação de *corpus* pergunta se os documentos recuperados correspondem ao tema. A validação de metadados pergunta se autores, instituições, referências

e palavras-chave estão corretos. A validação de indicadores pergunta se as métricas são adequadas à pergunta. A validação de rótulos pergunta se *clusters* foram nomeados de forma plausível. A validação interpretativa pergunta se conclusões estão sustentadas. A validação comunicacional pergunta se o produto final transmite incertezas e limites (CHEN; SONG, 2019; DONTHU *et al.*, 2021).

Quadro 20.3 - Tipos de validação e evidências de controle

Tipo de validação	Pergunta de controle	Exemplo de evidência
Corpus	Os documentos recuperados correspondem ao escopo?	Amostra de títulos/resumos; documentos canônicos incluídos.
Metadados	Autores, instituições e referências estão limpos?	Dicionário de normalização; taxa de duplicatas; revisão manual.
Indicadores	As métricas respondem à pergunta?	Justificativa teórica; comparação com alternativas.
Rótulos	*Clusters* foram nomeados de modo plausível?	Documentos representativos; revisão por especialista.
Interpretação	Conclusões derivam dos dados?	Matriz dado-inferência-hipótese; auditoria de causalidade.
Comunicação	O produto é claro e responsável?	Sumário executivo com limitações visíveis; revisão por público-alvo.

Fonte: Elaborado pelo autor (2026).

A triangulação é o mecanismo que impede a dependência excessiva de uma única fonte ou métrica. Um tema emergente pode ser analisado por crescimento de publicações, *burst* de termos, entrada de novos autores, aumento de citações recentes, presença em conferências e opinião de especialistas. Uma instituição pode ser avaliada por produção, redes, áreas, dados internos, repositórios e missão institucional. Um relatório de política científica deve combinar métricas, literatura substantiva e conhecimento dos atores. A IA pode ajudar a organizar a triangulação, mas a decisão sobre suficiência das evidências é humana (COBO *et al.*, 2011; CHEN, 2006).

Em protocolos com IA, a validação também deve considerar estabilidade de respostas. O mesmo *prompt* pode produzir saídas diferentes em execuções distintas. Para tarefas críticas, convém usar instruções estruturadas, temperatura baixa quando disponível, formatos tabulares, critérios explícitos e revisão por amostra. Em classificações temáticas, pode-se comparar saídas de rodadas independentes, medir concordância com especialistas e registrar casos ambíguos. Em redação interpretativa, deve-se auditar se a IA introduziu afirmações não presentes nos dados.

20.12 Matriz de prompts transversais

Os protocolos específicos podem ser complementados por *prompts* transversais. Eles funcionam como instrumentos de controle de qualidade e podem ser aplicados em qualquer tipo de estudo. A seguir são apresentados modelos que sintetizam práticas discutidas ao longo do livro.

Prompt 20.12.1 Seleção de indicadores

Dada a pergunta cientométrica [PERGUNTA], proponha indicadores adequados.

Para cada indicador, informe: dimensão analisada, unidade de análise, dado necessário, fórmula ou lógica de cálculo, interpretação possível, limitações, riscos de uso indevido e visualização recomendada.

Evite sugerir indicadores que não respondam à pergunta.

Prompt 20.12.2 Separar dado, inferência e hipótese

A partir dos resultados abaixo, organize a interpretação em três colunas: dado observado, inferência plausível e hipótese que exige evidência adicional.

Sinalize afirmações fortes demais, relações causais não demonstradas e conclusões que dependem de contexto externo.

Prompt 20.12.3 Revisão de limitações

Revise a seção de resultados abaixo e construa uma seção de limitações.

Considere cobertura de base, estratégia de busca, recorte temporal, idiomas, tipos documentais, qualidade de metadados, indicadores usados, parâmetros de rede, uso de IA, validação humana e generalização dos achados.

Prompt 20.12.4 Adaptação para público não técnico

Reescreva a interpretação abaixo para gestores não especialistas.

Mantenha rigor, explique indicadores em linguagem acessível, preserve incertezas, evite jargões desnecessários e não transforme evidências exploratórias em recomendações definitivas.

Prompt 20.12.5 Checagem de falsa causalidade

Analise o texto abaixo e identifique trechos em que associação, correlação, coocorrência, centralidade ou crescimento são apresentados como causalidade.

Para cada trecho, proponha uma redação alternativa mais prudente e indique que evidência seria necessária para sustentar causalidade.

20.13 Estudo de caso ilustrativo: da pergunta ao produto final

Considere uma agência pública interessada em compreender a evolução da pesquisa sobre inteligência artificial aplicada à saúde pública. O objetivo declarado é subsidiar um programa de fomento. A equipe decide combinar três protocolos: análise de área científica, detecção de temas emergentes e relatório de política científica. O primeiro passo é formular a pergunta: como se configurou a produção científica sobre IA em saúde pública entre 2014 e 2025, quais temas emergiram e que evidências podem orientar prioridades de fomento? A pergunta já indica domínio, período, temas e finalidade decisória.

Na etapa de fontes, a equipe seleciona Scopus, Web of Science e OpenAlex para comparação, além de PubMed para cobertura biomédica. A estratégia de busca é construída em blocos: inteligência artificial, aprendizado de máquina, modelos de linguagem, saúde pública, vigilância epidemiológica, sistemas de saúde, promoção da saúde e políticas de saúde. A IA é usada para expandir termos, mas a lista final é revisada por especialistas em saúde pública e ciência da informação. O protocolo registra três versões da *string* e explica por que determinados termos foram excluídos por gerar ruído.

A coleta produz quatro bases brutas. A equipe preserva arquivos originais e cria uma base integrada. A deduplicação usa DOI, título normalizado, ano e autores. A limpeza padroniza instituições, países e palavras-chave. A IA auxilia a agrupar termos como *machine learning*, ML, aprendizado de máquina e modelos preditivos, mas casos ambíguos são marcados para revisão. A documentação registra decisões de normalização e mantém tabela de termos preferenciais.

Os indicadores incluem crescimento anual, distribuição por áreas, países, instituições, periódicos, citações, redes de coautoria, coocorrência de palavras-chave, acoplamento bibliográfico e análise temporal de termos. Para temas emergentes, a equipe calcula crescimento em janelas de três anos, identifica *bursts* de termos e compara *clusters* recentes com *clusters* anteriores. A IA é usada para nomear *clusters*, mas cada rótulo é validado por leitura de documentos representativos.

Na interpretação, os resultados são organizados em três níveis: descrição, interpretação e hipótese. A descrição mostra aumento de publicações em determinados períodos, crescimento de termos associados a vigilância digital e modelos preditivos, e forte colaboração entre algumas regiões. A interpretação sugere que a agenda se deslocou de aplicações genéricas de aprendizado de máquina para usos em monitoramento, predição e apoio a decisões. A hipótese prospectiva indica que modelos de linguagem podem reconfigurar práticas de revisão, triagem e comunicação em saúde pública, mas essa hipótese exige análise qualitativa e acompanhamento temporal.

O relatório final apresenta sumário executivo, método, achados, temas emergentes, capacidades institucionais, lacunas, oportunidades de cooperação e recomendações condicionais. O texto evita dizer que a área deve receber recursos simplesmente porque cresceu. Em vez disso, relaciona crescimento, capacidade nacional, relevância social, lacunas de infraestrutura, riscos éticos e necessidade de avaliação. O anexo inclui *strings* de busca, *prompts*, parâmetros de rede, versões de ferramentas e limitações. O estudo de caso ilustra como protocolos evitam improvisação e permitem que a IA seja usada com transparência.

20.14 Erros frequentes na aplicação de protocolos

- Escolher indicadores antes de formular a pergunta. Esse erro transforma o estudo em vitrine de métricas, sem coerência analítica. O protocolo deve começar por objetivo e pergunta.
- Tratar uma base de dados como representação completa da ciência. Toda base tem critérios de indexação, cobertura linguística, lacunas disciplinares e limitações de metadados.
- Usar IA para gerar strings de busca sem revisão. Modelos podem sugerir termos inexistentes, inadequados, ambíguos ou excessivamente amplos.
- Não preservar dados brutos. Sem arquivos originais, torna-se difícil auditar exclusões, deduplicações e transformações.

- Rotular clusters apenas pela saída da IA. Rótulos devem ser validados por documentos representativos e conhecimento de domínio.
- Confundir colaboração por coautoria com colaboração substantiva. Coautoria é *proxy*, não descrição completa da prática colaborativa.
- Transformar tendências em previsões. Análise temporal produz hipóteses e sinais, não certezas sobre o futuro.
- Apresentar recomendações sem explicitar valores e critérios. Em política científica, dados informam decisões, mas não substituem deliberação.
- Omitir *prompts* e configurações. Se a IA contribuiu para análise, classificação, interpretação ou redação, seu uso deve ser documentado.
- Esconder limitações no final do relatório. Limitações devem acompanhar resultados e recomendações, sobretudo quando o estudo orienta decisões.

20.15 Checklist final para protocolos completos

Quadro 20.4 - Checklist final para protocolos completos

CHECKLIST FINAL PARA PROTOCOLOS COMPLETOS *(antes de encerrar o protocolo do estudo cientométrico)*	
☐	01. Objetivo substantivo do estudo foi explicitado e diferenciado do produto final esperado.
☐	02. Pergunta principal e perguntas secundárias estão alinhadas ao objetivo.
☐	03. Unidades de análise foram declaradas: documentos, autores, instituições, países, termos, referências ou *clusters*.
☐	04. Fontes de dados foram justificadas em termos de cobertura, qualidade, acesso e finalidade.
☐	05. *Strings* de busca, filtros, campos e datas foram registrados.
☐	06. Critérios de inclusão e exclusão foram definidos antes da análise interpretativa.
☐	07. Metadados necessários foram listados e sua completude foi verificada.
☐	08. Procedimentos de deduplicação e limpeza foram documentados.
☐	09. Indicadores foram selecionados por adequação à pergunta, não por disponibilidade automática.
☐	10. Ferramentas, versões e parâmetros foram registrados.
☐	11. *Prompts* usados foram armazenados com respostas, revisões e decisões humanas.
☐	12. Validação de *corpus*, metadados, indicadores, rótulos e interpretação foi planejada.
☐	13. Produtos finais foram adaptados ao público-alvo sem perder rigor.
☐	14. Limitações foram incorporadas à interpretação e às recomendações.
☐	15. Associações não foram apresentadas como causalidade sem evidência adicional.
☐	16. Uso de IA foi transparente, auditável e subordinado à responsabilidade autoral humana.

Registro da verificação

Protocolo ou versão verificada	Responsável	Data

Fonte: Elaborado pelo autor (2026).

20.16 Considerações finais

Protocolos completos de estudo são a condição prática para que a cientometria com inteligência artificial seja rigorosa, transparente e útil. Eles convertem a multiplicidade de tarefas - formular pergunta, buscar documentos, limpar dados, calcular indicadores, mapear redes, interpretar *clusters*, redigir resultados e comunicar decisões - em uma sequência documentada de escolhas. Sem protocolo, a IA tende a reforçar a aparência de fluidez: respostas rápidas, textos coerentes e sugestões plausíveis podem ocultar fragilidade de *corpus*, ausência de validação e extrapolações indevidas. Com protocolo, a IA se torna componente controlado de uma arquitetura metodológica mais ampla.

Os seis protocolos apresentados cobrem usos recorrentes da cientometria aplicada. A análise de área científica permite compreender estrutura e evolução de domínios. A análise de autores e grupos apoia memória, colaboração e diagnóstico, desde que evite reducionismo avaliativo. A análise institucional orienta planejamento e visibilidade, mas precisa respeitar missão e contexto. A análise de temas emergentes organiza sinais de mudança, sem prometer previsão. A análise de colaboração internacional revela redes e oportunidades, sem presumir causalidade ou qualidade. O relatório de política científica traduz evidências em cenários e opções, sem substituir deliberação.

Em todos os protocolos, os *prompts* desempenham papel de mediação. Eles ajudam a estruturar raciocínio, formular alternativas, revisar riscos, adaptar linguagem e documentar decisões. Entretanto, *prompts* não são atalhos para evidência. A resposta da IA deve ser tratada como proposta, não como resultado final. A validação humana, a triangulação de fontes e a transparência metodológica continuam sendo requisitos centrais.

Um livro sobre cientometria com IA deve insistir nessa tese: a inovação não está apenas em usar modelos generativos, mas em redesenhar práticas de análise para que a IA seja incorporada de forma responsável. Protocolos completos são instrumentos dessa responsabilidade. Eles permitem que pesquisadores e gestores saibam não apenas o que foi encontrado, mas como foi encontrado, com que dados, sob quais premissas, com que incertezas e com que limites. Essa é a diferença entre automação bibliométrica e cientometria cientificamente orientada.

Apêndice A.20 - Operacionalização, governança e anexos dos protocolos

Este apêndice reúne as orientações operacionais que, no manuscrito original, estavam distribuídas em dois apêndices e em uma nota final. A consolidação elimina repetições, preserva o conteúdo metodológico e organiza a implementação em sete componentes: papéis, maturidade, governança de dados, parâmetros analíticos, ciclos de revisão, anexos e melhoria contínua.

A.20.1 - Papéis e responsabilidades

A operacionalização dos protocolos requer uma divisão clara de responsabilidades. Em estudos cientométricos com IA, raramente uma única pessoa domina todas as dimensões envolvidas: conhecimento substantivo da área, estratégia de busca, metadados bibliográficos, estatística, análise de redes, visualização, linguagem de programação, redação científica, governança de dados e ética. Por isso, o protocolo deve indicar papéis. O pesquisador de domínio valida vocabulário e interpretações; o bibliotecário ou cientista da informação revisa a estratégia de busca; o analista de dados implementa limpeza e indicadores; o cientometrista conecta métodos e perguntas; o gestor define uso decisório; e a IA atua como apoio textual, heurístico e analítico limitado.

Quadro 20.5 - Distribuição recomendada de responsabilidades

Papel	Responsabilidade principal
Coordenação científica	Define a pergunta, o escopo e os critérios substantivos; aprova alterações no protocolo.
Especialista de domínio	Valida vocabulário, documentos representativos, rótulos e interpretações.
Bibliotecário ou cientista da informação	Revisa fontes, estratégia de busca, campos, filtros e documentação da recuperação.
Analista de dados/ especialista em cientometria	Executa limpeza, normalização, indicadores, redes, testes de sensibilidade e visualizações.
Gestor ou usuário decisório	Explicita a decisão a apoiar, os valores envolvidos e as condições de uso do produto.
Sistema de IA	Apoia tarefas textuais e heurísticas delimitadas; não assume decisão metodológica, autoria científica ou julgamento final.

Fonte: Elaborado pelo autor (2026).

A.20.2 - Níveis de maturidade e critérios de progressão

Um protocolo também deve prever níveis de maturidade do estudo. No nível exploratório, o objetivo é compreender o perímetro de um tema, testar termos e identificar fontes. No nível analítico, o *corpus* já está estabilizado e os indicadores são calculados com parâmetros documentados. No nível interpretativo, os achados são triangulados com literatura e especialistas. No nível decisório, os resultados são traduzidos em cenários, recomendações condicionais e indicadores de monitoramento. A IA pode participar de todos os níveis, mas a exigência de validação aumenta conforme o resultado se aproxima de avaliação institucional ou política científica.

Quadro 20.6 - Níveis de maturidade do estudo

Nível	Finalidade	Condição mínima para avançar
Exploratório	Delimitar tema, testar termos, identificar fontes e avaliar viabilidade.	Vocabulário e escopo preliminar validados.
Analítico	Estabilizar o *corpus*, executar limpeza e calcular indicadores com parâmetros registrados.	Base limpa, dicionários e scripts auditáveis.

Nível	Finalidade	Condição mínima para avançar
Interpretativo	Triangular resultados com literatura, documentos e especialistas.	Matriz de evidências e divergências registrada.
Decisório	Traduzir resultados em cenários, recomendações condicionais e monitoramento.	Revisão ética, validação deliberativa e limitações visíveis.

Fonte: Elaborado pelo autor (2026).

A.20.3 - Governança de dados e curadoria de metadados

A governança dos dados deve ser definida antes da coleta. O protocolo deve informar onde os arquivos brutos serão armazenados, quem terá acesso, como versões serão nomeadas, que campos serão preservados e que procedimentos serão adotados para dados sensíveis. Embora metadados bibliográficos sejam geralmente públicos, análises institucionais podem envolver dados internos, identificação de indivíduos, produtividade de grupos pequenos ou informações estratégicas. Nesses casos, a equipe deve decidir se os resultados serão agregados, anonimizados ou restritos. A IA não deve receber dados sensíveis quando não houver base legal, consentimento ou segurança adequada.

A curadoria de metadados é parte do protocolo e não uma tarefa meramente auxiliar. Erros em nomes de autores, instituições, países e referências podem alterar redes, rankings e interpretações. Antes de calcular indicadores, a equipe deve examinar campos ausentes, duplicatas, inconsistências de DOI, variações de afiliação, grafias alternativas e registros incompletos. Em estudos com várias bases, é necessário definir regra de prioridade entre metadados conflitantes. O protocolo deve registrar se a fonte principal será preservada ou se será criado um registro consolidado a partir de múltiplas fontes.

Os arquivos brutos devem ser preservados separadamente das versões limpas; cada transformação precisa ter responsável, data, regra e justificativa. Dados sensíveis ou internos somente devem ser enviados a ferramentas de IA quando houver base legal, autorização institucional e proteção adequada. Nos demais casos, recomenda-se agregação, anonimização ou uso de descrições estruturais.

A.20.4 - Parâmetros, redes e rotulagem de clusters

Os parâmetros de rede devem ser explicitados com o mesmo cuidado dedicado à estratégia de busca. Medidas de centralidade, limiares de ocorrência, normalização, resolução de *clusters* e métodos de layout não são escolhas neutras. Um limiar alto pode ocultar atores periféricos e temas emergentes; um limiar baixo pode gerar mapas densos e pouco interpretáveis. A escolha deve ser justificada em relação à pergunta. Quando possível, o protocolo deve prever análise de sensibilidade, apresentando resultados com diferentes limiares e explicando o que se mantém estável ou muda.

Em mapas temáticos, a nomeação de *clusters* precisa seguir uma sequência verificável. Primeiro, selecionar termos e documentos representativos. Segundo, gerar rótulos candidatos com ou sem IA. Terceiro, comparar os rótulos com títulos, resumos e palavras-chave. Quarto, submeter os rótulos a especialistas. Quinto, documentar casos ambíguos. Essa sequência evita que rótulos elegantes, mas imprecisos, sejam incorporados ao relatório. Um *cluster* não é uma entidade natural; é resultado de escolhas de dados, método e interpretação.

A redação de resultados deve preservar a diferença entre achado estatístico e afirmação substantiva. Dizer que uma instituição ocupa posição central na rede de coautoria é diferente de afirmar que ela coordena a área. Dizer que um termo cresceu em frequência é diferente de afirmar que o tema é promissor. Dizer que publicações internacionais recebem mais citações é diferente de afirmar que internacionalização causa impacto. O protocolo deve exigir que cada conclusão seja acompanhada de evidência, escopo e limite.

Sempre que o resultado depender de limiar, normalização, resolução de agrupamento ou método de contagem, o protocolo deve prever teste de sensibilidade. A estabilidade deve ser descrita por aquilo que permanece, muda ou desaparece quando os parâmetros são alterados.

A.20.5 - Ciclo de revisão, critérios de parada e auditabilidade

A inteligência artificial pode ser usada como interlocutor crítico. Nesse papel, ela não produz a conclusão final, mas questiona coerência, lacunas, vieses e alternativas interpretativas. Um *prompt* de interlocução crítica deve pedir que o modelo identifique afirmações fortes demais, dependência de uma única métrica, ausência de validação, problemas de fonte e falta de contextualização. Essa função é útil porque obriga a equipe a revisar a narrativa antes da entrega. Ainda assim, a crítica gerada pela IA deve ser examinada por humanos.

A aplicação dos protocolos em equipes exige um ciclo de revisão. Uma versão inicial pode ser construída antes da coleta; uma segunda versão após testes de busca; uma terceira após limpeza; uma quarta após indicadores preliminares; e uma quinta após validação interpretativa. Esse ciclo permite que o protocolo seja estável sem ser rígido. Mudanças são aceitáveis quando justificadas e registradas. O problema não é ajustar o protocolo; o problema é ajustar silenciosamente o método para acomodar resultados desejados.

Protocolos também devem prever critérios de parada. Em buscas bibliográficas, sempre é possível incluir mais termos, bases ou documentos. Em análise de redes, sempre é possível testar novos limiares. Em *prompts*, sempre é possível solicitar mais interpretações. Sem critérios de parada, o estudo se torna interminável ou seletivo. O protocolo deve definir quando a busca será considerada suficiente, quando a limpeza será encerrada, quais análises serão principais e quais serão suplementares. Critérios de parada fortalecem reprodutibilidade.

A qualidade de um protocolo pode ser avaliada por sua auditabilidade. Um leitor externo consegue reconstruir a pergunta, entender a busca, identificar o *corpus*, verificar a limpeza, reproduzir os indicadores principais, compreender os *prompts* usados, examinar a validação e reconhecer limitações? Se a resposta for negativa, o protocolo precisa ser ampliado. A auditabilidade não exige abertura irrestrita de todos os dados em todos os contextos, mas exige que decisões essenciais sejam transparentes para quem tem legitimidade de revisão.

O protocolo deve prever uma seção de incerteza. Essa seção não é sinal de fraqueza, mas de rigor. Ela deve indicar incerteza de cobertura, incerteza de identificação, incerteza de classificação, incerteza de interpretação e incerteza prospectiva. Em estudos com IA, deve incluir incerteza de modelo: variação de respostas, risco de alucinação, dependência de versão, influência do *prompt* e dificuldade de reprodução exata. Explicitar incertezas ajuda o leitor a usar os resultados de forma proporcional.

Uma prática recomendável é manter uma matriz de evidências. Nela, cada conclusão relevante é associada aos dados que a sustentam, aos indicadores usados, à visualização correspondente, ao *prompt* eventualmente empregado, à validação realizada e ao nível de confiança. Essa matriz permite revisar rapidamente se uma conclusão está sustentada ou se

depende de inferência frágil. Em relatórios executivos, a matriz pode permanecer no anexo, mas deve orientar a redação principal.

Outro recurso útil é a matriz de decisões humanas. Ela registra pontos em que a equipe escolheu entre alternativas: usar uma base ou outra, incluir *preprints* ou não, adotar contagem inteira ou fracionária, manter ou excluir um termo ambíguo, aceitar ou modificar um rótulo de *cluster*. A IA pode ter sugerido opções, mas a decisão final deve ter responsável humano. Essa matriz protege a integridade do estudo e facilita revisão posterior.

Quadro 20.7 - Ciclo mínimo de versões do protocolo

Versão	Conteúdo de controle
Versão 1 – Planejamento	Pergunta, finalidade, fontes e responsabilidades.
Versão 2 – Busca piloto	Termos testados, ruídos, documentos canônicos e ajustes.
Versão 3 – *Corpus* limpo	Deduplicação, dicionários e regras de normalização.
Versão 4 – Análise preliminar	Indicadores, parâmetros, visualizações e sensibilidade.
Versão 5 – Interpretação validada	Triangulação, divergências, limitações e decisões humanas.
Versão 6 – Produto final	Adequação ao público, revisão ética e ata de encerramento metodológico.

Fonte: Elaborado pelo autor (2026).

A temporalidade deve ser incorporada aos protocolos. Bases de dados mudam, registros são corrigidos, citações aumentam, periódicos entram ou saem de indexação, autores mudam de afiliação e termos científicos se transformam. Um estudo cientométrico é sempre fotografia de um momento metodologicamente definido. O protocolo deve registrar a data de coleta e, em análises recorrentes, estabelecer periodicidade de atualização. Quando resultados forem usados em *dashboards*, deve haver controle de versão e histórico de mudanças.

O uso de IA deve ser declarado de forma específica. Não basta dizer que a IA foi utilizada. É preciso indicar se ela apoiou expansão de termos, classificação, sumarização, rotulagem, redação, revisão de limitações ou tradução de linguagem. Também é necessário indicar que as saídas foram revisadas por humanos. Essa declaração protege a responsabilidade autoral e permite que leitores avaliem a extensão da mediação algorítmica.

Por fim, protocolos completos devem ser compreendidos como objetos vivos. Eles podem ser aprimorados após críticas, replicações, mudanças de base e desenvolvimento de novas ferramentas. A boa prática não é congelar procedimentos, mas manter rastreabilidade. Em uma área tão dinâmica quanto a cientometria com IA, a combinação de flexibilidade e documentação é essencial. O protocolo é o instrumento que permite inovar sem perder controle metodológico.

Como reforço final, recomenda-se que cada protocolo seja encerrado com uma página de síntese contendo a pergunta, o *corpus*, os indicadores principais, os *prompts* usados, a validação realizada, as limitações e o produto final. Essa síntese facilita a leitura por avaliadores e reduz a possibilidade de que decisões metodológicas fiquem dispersas ao longo do documento.

A.20.6 - Cautelas específicas por protocolo

Na análise de autores e grupos, a dimensão ética é especialmente relevante. Resultados podem ser usados em avaliação de desempenho, progressão, contratação ou distribuição de recursos. Por isso, o protocolo deve impedir comparações injustas entre áreas, estágios de carreira e regimes de autoria. Deve também evitar exposição individual indevida quando o objetivo é diagnóstico coletivo. Indicadores de carreira precisam ser contextualizados por tempo, campo, trajetória, licença, colaboração e oportunidades. A IA não deve gerar julgamentos sobre valor de pesquisadores.

Na análise institucional, a comparação com pares deve ser cuidadosamente documentada. A seleção de instituições comparáveis pode alterar conclusões tanto quanto a escolha de indicadores. Critérios possíveis incluem missão, porte, região, área de especialização, orçamento, idade institucional e perfil de pós-graduação. Comparar uma instituição regional com universidades globais intensivas em pesquisa pode produzir diagnóstico injusto. O protocolo deve explicar por que os pares foram escolhidos e que tipo de comparação é legítima.

Na detecção de temas emergentes, é útil combinar sinais quantitativos e qualitativos. Sinais quantitativos incluem crescimento de documentos, citações recentes, aumento de termos, novos *clusters* e entrada de atores. Sinais qualitativos incluem editoriais, chamadas de conferências, programas de financiamento, debates regulatórios e mudanças tecnológicas. A IA pode ajudar a organizar esses sinais em categorias, mas não deve converter sinais fracos em previsões. O produto final deve comunicar graus de evidência e incerteza.

Na colaboração internacional, a contagem fracionária pode ser necessária para evitar que documentos com muitos países distorçam indicadores. A contagem inteira valoriza a presença de todos os países em cada documento; a fracionária distribui o crédito. Ambas podem ser justificadas, mas respondem a perguntas diferentes. O protocolo deve indicar qual método foi usado e, em estudos sensíveis, apresentar ambos. A IA pode explicar a diferença a gestores, mas a decisão metodológica deve ser registrada.

Em relatórios de política científica, recomendações devem ser condicionais. Em vez de afirmar que uma área deve ser priorizada, o relatório pode dizer: se o objetivo for aumentar capacidade nacional em determinado tema, então os dados indicam tais lacunas e oportunidades. Se o objetivo for ampliar impacto internacional, então os dados sugerem outro conjunto de ações. Essa redação evita que indicadores pareçam determinar decisões. A política envolve valores, prioridades e negociação; a cientometria oferece evidências para qualificar esse processo.

A.20.7 - Modelos de anexos obrigatórios

Um anexo de estratégia de busca deve conter mais do que a *string* final. Deve apresentar o raciocínio que levou à construção dos blocos conceituais, os termos testados, os termos descartados, os campos pesquisados, a data da busca, a base consultada, o número de registros por versão e observações sobre ruído. Essa prática é importante porque estratégias de busca em cientometria costumam ser ajustadas iterativamente. Sem o histórico, o leitor vê apenas a versão final e não compreende por que alternativas foram abandonadas. A IA pode ajudar a reconstruir esse histórico a partir de notas de trabalho, mas não deve inventar justificativas retrospectivas. O registro deve refletir decisões reais.

Um anexo de limpeza de dados deve conter dicionários de normalização. Para autores, pode registrar variantes de nome, identificadores, afiliações e decisões sobre homônimos. Para instituições, pode registrar nomes oficiais, siglas, traduções, unidades agregadas e

identificadores ROR. Para palavras-chave, pode registrar termos preferenciais, sinônimos, acrônimos, grafias alternativas e termos mantidos separados por ambiguidade. Esse anexo torna a análise mais transparente e permite que especialistas contestem escolhas terminológicas. Em estudos com IA, recomenda-se marcar quais agrupamentos foram sugeridos pelo modelo e quais foram confirmados por humanos.

Um anexo de *prompts* deve organizar as interações por etapa do estudo. *Prompts* de busca, limpeza, classificação, interpretação e redação não devem ficar misturados. Para cada *prompt*, é útil registrar objetivo, texto, dados fornecidos, saída esperada, modelo usado, data, resposta aproveitada e revisão humana. Quando a saída da IA não foi usada, isso também pode ser registrado, pois revela tentativa metodológica descartada. O anexo de *prompts* contribui para a transparência e ajuda a distinguir autoria humana de apoio algorítmico.

Um anexo de parâmetros de visualização deve indicar escolhas que afetam mapas e gráficos. Em redes, isso inclui tipo de relação, método de contagem, limiar, normalização, algoritmo de layout, resolução de comunidades e tratamento de componentes. Em séries temporais, inclui janela, agregação, suavização e normalização. Em mapas temáticos, inclui critérios de centralidade e densidade. Visualizações cientométricas podem parecer autoexplicativas, mas são produtos de decisões técnicas. O protocolo deve permitir que essas decisões sejam examinadas.

Um anexo de validação deve apresentar evidências de controle de qualidade. Pode incluir número de registros revisados manualmente, critérios de amostragem, taxa de erro encontrada, exemplos de correções, documentos canônicos verificados, rótulos de *cluster* revisados, divergências entre avaliadores e decisões finais. Esse anexo não precisa ser excessivamente longo, mas deve demonstrar que a equipe não aceitou automaticamente resultados de softwares ou modelos de IA. A validação é especialmente importante quando o estudo será usado em avaliação institucional ou política científica.

Um modelo de governança do estudo deve definir instâncias de aprovação. A etapa de pergunta pode ser aprovada pelo coordenador científico; a estratégia de busca por especialista em informação; a limpeza por analista de dados e especialista de domínio; a interpretação por grupo multidisciplinar; e o relatório final por equipe responsável pela decisão. Essa divisão reduz o risco de que uma única perspectiva domine o estudo. Em contextos institucionais, também protege contra conflitos de interesse e uso seletivo de indicadores.

O protocolo deve prever tratamento de divergências. Especialistas podem discordar sobre rótulos, inclusão de termos, relevância de documentos ou interpretação de tendências. Em vez de esconder divergências, o protocolo pode registrá-las e indicar decisão final. Algumas divergências podem ser resolvidas por leitura adicional; outras refletem disputas legítimas do campo. A IA pode ajudar a sintetizar posições, mas não deve arbitrar controvérsias científicas. Registrar divergências aumenta a honestidade analítica.

Em estudos recorrentes, como observatórios de pesquisa, recomenda-se criar um protocolo de atualização. Ele deve indicar periodicidade, critérios de inclusão de novos registros, tratamento de mudanças de base, revisão de termos emergentes, recalculo de indicadores e comparação com versões anteriores. Atualizar não significa apenas adicionar documentos; pode exigir refazer limpeza, revisar *clusters* e reinterpretar séries. A IA pode apoiar comparação entre versões, identificando mudanças, mas a equipe deve validar se tais mudanças são substantivas ou artefatos.

Quadro 20.8 - Conjunto mínimo de anexos

Anexo	Conteúdo mínimo
Estratégia de busca	Blocos conceituais, versões da *string*, campos, filtros, datas, resultados e justificativas.
Limpeza e normalização	Variantes, identificadores, regras, exclusões, ambiguidades e dicionários.
Prompts e IA	Objetivo, texto, modelo, data, dados fornecidos, saída, revisão e decisão humana.
Parâmetros e visualizações	Relação, contagem, limiar, normalização, algoritmo, resolução, layout e janelas.
Validação	Amostra, critérios, erros, correções, divergências e decisões finais.
Governança e atualização	Responsáveis, aprovações, acesso, periodicidade, histórico e tratamento de mudanças.

Fonte: Elaborado pelo autor (2026).

A.20.8 - Aplicação prática, comunicação e melhoria contínua

Para relatórios executivos, é útil adotar uma estrutura de uma página por decisão. Cada página deve conter pergunta de decisão, evidência principal, gráfico ou mapa, interpretação prudente, riscos, opções e limitações. Essa estrutura impede que gestores recebam apenas listas de indicadores. A IA pode ajudar a transformar resultados técnicos em linguagem acessível, mas deve manter a ligação entre evidência e conclusão. Uma síntese executiva responsável comunica também o que os dados não permitem decidir.

Em artigos científicos, o protocolo pode ser convertido em seção metodológica. A redação deve informar desenho do estudo, fontes, busca, critérios de inclusão, extração, limpeza, análise, visualização, uso de IA, validação e limitações. Muitos artigos bibliométricos são criticados por relatar ferramentas, mas não decisões. Dizer que se usou um software não explica parâmetros, *corpus* ou interpretação. Um bom protocolo fornece os elementos necessários para uma seção de método robusta e revisável por pares.

No caso de programas de pós-graduação, o protocolo deve ser sensível à missão formativa. Indicadores de produção e citação não capturam orientação, formação de recursos humanos, impacto regional, inovação pedagógica ou inserção social. Uma análise cientométrica pode apoiar o diagnóstico, mas deve ser combinada com dados de formação, egressos, projetos, redes e contribuição social. A IA pode auxiliar a organizar evidências, mas não deve produzir avaliação total do programa com base apenas em registros bibliográficos.

Em prospecção tecnológica, o protocolo deve integrar publicações e patentes com cuidado. Publicações indicam produção científica; patentes indicam proteção de invenções em contextos jurídicos e econômicos específicos. A relação entre ciência e tecnologia não é linear. Um tema com muitas publicações pode não gerar patentes; uma tecnologia patenteada pode depender de conhecimento tácito ou industrial não publicado. O protocolo deve explicitar a lógica de integração entre literatura científica, patentes, projetos e mercado.

Na organização do conhecimento e curadoria de acervos, protocolos cientométricos podem apoiar construção de vocabulários, identificação de núcleos temáticos, seleção de coleções e desenvolvimento de serviços de informação. Nesse caso, o foco não é avaliar desempenho, mas organizar acesso ao conhecimento. Indicadores de coocorrência, cocitação e acoplamento podem revelar relações úteis para taxonomias e mapas de navegação. A IA pode sugerir categorias, mas bibliotecários e especialistas devem validar coerência semântica e adequação ao público.

A comunicação das limitações deve ser proporcional ao uso. Em um artigo científico, limitações podem ser detalhadas em seção específica. Em um briefing executivo, devem aparecer junto aos achados principais, para evitar leitura seletiva. Em *dashboards*, limitações podem aparecer como notas, metadados e alertas. A IA pode ajudar a redigir limitações em linguagem clara, mas não deve suavizar riscos para tornar o relatório mais persuasivo. A utilidade de um estudo depende da honestidade com que comunica incertezas.

A avaliação posterior do protocolo é uma etapa frequentemente esquecida. Após a entrega do produto, a equipe deve registrar o que funcionou, o que precisou ser ajustado, quais etapas consumiram mais tempo, quais dados apresentaram maior problema, quais *prompts* foram úteis e quais produziram ruído. Esse aprendizado deve alimentar versões futuras. A cientometria com IA evolui rapidamente; a melhoria contínua dos protocolos é uma forma de preservar rigor em ambiente de mudança tecnológica.

Um princípio final é a proporcionalidade metodológica. Nem todo estudo exige todas as técnicas disponíveis. Uma pergunta simples pode ser respondida com indicadores descritivos e revisão cuidadosa; uma pergunta complexa pode exigir redes, séries temporais, classificação textual e validação por especialistas. O protocolo deve evitar tanto simplificação excessiva quanto sofisticação desnecessária. A IA pode sugerir muitas análises, mas cabe à equipe selecionar apenas aquelas que contribuem para responder à pergunta de forma rigorosa.

Protocolos completos, quando bem aplicados, transformam a cientometria com IA em prática acumulativa. Estudos futuros podem reaproveitar *strings*, dicionários, *prompts*, scripts e critérios, ajustando-os a novos contextos. Essa acumulação metodológica reduz retrabalho e aumenta comparabilidade. Ao mesmo tempo, cada aplicação deve ser sensível ao domínio, ao idioma, à escala e ao uso social dos resultados. A padronização é ponto de partida, não substituto de julgamento.

A dimensão pedagógica dos protocolos merece destaque. Ao explicitar cada etapa, eles ensinam novos analistas a pensar cientometricamente. O estudante aprende que uma pergunta de pesquisa precisa de unidade de análise; que busca bibliográfica é construção metodológica; que indicadores têm pressupostos; que mapas são representações; que IA exige validação; e que comunicação científica deve ser responsável. Assim, protocolos são também instrumentos de formação em cultura analítica.

A adoção institucional de protocolos pode criar padrões mínimos de qualidade. Universidades, agências e laboratórios podem estabelecer que todo relatório cientométrico inclua anexos de busca, limpeza, indicadores, *prompts* e limitações. Isso não impede criatividade metodológica; apenas garante transparência básica. Em longo prazo, padrões desse tipo reduzem disputas sobre números e deslocam a discussão para interpretação, missão e decisão.

A implementação dos protocolos em ambientes reais requer negociação entre rigor e viabilidade. Equipes de pesquisa trabalham com prazos, acesso desigual a bases, limitações de pessoal, restrições de licença e expectativas de gestores. O protocolo não deve ignorar essas condições; deve torná-las visíveis. Quando uma base comercial não estiver disponível, pode-se usar fonte aberta, desde que a limitação seja declarada. Quando não houver tempo para validação extensa, pode-se realizar validação amostral, desde que a margem de incerteza seja comunicada. Quando a IA for usada para acelerar tarefas, deve-se registrar onde a aceleração ocorreu e quais controles compensatórios foram aplicados. A transparência sobre restrições aumenta a confiança no estudo porque mostra que a equipe não confunde ideal metodológico com execução concreta.

Outro aspecto prático é a construção de memória institucional. Muitas organizações produzem relatórios cientométricos isolados, sem preservar dados, scripts, *prompts* ou decisões. A cada novo estudo, a equipe recomeça do zero. Protocolos completos reduzem essa perda de conhecimento. Eles permitem criar repositórios internos de estratégias de busca, dicionários de normalização, modelos de relatório, exemplos de visualização e *prompts* validados. Com isso, estudos futuros podem ser mais rápidos sem serem menos rigorosos. A IA pode ajudar a recuperar e adaptar materiais anteriores, mas o acervo metodológico deve permanecer sob curadoria humana.

A adoção de protocolos também favorece diálogo entre áreas. Pesquisadores de diferentes campos frequentemente discordam sobre o que conta como produção relevante, qual base é adequada, que indicadores são aceitáveis e como interpretar citações. O protocolo oferece uma linguagem comum para explicitar essas diferenças. Em vez de discutir apenas resultados finais, a equipe pode discutir unidades de análise, cobertura, normalização, métricas e validação. Esse deslocamento melhora a qualidade da deliberação e reduz conflitos baseados em expectativas implícitas.

Em ambientes de política científica, a comunicação dos protocolos deve ser adaptada sem perder densidade. Gestores não precisam ler todos os detalhes técnicos, mas devem ter acesso a anexos que sustentem os achados. A síntese executiva deve declarar, em poucas linhas, que bases foram usadas, que período foi coberto, que limitações são relevantes e que papel a IA desempenhou. Essa prática evita que relatórios sejam consumidos apenas como narrativas persuasivas. Um bom briefing apresenta evidências e também ensina como interpretá-las.

Por fim, protocolos completos contribuem para uma ética da moderação. A disponibilidade de IA generativa tende a estimular produção abundante de textos, mapas e interpretações. No entanto, a credibilidade da cientometria não depende da quantidade de produtos gerados, mas da qualidade das perguntas, da adequação dos dados, da clareza dos métodos e da prudência das conclusões. A função do protocolo é moderar a velocidade tecnológica com responsabilidade científica. Ele obriga a perguntar: que problema estamos respondendo, com quais dados, mediante quais escolhas, com que validação e com quais limites? Essas perguntas preservam o caráter científico da análise em um contexto de automação crescente.

Uma recomendação operacional adicional é realizar uma reunião de encerramento metodológico antes da publicação ou entrega do produto. Nessa reunião, a equipe deve revisar a pergunta, confirmar se o *corpus* responde ao escopo, examinar se os indicadores escolhidos são proporcionais ao problema, verificar se as visualizações não induzem interpretações exageradas, conferir se os *prompts* foram documentados e validar se as limitações estão visíveis. Esse procedimento simples reduz erros de comunicação e impede que achados preliminares sejam convertidos em conclusões fortes demais. A reunião também deve registrar pendências e decisões finais, formando uma ata metodológica que acompanha o relatório. Em estudos recorrentes, essa ata se torna fonte de aprendizado para versões futuras.

Também é útil distinguir produtos de exploração, produtos de análise e produtos de decisão. Produtos de exploração servem para orientar perguntas e testar *corpus*; não devem circular como diagnóstico final. Produtos de análise apresentam evidências estabilizadas, mas ainda exigem interpretação. Produtos de decisão traduzem evidências em cenários ou opções, assumindo maior responsabilidade comunicacional. A IA pode gerar textos em todos esses níveis, mas o protocolo deve deixar claro em qual estágio cada saída se encontra. Essa distinção evita que rascunhos analíticos sejam tratados como pareceres conclusivos e preserva a integridade do processo cientométrico.

Finalmente, a sustentabilidade dos protocolos depende de atualização. Novas bases, métricas, modelos de linguagem e diretrizes de avaliação surgem continuamente. Portanto, cada protocolo deve conter data de revisão prevista e responsável pela atualização. A cientometria com IA não pode depender de modelos estáticos em um ambiente técnico mutável. O compromisso com atualização, documentação e validação é o que permite transformar inovação tecnológica em prática científica robusta.

A adoção desses procedimentos transforma a relação entre IA e cientometria em uma prática de responsabilidade compartilhada. A máquina pode sugerir, organizar, comparar e redigir, mas a equipe humana define o problema, seleciona as evidências, interpreta os resultados e responde por suas consequências. Essa separação de funções deve ser preservada em todos os protocolos. Quando ela desaparece, a IA deixa de ser instrumento e passa a ocupar indevidamente o lugar de autoridade analítica. Quando ela é explicitada, a IA amplia a capacidade de investigação sem enfraquecer os fundamentos científicos da análise.

Essa é a base operacional para uma cientometria com IA transparente, crítica, documentada, validada e metodologicamente sustentável.

A página-síntese de cada protocolo deve registrar pergunta, *corpus*, indicadores, *prompts*, validação, limitações, decisões humanas, versão e data prevista de atualização. Essa síntese funciona como porta de entrada para avaliação, reutilização e memória institucional.

Capítulo 21

Biblioteca de prompts cientométricos

A consolidação da inteligência artificial generativa como instrumento de apoio à pesquisa científica transformou a maneira como pesquisadores formulam perguntas, exploram literatura, constroem estratégias de busca, organizam metadados, interpretam indicadores e redigem produtos analíticos. Na cientometria, essa transformação é particularmente relevante porque a área lida com grandes volumes de registros bibliográficos, citações, palavras-chave, redes de colaboração, mapas temáticos e séries temporais. A possibilidade de usar modelos de linguagem para interagir com esses objetos não elimina a necessidade de métodos quantitativos, mas cria uma camada adicional de mediação: a camada dos *prompts*. Um *prompt* não é apenas uma pergunta feita a uma máquina; é uma formulação metodológica que define papel, contexto, dados, restrições, critérios de interpretação e forma esperada de saída (BENDER *et al.*, 2021; JI *et al.*, 2023; OUYANG *et al.*, 2022).

Ao longo dos capítulos anteriores, foram discutidos fundamentos de cientometria, aplicação de IA, engenharia de *prompts*, formulação de perguntas, estratégias de busca, coleta e limpeza de dados, indicadores, redes, mapas temáticos, análise temporal, sumarização de literatura, visualização, interpretação e escrita de relatórios. Este capítulo funciona como síntese prática dessa trajetória. Seu propósito é reunir uma biblioteca de *prompts* que possa ser reutilizada, adaptada e documentada em estudos reais. A biblioteca cobre desde o momento inicial, em que o pesquisador ainda possui apenas um tema amplo, até a etapa final de auditoria do relatório e do protocolo.

A cientometria depende de decisões sucessivas: qual base consultar, que campos exportar, quais termos incluir, que indicadores calcular, como tratar duplicatas, como nomear *clusters*, como interpretar crescimento de publicações e como comunicar resultados a diferentes públicos. Em cada uma dessas decisões, a IA pode oferecer alternativas, organizar justificativas, sugerir verificações e produzir textos preliminares. Contudo, a IA também pode fabricar referências, confundir conceitos, exagerar conclusões, transformar associação em causalidade, omitir limitações ou apresentar resultados plausíveis sem fundamento empírico. Por isso, uma biblioteca de *prompts* cientométricos precisa incorporar mecanismos de controle: pedir grau de incerteza, exigir distinção entre dado e inferência, obrigar indicação de evidências necessárias, solicitar riscos e documentar decisões.

A noção de *prompt* adotada neste capítulo é compatível com a literatura recente sobre aprendizagem baseada em *prompts* e engenharia de *prompts*, segundo a qual o comportamento de modelos de linguagem pode ser orientado por instruções, exemplos, formatos de saída e restrições explícitas (BROWN *et al.*, 2020; LIU *et al.*, 2023; WEI *et al.*, 2022; WANG *et al.*, 2023; YAO *et al.*, 2023). Entretanto, no contexto cientométrico, a eficácia do *prompt* não deve ser avaliada apenas pela fluência da resposta. Ela deve ser avaliada pela aderência aos dados, pela verificabilidade das afirmações, pela utilidade metodológica, pela capacidade de explicitar limitações e pela compatibilidade com princípios de avaliação responsável da pesquisa (HICKS *et al.*, 2015).

A biblioteca apresentada a seguir foi organizada em dez grupos: *prompts* para formular perguntas, busca bibliográfica, limpeza de dados, indicadores, redes, mapas temáticos, interpretação, redação, revisão crítica e auditoria/reprodutibilidade. Cada grupo contém *prompts* que podem ser utilizados de modo independente ou em cadeias de *prompts*. Em um estudo completo, por exemplo, o pesquisador pode iniciar com um *prompt* para transformar tema

amplo em pergunta, seguir com outro para criar estratégia de busca, usar *prompts* de limpeza para padronizar palavras-chave, aplicar *prompts* de indicadores para selecionar métricas, solicitar interpretação prudente de redes e, ao final, usar *prompts* de auditoria para revisar limitações. Essa lógica transforma a interação com a IA em fluxo metodológico documentado.

Princípios de uso da biblioteca:

1. Nenhum *prompt* deve ser usado como substituto de decisão metodológica. O *prompt* deve ajudar a formular alternativas e explicitar consequências, mas a escolha final permanece sob responsabilidade do pesquisador.
2. Os dados fornecidos ao modelo devem ser identificáveis quanto à origem, data, base, recorte temporal, estratégia de busca, campos utilizados e procedimentos de limpeza. Respostas obtidas sem rastreabilidade não devem sustentar conclusões substantivas.
3. Os *prompts* devem ser salvos em versão integral, junto com a data de uso, o modelo utilizado, os parâmetros relevantes, os dados de entrada e a saída recebida. Essa documentação transforma a conversa com a IA em parte auditável do protocolo.
4. O pesquisador deve distinguir *prompts* exploratórios, analíticos, redacionais e críticos. *Prompts* exploratórios servem para abrir possibilidades; analíticos organizam evidências; redacionais transformam resultados em texto; críticos identificam riscos, vieses e fragilidades.
5. Toda saída da IA que envolva autores, obras, números, bases de dados, indicadores ou citações deve ser conferida em fonte externa. A fluência da resposta não é evidência de validade.
6. A IA deve ser solicitada a registrar incerteza, limitações, pressupostos e possíveis vieses. Uma boa resposta cientométrica não é apenas afirmativa; é também prudente.
7. *Prompts* que geram classificações, rótulos de *clusters* ou agrupamentos terminológicos devem produzir grau de confiança e lista de casos ambíguos para revisão humana.
8. *Prompts* usados para redação devem ser instruídos a não extrapolar dados, não afirmar causalidade sem desenho causal e não converter indicadores em juízos absolutos de qualidade científica.
9. Respostas da IA devem ser tratadas como insumos intermediários. O produto final deve resultar de triangulação entre métricas, literatura substantiva, conhecimento do domínio e avaliação crítica.
10. A biblioteca deve ser adaptada ao contexto institucional, disciplinar e linguístico do estudo. O mesmo *prompt* pode ter desempenho diferente em bases, idiomas, áreas e modelos distintos.

21.1 Prompts para formular perguntas

A formulação da pergunta é a etapa que define o alcance de todo estudo cientométrico. Perguntas vagas tendem a produzir estratégias de busca dispersas, indicadores mal escolhidos, interpretações excessivas e relatórios pouco úteis. A IA pode apoiar essa etapa ao decompor temas amplos em perguntas operacionalizáveis, sugerir unidades de análise, explicitar variáveis, propor recortes temporais e apontar limites. O pesquisador, entretanto, deve impedir

que a IA substitua a relação entre pergunta, teoria, método e dado. Em cientometria, uma pergunta adequada precisa indicar o fenômeno investigado, a unidade de análise, a fonte de dados, o período, o tipo de evidência desejado e o produto esperado. Perguntar "como está a pesquisa em inteligência artificial?" é insuficiente. Perguntar "como evoluiu a produção científica sobre explicabilidade em IA em artigos indexados na Scopus entre 2015 e 2025, considerando crescimento anual, países líderes, periódicos, redes de colaboração e temas emergentes?" já orienta decisões metodológicas concretas.

Prompt 21.1.1 Transformação de tema amplo em problema cientométrico

Objetivo: Converter um tema inicial em problema cientométrico delimitado, com unidade de análise, escopo e possíveis indicadores.

Contexto de uso: Fase inicial de planejamento, antes da construção da estratégia de busca.

Atue como especialista em cientometria.

A partir do tema [TEMA], formule cinco problemas de pesquisa cientométrica.

Para cada problema, indique: pergunta principal, unidade de análise, base de dados recomendada, período temporal plausível, indicadores possíveis, visualizações úteis, limitações metodológicas e decisão humana necessária antes da coleta.

Não invente dados empíricos; apresente apenas possibilidades metodológicas.

Saída esperada: Tabela textual com cinco problemas, cada um acompanhado de justificativa e limitações.

Validação recomendada: Conferir se cada pergunta pode ser respondida por metadados bibliográficos disponíveis e se não exige dados inexistentes na base selecionada.

Risco metodológico: A IA pode sugerir escopos amplos demais, bases inadequadas ou indicadores que não respondem diretamente à pergunta.

Prompt 21.1.2 Decomposição de objetivo geral em perguntas específicas

Objetivo: Transformar um objetivo geral em conjunto de perguntas secundárias alinhadas a dimensões cientométricas.

Contexto de uso: Construção de projeto de pesquisa, dissertação, tese, artigo ou relatório institucional.

Receba o objetivo geral abaixo e decomponha-o em perguntas cientométricas específicas.

Organize as perguntas nas dimensões produção, impacto, colaboração, estrutura temática, evolução temporal, internacionalização e aplicações de política científica.

Para cada pergunta, indique que tipo de dado é necessário e que evidência seria insuficiente para respondê-la. Objetivo: [OBJETIVO].

Para cada pergunta, indique que tipo de dado é necessário e que evidência seria insuficiente para respondê-la. Objetivo: [OBJETIVO].

Saída esperada: Conjunto de perguntas secundárias com dimensões analíticas e requisitos de dados.

Validação recomendada: Verificar se as perguntas não repetem a mesma dimensão com nomes diferentes e se há equilíbrio entre descrição e interpretação.

Risco metodológico: Pode ocorrer proliferação de perguntas sem prioridade analítica, tornando o estudo inviável.

Prompt 21.1.3 Escolha de pergunta principal

Objetivo: Selecionar a pergunta mais robusta entre alternativas, considerando viabilidade, relevância e validade metodológica.

Contexto de uso: Definição final do desenho de estudo.

Analise as perguntas cientométricas listadas a seguir e classifique-as segundo relevância científica, viabilidade de coleta, clareza da unidade de análise, adequação dos indicadores e risco de interpretação indevida. Escolha uma pergunta principal e justifique.

Perguntas: [LISTA].

Saída esperada: Ranking argumentado das perguntas e recomendação da pergunta principal.

Validação recomendada: Comparar a escolha da IA com interesses do projeto, disponibilidade de bases e capacidade de análise da equipe.

Risco metodológico: A IA pode privilegiar perguntas mais gerais ou mais elegantes, mas menos relevantes para o contexto institucional.

Prompt 21.1.4 Delimitação de escopo temporal e documental

Objetivo: Definir período, tipos documentais e critérios de inclusão coerentes com a pergunta.

Contexto de uso: Planejamento da busca e definição de critérios de elegibilidade.

Para a pergunta cientométrica [PERGUNTA], proponha três opções de recorte temporal e documental.

Para cada opção, explique vantagens, riscos, possíveis vieses e impacto esperado sobre a interpretação dos resultados.

Considere artigos, revisões, anais de congresso, livros, capítulos e *preprints* quando forem pertinentes à área.

Saída esperada: Três desenhos possíveis de escopo, com comparação de vantagens e limitações.

Validação recomendada: Testar recortes em busca piloto e avaliar volume de registros, ruído e cobertura da área.

Risco metodológico: A IA pode sugerir períodos convencionais sem considerar a história específica do tema.

Prompt 21.1.5 Formulação de pergunta para gestores

Objetivo: Adaptar uma pergunta acadêmica para finalidade de tomada de decisão institucional.

Contexto de uso: Relatórios para pró-reitorias, agências, programas de pós-graduação e unidades acadêmicas.

Reformule a pergunta cientométrica [PERGUNTA] para uso em gestão de pesquisa.

A nova pergunta deve produzir evidências úteis para decisão institucional, sem reduzir a análise a ranking.

Indique indicadores, visualizações, riscos de uso indevido e tipo de recomendação que poderia ser produzido.

Saída esperada: Pergunta orientada à decisão, acompanhada de indicadores e cautelas.

Validação recomendada: Validar com gestores se a pergunta responde a uma necessidade real e se evita incentivos perversos.

Risco metodológico: Pode transformar análise científica em instrumento simplificado de competição institucional.

Prompt 21.1.6 Pergunta para identificação de lacunas

Objetivo: Construir pergunta voltada a identificar lacunas temáticas e oportunidades de pesquisa.

Contexto de uso: Planejamento de agendas de pesquisa e prospecção científica.

A partir do tema [TEMA] e do objetivo [OBJETIVO], formule perguntas cientométricas voltadas a lacunas de pesquisa.

Diferencie lacunas de baixa produção, lacunas de baixa colaboração, lacunas metodológicas, lacunas geográficas e lacunas de tradução entre áreas.

Indique quais evidências bibliométricas seriam necessárias para sustentar cada tipo de lacuna.

Saída esperada: Lista de perguntas sobre lacunas, separadas por tipo e evidência necessária.

Validação recomendada: Conferir se a lacuna não resulta apenas de estratégia de busca estreita ou base de dados limitada.

Risco metodológico: A IA pode confundir baixa visibilidade em uma base com ausência real de pesquisa.

Prompt 21.1.7 Pergunta para comparação entre áreas

Objetivo: Formular pergunta comparativa entre campos, temas ou subáreas.

Contexto de uso: Estudos interdisciplinares, avaliação institucional e análise de portfólios científicos.

Ajude-me a formular uma pergunta cientométrica comparativa entre [ÁREA A] e [ÁREA B].

A pergunta deve evitar comparações injustas entre campos com culturas de publicação diferentes.

Indique normalizações necessárias, indicadores apropriados, bases recomendadas e limitações de comparabilidade.

Saída esperada: Pergunta comparativa com cautelas de normalização e comparabilidade.

Validação recomendada: Consultar especialistas de ambas as áreas para verificar diferenças de publicação, idioma, indexação e citação.

Risco metodológico: Comparações podem naturalizar hierarquias entre áreas sem considerar culturas epistêmicas distintas.

Prompt 21.1.8 Auditoria da pergunta de pesquisa

Objetivo: Revisar criticamente uma pergunta antes da coleta de dados.

Contexto de uso: Revisão de projeto, qualificação de pesquisa e planejamento de relatório.

Audite a pergunta cientométrica abaixo.

Identifique ambiguidades, pressupostos ocultos, termos indefinidos, riscos de viés, unidades de análise ausentes, indicadores inadequados e conclusões que a pergunta não permitiria sustentar. Em seguida, proponha uma versão revisada.

Pergunta: [PERGUNTA].

Saída esperada: Diagnóstico crítico e versão revisada da pergunta.

Validação recomendada: Checar se a versão revisada continua alinhada ao objetivo do projeto e aos recursos disponíveis.

Risco metodológico: A IA pode tornar a pergunta tecnicamente mais sofisticada, mas menos compreensível para o público-alvo.

21.2 Prompts para busca bibliográfica

A busca bibliográfica é a principal porta de entrada para o *corpus* cientométrico. Ela define o universo empírico a partir do qual serão calculados indicadores, construídas redes, identificados temas e formuladas interpretações. Com apoio de IA, é possível expandir vocabulários, traduzir termos entre idiomas e áreas, identificar sinônimos, construir *strings* booleanas e revisar lacunas. Porém, o modelo de linguagem não deve ser tratado como autoridade terminológica. Ele pode sugerir termos inexistentes, descritores imprecisos, autores irrelevantes ou sintaxes incompatíveis com bases específicas. Por isso, todo *prompt* de busca deve exigir separação entre termos principais, variantes, sinônimos, siglas, termos correlatos e termos duvidosos. Deve também solicitar adaptação por base, uma vez que Scopus, Web of Science, PubMed, Dimensions, OpenAlex e repositórios institucionais operam com campos e sintaxes distintas (MOHER *et al.*, 2009; PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021; MONGEON; PAUL-HUS, 2016; MARTÍN-MARTÍN *et al.*, 2021).

Prompt 21.2.1 Expansão de vocabulário de busca

Objetivo: Ampliar repertório terminológico sem perder rastreabilidade.

Contexto de uso: Antes da elaboração da *string* booleana.

Atue como especialista em estratégias de busca para estudos cientométricos.

A partir do tema [TEMA], gere uma lista de termos principais, sinônimos, variantes ortográficas, siglas, expressões relacionadas, termos em inglês, termos em português e termos emergentes. Separe termos consolidados de termos duvidosos.

Para cada termo, explique por que deve ou não entrar na *string* de busca.

Saída esperada: Vocabulário estruturado com justificativa de inclusão ou exclusão.

Validação recomendada: Conferir termos em tesauros, artigos sentinela, descritores de área e buscas piloto.

Risco metodológico: A IA pode produzir traduções literais que não são usadas pela comunidade científica.

Prompt 21.2.2 Construção de string booleana

Objetivo: Criar *string* inicial compatível com bases bibliográficas.

Contexto de uso: Construção de busca em Scopus, Web of Science, PubMed, Dimensions, OpenAlex ou bases institucionais.

Com base no vocabulário abaixo, construa uma estratégia booleana para [BASE].

Use campos apropriados, operadores, truncamentos e aspas conforme a sintaxe da base. Separe blocos conceituais e explique a função de cada bloco.

Vocabulário: [VOCABULÁRIO].

Saída esperada: *String* booleana segmentada em blocos, com explicação da lógica de busca.

Validação recomendada: Executar busca piloto na base real e ajustar erros de sintaxe, excesso de ruído ou baixa recuperação.

Risco metodológico: A IA pode aplicar sintaxe de uma base em outra ou usar operadores não suportados.

Prompt 21.2.3 Busca sensível versus específica

Objetivo: Comparar versões amplas e restritivas da estratégia.

Contexto de uso: Ajuste fino da estratégia antes da coleta definitiva.

Crie duas versões da estratégia de busca para [TEMA]: uma sensível, com maior recuperação, e outra específica, com maior precisão.

Para cada versão, indique quais termos aumentam ruído, quais reduzem lacunas, que tipos de documentos provavelmente serão recuperados e que testes devem ser realizados.

Saída esperada: Duas *strings* com análise de sensibilidade, especificidade, ruído e lacunas.

Validação recomendada: Comparar amostras dos primeiros resultados e calcular proporção de registros relevantes em teste piloto.

Risco metodológico: Busca sensível pode gerar *corpus* inviável; busca específica pode excluir frentes emergentes.

Prompt 21.2.4 Revisão de ruído e lacunas

Objetivo: Diagnosticar problemas em resultados preliminares de busca.

Contexto de uso: Iteração entre busca piloto e busca final.

Analise a amostra de títulos e resumos recuperados pela estratégia [STRING].

Identifique fontes de ruído, temas relevantes ausentes, termos ambíguos, possíveis falsos positivos e possíveis falsos negativos.

Proponha ajustes na *string*, justificando cada alteração.

Saída esperada: Lista de ruídos, lacunas e ajustes recomendados.

Validação recomendada: Reexecutar a busca e comparar variação de volume e relevância por amostragem.

Risco metodológico: A IA pode inferir lacunas sem conhecimento completo da área ou da base.

Prompt 21.2.5 Adaptação por base de dados

Objetivo: Converter uma estratégia geral para sintaxes específicas.

Contexto de uso: Estudos comparativos ou multimodais de coleta bibliográfica.

Adapte a estratégia de busca abaixo para Scopus, Web of Science, PubMed, Dimensions e OpenAlex.

Para cada base, indique campos equivalentes, operadores compatíveis, limitações de sintaxe e cuidados na exportação.

Estratégia geral: [STRING].

Saída esperada: Versões específicas da estratégia por base, com observações técnicas.

Validação recomendada: Testar cada *string* diretamente na respectiva base e registrar data e número de resultados.

Risco metodológico: A IA pode desconhecer mudanças recentes na sintaxe ou nas interfaces das bases.

Prompt 21.2.6 Tradução terminológica entre áreas

Objetivo: Traduzir conceitos entre campos disciplinares sem perder sentido.

Contexto de uso: Estudos interdisciplinares e mapeamentos conceituais.

O conceito [CONCEITO] é usado em [ÁREA A]. Identifique termos equivalentes, próximos ou contrastantes usados em [ÁREA B]. Para cada termo, indique se é sinônimo, analogia, tradução parcial, termo guarda-chuva ou falso equivalente. Sugira como representar essa relação na estratégia de busca.

Saída esperada: Mapa terminológico entre áreas, com alerta para falsos equivalentes.

Validação recomendada: Validar com literatura de ambas as áreas e especialistas de domínio.

Risco metodológico: Pode haver aproximações indevidas entre tradições teóricas diferentes.

Prompt 21.2.7 Documentação da busca

Objetivo: Produzir registro reprodutível da estratégia de busca.

Contexto de uso: Redação de metodologia em artigo, dissertação ou relatório.

Transforme as informações abaixo em uma seção metodológica de busca bibliográfica. Inclua base, data, *string*, campos pesquisados, filtros, tipos documentais, idioma, período, número de resultados, critérios de inclusão/exclusão e limitações.

Dados: [DADOS DA BUSCA].

Saída esperada: Texto metodológico claro e auditável.

Validação recomendada: Conferir todos os números, datas e filtros com registros da busca real.

Risco metodológico: A IA pode completar lacunas com suposições se os dados estiverem incompletos.

Prompt 21.2.8 Detecção de termos inventados pela IA

Objetivo: Identificar termos, descritores ou autores suspeitos em estratégia gerada por IA.

Contexto de uso: Controle de qualidade da estratégia antes da coleta.

Revise a lista abaixo e marque termos que parecem não consolidados, ambíguos, inventados, traduzidos literalmente ou inadequados ao campo.

Para cada termo suspeito, explique o risco e sugira uma forma de verificação externa.

Lista: [TERMOS].

Saída esperada: Lista auditada de termos confiáveis, duvidosos e inadequados.

Validação recomendada: Verificar em bases bibliográficas, tesauros, artigos de revisão e literatura seminal.

Risco metodológico: A IA pode não reconhecer que um termo raro é legítimo em subárea específica.

21.3 Prompts para limpeza de dados

A limpeza de dados é uma etapa decisiva porque pequenas inconsistências podem alterar redes, indicadores e mapas. Um mesmo autor pode aparecer com iniciais diferentes; uma instituição pode ser registrada em múltiplas línguas; palavras-chave podem variar no singular, plural, sigla ou tradução; e duplicatas podem inflar a produção de um tema. A IA pode auxiliar a identificar padrões, propor agrupamentos e gerar regras de normalização, mas nunca deve eliminar registros automaticamente sem documentação. Em cientometria, a limpeza é ato interpretativo: decidir que dois termos são equivalentes ou que duas afiliações pertencem à mesma instituição implica conhecimento de contexto. Por isso, *prompts* de limpeza devem solicitar grau de confiança, justificativa, casos ambíguos e lista de decisões que requerem revisão humana (ARIA; CUCCURULLO, 2017; COBO *et al.*, 2012).

Prompt 21.3.1 Normalização de palavras-chave

Objetivo: Agrupar variantes terminológicas preservando ambiguidade.

Contexto de uso: Antes de análise de coocorrência, mapas temáticos e nuvens de termos.

Receba a lista de palavras-chave abaixo e agrupe termos equivalentes ou muito próximos.

Para cada grupo, indique o termo preferencial, os termos agrupados, a justificativa semântica e o grau de confiança.

Não elimine termos ambíguos; marque-os para revisão humana.

Lista: [PALAVRAS-CHAVE].

Saída esperada: Tabela textual com grupos, termo preferencial, justificativa e grau de confiança.

Validação recomendada: Revisar grupos de baixa confiança e comparar com tesauros ou literatura da área.

Risco metodológico: A IA pode agrupar termos teoricamente diferentes por semelhança lexical superficial.

Prompt 21.3.2 Padronização de nomes de autores

Objetivo: Identificar possíveis variantes de nomes autorais.

Contexto de uso: Análise de produtividade, autoria, colaboração e trajetórias.

Analise a lista de autores abaixo e sugira possíveis agrupamentos de variantes do mesmo nome.

Use indícios como sobrenome, iniciais, afiliação, país, coautores e tema.

Classifique cada agrupamento como alto, médio ou baixo grau de confiança.

Não fundir casos incertos: Lista: [AUTORES].

Saída esperada: Agrupamentos potenciais de autores com confiança e justificativas.

Validação recomendada: Confirmar com ORCID, páginas institucionais, perfis bibliográficos e amostras de publicações.

Risco metodológico: Homônimos podem ser fundidos indevidamente; variantes podem permanecer separadas.

Prompt 21.3.3 Desambiguação de instituições

Objetivo: Padronizar afiliações institucionais em diferentes línguas e formas.

Contexto de uso: Análise institucional, colaboração, internacionalização e geografia da ciência.

Receba a lista de afiliações abaixo e proponha uma padronização institucional.

Para cada variação, indique nome padronizado, país, possíveis unidades subordinadas, justificativa e grau de confiança.

Marque universidades, hospitais, laboratórios, empresas e órgãos públicos separadamente quando necessário.

Lista: [AFILIAÇÕES].

Saída esperada: Lista de instituições padronizadas e variações associadas.

Validação recomendada: Verificar nomes oficiais, identificadores ROR, sites institucionais e padrões nacionais.

Risco metodológico: Instituições diferentes podem ter nomes semelhantes; unidades podem ser confundidas com matriz.

Prompt 21.3.4 Tratamento de duplicatas

Objetivo: Identificar registros potencialmente duplicados antes da análise.

Contexto de uso: Combinação de bases bibliográficas e limpeza de *corpus*.

Analise os registros abaixo e identifique possíveis duplicatas.

Considere DOI, título, autores, ano, periódico, volume, páginas e resumo. Classifique como duplicata provável, duplicata possível ou registro distinto.

Explique o critério usado e indique o registro que deve ser preservado.

Registros: [REGISTROS].

Saída esperada: Diagnóstico de duplicatas com critérios de decisão.

Validação recomendada: Aplicar regras automatizadas e revisão manual em casos sem DOI ou com metadados incompletos.

Risco metodológico: Títulos semelhantes de versões *preprint* e artigo final podem ser tratados incorretamente.

Prompt 21.3.5 Criação de dicionário de normalização

Objetivo: Transformar decisões de limpeza em dicionário reprodutível.

Contexto de uso: Reprodutibilidade e auditoria de dados.

Com base nas decisões de padronização abaixo, construa um dicionário de normalização com colunas: valor original, valor padronizado, tipo de campo, regra aplicada, justificativa, grau de confiança e responsável pela validação.

Decisões: [DECISÕES].

Saída esperada. Dicionário reprodutível para limpeza de autores, instituições ou termos.

Validação recomendada: Salvar o dicionário junto aos dados brutos e ao script de limpeza.

Risco metodológico: A IA pode criar regras genéricas sem preservar exceções importantes.

Prompt 21.3.6 Amostragem de qualidade

Objetivo: Planejar verificação humana de dados limpos.

Contexto de uso: Controle de qualidade antes da análise final.

Proponha um plano de amostragem para verificar a qualidade da limpeza de dados do *corpus* [DESCRIÇÃO].

Inclua tamanho de amostra, critérios de estratificação, campos a verificar, taxa de erro aceitável, procedimento para resolver divergências e forma de registrar correções.

Saída esperada: Plano de auditoria por amostragem.

Validação recomendada: Aplicar amostra real e calcular taxa de erro por campo.

Risco metodológico: A amostragem pode subestimar erros raros, mas de alto impacto, como homônimos autorais.

Prompt 21.3.7 Detecção de campos ausentes

Objetivo: Identificar lacunas nos metadados necessários ao estudo.

Contexto de uso: Preparação para análise de redes, indicadores e mapas temáticos.

Analise a descrição do *corpus* abaixo e identifique campos ausentes ou incompletos que podem comprometer o estudo.

Relacione cada campo ausente ao tipo de análise afetada, ao risco interpretativo e a possíveis soluções.

Corpus: [DESCRIÇÃO DOS CAMPOS].

Saída esperada: Mapa de lacunas de metadados e impactos analíticos.

Validação recomendada: Checar porcentagem real de completude por campo.

Risco metodológico: A IA pode presumir disponibilidade de campos que não existem na base exportada.

Prompt 21.3.8 Relato metodológico da limpeza

Objetivo: Redigir seção transparente de pré-processamento.

Contexto de uso: Redação de artigo, relatório ou apêndice metodológico.

Escreva uma seção metodológica sobre o pré-processamento do *corpus*.

Use os dados abaixo e descreva: remoção de duplicatas, padronização de autores, desambiguação institucional, normalização de palavras-chave, tratamento de campos ausentes, amostragem de qualidade e limitações.

Dados: [PROCESSO DE LIMPEZA].

Saída esperada: Texto metodológico com procedimentos e limitações.

Validação recomendada: Verificar se todos os procedimentos descritos foram de fato executados.

Risco metodológico: A IA pode redigir uma metodologia mais completa do que a prática realizada.

21.4 Prompts para indicadores

Indicadores cientométricos sintetizam propriedades da produção científica, mas não são equivalentes diretos de qualidade. Número de publicações, citações, índice h, colaboração internacional, centralidade em redes e atenção social medem aspectos diferentes e carregam vieses disciplinares, linguísticos, temporais e de base. A IA pode auxiliar na escolha e explicação de indicadores, na comparação entre métricas e na redação de resultados, desde que seja instruída a não extrapolar. *Prompts* para indicadores devem exigir compatibilidade entre pergunta, unidade de análise e métrica; devem pedir limitações; e devem impedir que métricas sejam convertidas em juízos absolutos sobre mérito científico (GARFIELD, 1972; HIRSCH, 2005; MOED, 2008; BORNMANN; MUTZ, 2015; HICKS *et al.*, 2015; PIWOWAR, 2013).

Prompt 21.4.1 Seleção de indicadores

Objetivo: Escolher indicadores coerentes com a pergunta de pesquisa.

Contexto de uso: Planejamento analítico e desenho do relatório.

Para a pergunta cientométrica [PERGUNTA], selecione indicadores de produção, impacto, colaboração, interdisciplinaridade, internacionalização e atenção social.

Para cada indicador, explique: o que mede, o que não mede, unidade de análise, dado necessário, limitação e visualização recomendada.

Saída esperada: Matriz de indicadores com escopo, dado e limitação.

Validação recomendada: Conferir se todos os indicadores podem ser calculados com os metadados disponíveis.

Risco metodológico: A IA pode sugerir métricas incompatíveis com a base ou com o nível de análise.

Prompt 21.4.2 Limitações do índice h

Objetivo: Explicar limitações de indicador amplamente usado.

Contexto de uso: Relatórios institucionais, pareceres e comunicação com gestores.

Explique as limitações do índice h para o público [PÚBLICO].

A explicação deve abordar dependência de tempo de carreira, área, base de dados, autoria coletiva, autocitação, idioma, idade das publicações e impossibilidade de medir qualidade intrínseca.

Use linguagem formal, mas acessível.

Saída esperada: Texto explicativo adaptado ao público selecionado.

Validação recomendada: Comparar com literatura sobre métricas responsáveis e exemplos de diferentes áreas.

Risco metodológico: A IA pode simplificar o indicador ou reforçar usos inadequados.

Prompt 21.4.3 Interpretação de crescimento de publicações

Objetivo: Interpretar aumento de produção sem confundir crescimento com qualidade.

Contexto de uso: Análise temporal de produção científica.

Interprete a série temporal de publicações abaixo.

Separe: descrição objetiva, possíveis explicações, hipóteses que exigem evidência adicional, artefatos de busca possíveis e afirmações que não podem ser feitas apenas com esses dados.

Série: [DADOS].

Saída esperada: Interpretação em níveis, com cautelas metodológicas.

Validação recomendada: Verificar mudanças na indexação, expansão da base, variações terminológicas e eventos externos.

Risco metodológico: Crescimento pode ser interpretado como relevância sem considerar ruído ou mudanças de cobertura.

Comparação de métricas de impacto

Objetivo: Comparar citações brutas, normalizadas e indicadores de periódicos.

Contexto de uso: Escolha de indicadores para avaliação e relatórios.

Compare as métricas [MÉTRICAS] para responder à pergunta [PERGUNTA].

Explique quando cada métrica é adequada, quais vieses carrega, quais normalizações são necessárias e que combinações podem produzir interpretação mais equilibrada.

Saída esperada: Comparação argumentada das métricas de impacto.

Validação recomendada: Conferir fórmulas, fontes e nível de agregação de cada métrica.

Risco metodológico: Métricas de periódico podem ser aplicadas indevidamente a artigos ou autores.

Prompt 21.4.4 Indicadores de colaboração

Objetivo: Selecionar métricas para colaboração científica.

Contexto de uso: Estudos de redes, internacionalização e política científica.

Para o *corpus* [DESCRIÇÃO], proponha indicadores de colaboração. Inclua coautoria média, colaboração internacional, colaboração interinstitucional, densidade de rede, centralidade, componentes e comunidades. Para cada indicador, indique interpretação e limitação.

Saída esperada: Plano de indicadores relacionais e colaborativos.

Validação recomendada: Checar qualidade dos dados de autoria e afiliação.

Risco metodológico: Coautoria pode ser confundida com colaboração substantiva sem evidência adicional.

Prompt 21.4.5 Indicadores alternativos

Objetivo: Avaliar usos e limites de *altmetrics*.

Contexto de uso: Relatórios de comunicação científica, impacto social e circulação pública.

Explique se indicadores alternativos de atenção social são pertinentes para [TEMA]. Diferencie atenção, uso, circulação, discussão pública e impacto científico. Indique fontes possíveis, limitações, riscos de manipulação e formas de triangulação com citações e conteúdo.

Saída esperada: Análise de pertinência de *altmetrics*.

Validação recomendada: Conferir disponibilidade real de dados altmétricos e variação por área.

Risco metodológico: Atenção social pode ser interpretada indevidamente como qualidade ou impacto científico.

Prompt 21.4.6 Resultado narrativo a partir de tabela

Objetivo: Transformar tabela de indicadores em texto sem extrapolação.

Contexto de uso: Redação de resultados de artigo ou relatório.

A partir da tabela abaixo, escreva uma seção de resultados. Descreva padrões, contrastes e mudanças observáveis. Não atribua causalidade. Não use termos avaliativos como “melhor” ou “superior” sem critério explícito. Aponte limitações da tabela. Tabela: [TABELA].

Saída esperada: Texto de resultados descritivo e prudente.

Validação recomendada. Conferir números, percentuais e interpretação com a tabela original.

Risco metodológico: A IA pode criar explicações causais não sustentadas pelos dados.

Prompt 21.4.7 Auditoria de indicador inadequado

Objetivo: Identificar uso indevido de métrica em conclusão.

Contexto de uso: Revisão crítica de relatórios, pareceres e artigos.

Avalie se a conclusão abaixo usa indicadores de forma adequada. Identifique extrapolações, confusão entre quantidade e qualidade, ausência de normalização, comparações injustas, vieses de base e afirmações sem suporte. Conclusão: [TEXTO]. Indicadores usados: [INDICADORES].

Saída esperada: Parecer crítico sobre adequação de métricas.

Validação recomendada: Comparar com princípios de métricas responsáveis e literatura da área.

Risco metodológico: A IA pode ser excessivamente conservadora ou permissiva dependendo da formulação.

21.5 Prompts para redes

Redes científicas representam relações entre autores, documentos, instituições, países, referências ou termos. Elas permitem examinar colaboração, cocitação, acoplamento bibliográfico, coocorrência de palavras-chave e circulação internacional. A IA pode apoiar a leitura de propriedades estruturais e a redação de interpretações, mas é especialmente arriscada nesse domínio porque tende a transformar proximidade visual em relação substantiva e associação em causalidade. *Prompts* para redes devem exigir separação entre descrição estrutural, possível explicação, hipótese interpretativa, evidência necessária e risco de interpretação indevida. Uma comunidade detectada por modularidade, por exemplo, não é automaticamente uma escola teórica; um nó central não é necessariamente líder intelectual; e uma ligação de coautoria não prova colaboração efetiva (KESSLER, 1963; SMALL, 1973; CALLON; COURTIAL; LAVILLE, 1991; WALTMAN; VAN ECK; NOYONS, 2010).

Prompt 21.5.1 Interpretação prudente de rede

Objetivo: Interpretar rede sem inventar causalidade.

Contexto de uso: Análise de redes de coautoria, cocitação, acoplamento e palavras-chave.

Interprete a rede descrita abaixo.

Separe a resposta em: descrição estrutural, possíveis explicações, hipóteses interpretativas, evidências necessárias para confirmação e riscos de interpretação indevida.

Não afirme causalidade quando houver apenas associação.

Descrição: [DESCRIÇÃO DA REDE].

Saída esperada: Interpretação organizada por níveis de certeza.

Validação recomendada: Comparar com métricas, visualização, literatura substantiva e conhecimento do campo.

Risco metodológico: A IA pode inferir relações sociais ou intelectuais não demonstradas pela rede.

Prompt 21.5.2 Rede de coautoria

Objetivo: Analisar colaboração autoral com cautela.

Contexto de uso: Estudos de colaboração e formação de grupos de pesquisa.

Analise a rede de coautoria abaixo.

Identifique componentes, autores centrais, grupos, pontes e possíveis padrões de colaboração.

Explique o que a rede permite afirmar e o que não permite afirmar sobre colaboração científica.

Dados: [MÉTRICAS E DESCRIÇÃO].

Saída esperada: Análise de estrutura colaborativa e limites interpretativos.

Validação recomendada: Verificar desambiguação de autores e regras de contagem fracionária ou integral.

Risco metodológico: Homônimos, autoria honorária e grandes consórcios podem distorcer a rede.

Prompt 21.5.3 Rede de cocitação

Objetivo: Interpretar estrutura intelectual a partir de referências cocitadas.

Contexto de uso: Mapeamento da estrutura intelectual de uma área.

A partir da rede de cocitação descrita, identifique possíveis tradições intelectuais, referências estruturantes e comunidades.

Para cada comunidade, proponha rótulo provisório, evidências que sustentam o rótulo e leituras necessárias para validação.

Rede: [DESCRIÇÃO].

Saída esperada: Leitura preliminar de comunidades de cocitação.

Validação recomendada. Ler referências centrais antes de confirmar rótulos de escola ou tradição.

Risco metodológico: A IA pode nomear comunidades com base em títulos sem compreender debates teóricos.

Prompt 21.5.4 Acoplamento bibliográfico

Objetivo: Interpretar proximidade entre documentos recentes.

Contexto de uso: Estudos de temas recentes e literatura ainda pouco citada.

Analise a rede de acoplamento bibliográfico abaixo.

Explique quais frentes de pesquisa podem estar representadas, quais documentos parecem conectar subtemas e quais evidências adicionais seriam necessárias para confirmar frentes emergentes.

Dados: [REDE].

Saída esperada: Análise de frentes de pesquisa com cautela temporal.

Validação recomendada: Conferir referências compartilhadas e conteúdo dos documentos centrais.

Risco metodológico: Acoplamento pode refletir referências comuns genéricas, não proximidade substantiva.

Prompt 21.5.5 Coocorrência de palavras-chave

Objetivo: Interpretar mapa de termos e *clusters* conceituais.

Contexto de uso: Mapas temáticos e análise conceitual.

Interprete a rede de coocorrência de palavras-chave abaixo.

Para cada *cluster*, proponha rótulo, termos centrais, subtemas, possíveis ambiguidades e necessidade de validação.

Não trate frequência de termo como importância teórica automática.

Rede: [DADOS].

Saída esperada: Leitura temática preliminar com rótulos e ambiguidades.

Validação recomendada: Validar normalização de palavras-chave e leitura de resumos representativos.

Risco metodológico: Termos genéricos ou mal normalizados podem gerar *clusters* artificiais.

Prompt 21.5.6 Centralidade e modularidade

Objetivo: Explicar métricas de rede para interpretação cientométrica.

Contexto de uso: Capacitação de equipe e redação metodológica.

Explique as métricas de centralidade, densidade, modularidade, grau, intermediação e proximidade no contexto da rede [TIPO DE REDE].

Para cada métrica, indique interpretação possível, limitação e exemplo de mau uso.

Saída esperada: Glossário interpretativo de métricas de rede.

Validação recomendada: Checar se as métricas são adequadas ao tipo de rede e ao modo de normalização.

Risco metodológico: Métricas de rede podem ser usadas como rankings sem interpretação contextual.

Prompt 21.5.7 Comparação entre redes

Objetivo: Comparar redes de períodos, áreas ou instituições diferentes.

Contexto de uso: Estudos comparativos temporais, institucionais ou internacionais.

Compare as duas redes descritas abaixo.

Identifique diferenças de tamanho, densidade, fragmentação, centralização, comunidades e conectividade.

Indique quais diferenças podem decorrer de desenho de coleta, cobertura de base ou normalização.

Rede A: [A]. Rede B: [B].

Saída esperada: Comparação estruturada de redes com cautelas.

Validação recomendada. Normalizar por tamanho, período, base e critérios de inclusão antes de concluir.

Risco metodológico: Diferenças estruturais podem resultar de artefatos de coleta, não de fenômenos sociais.

Prompt 21.5.8 Auditoria causal de redes

Objetivo: Impedir interpretações causais indevidas.

Contexto de uso: Revisão de discussão e conclusão de estudos relacionais.

Revise a interpretação de rede abaixo e identifique frases que sugerem causalidade sem evidência.

Para cada frase, proponha reescrita prudente em termos de associação, proximidade, correlação, hipótese ou necessidade de evidência adicional.

Texto: [INTERPRETAÇÃO].

Saída esperada: Lista de extrapolações causais e reescritas prudentes.

Validação recomendada: Verificar se as conclusões têm suporte em desenho longitudinal, documental ou qualitativo.

Risco metodológico: A IA pode remover causalidade legítima se houver evidência externa não fornecida.

21.6 Prompts para mapas temáticos

Mapas temáticos organizam a estrutura cognitiva de um campo a partir de termos, documentos, referências ou tópicos. Eles permitem identificar temas motores, básicos, emergentes, declinantes, frentes de pesquisa e comunidades epistêmicas. A IA pode auxiliar a rotular *clusters*, comparar alternativas de nomenclatura, construir narrativas da estrutura intelectual e indicar leituras necessárias para validação. Entretanto, a rotulagem automática é uma das tarefas mais sensíveis: nomes de *clusters* podem parecer convincentes mesmo quando não correspondem ao conteúdo dos documentos. Um rótulo deve ser considerado hipótese interpretativa, não resultado automático. *Prompts* para mapas temáticos devem pedir justificativas baseadas em termos centrais, títulos, resumos e documentos representativos (CHEN, 2006; COBO *et al.*, 2012; VAN ECK; WALTMAN, 2010; WALTMAN; VAN ECK; NOYONS, 2010).

Prompt 21.6.1 Nomeação de clusters temáticos

Objetivo: Atribuir rótulos provisórios a *clusters* com justificativa.

Contexto de uso: Mapas de coocorrência, tópicos e comunidades temáticas.

Nomeie os *clusters* abaixo com base em termos centrais, documentos representativos e resumos.

Para cada *cluster*, forneça três opções de rótulo: descritiva, conceitual e aplicada. Indique justificativa, grau de confiança e evidências necessárias para validação.

Clusters: [DADOS].

Saída esperada: Rótulos alternativos por *cluster*, com confiança e justificativas.

Validação recomendada: Ler documentos representativos e validar com especialista da área.

Risco metodológico: Rótulos podem ser excessivamente amplos, sedutores ou desalinhados ao conteúdo real.

Prompt 21.6.2 - Comparação de rótulos alternativos

Objetivo: Escolher rótulo mais adequado para um *cluster*.

Contexto de uso: Refinamento de mapas temáticos e redação de resultados.

Compare os rótulos propostos para o *cluster* [*Cluster*].

Avalie precisão, abrangência, especificidade, neutralidade, aderência aos termos centrais e risco de interpretação.

Recomende o rótulo mais adequado e explique por que os demais são menos apropriados.

Saída esperada: Avaliação crítica dos rótulos e recomendação justificada.

Validação recomendada. Submeter rótulos a revisão por especialistas ou dupla codificação.

Risco metodológico: A IA pode preferir rótulos elegantes em vez de rótulos tecnicamente precisos.

Prompt 21.6.3 - Tema central e subtemas

Objetivo: Identificar estrutura interna de um *cluster*.

Contexto de uso: Interpretação detalhada de *clusters* temáticos.

A partir dos termos e documentos do *cluster* [*Cluster*], identifique tema central, subtemas, métodos recorrentes, objetos de estudo, áreas de aplicação e possíveis tensões conceituais.

Classifique cada inferência por grau de evidência.

Saída esperada: Descrição interna do *cluster* com graus de evidência.

Validação recomendada: Conferir nos resumos e artigos completos dos documentos centrais.

Risco metodológico: Subtemas podem ser inferidos de termos isolados sem suporte textual suficiente.

Prompt 21.6.4 Temas motores, básicos, emergentes e declinantes

Objetivo: Interpretar mapa estratégico ou matriz temática.

Contexto de uso: Análise estratégica de campos científicos.

Interprete a posição dos temas abaixo em uma matriz de centralidade e densidade.

Classifique-os como motores, básicos, emergentes, declinantes ou periféricos.

Explique o significado metodológico de cada classificação e ressalte limites da interpretação.

Dados: [MATRIZ].

Saída esperada: Classificação de temas e interpretação da matriz.

Validação recomendada: Confirmar métricas de centralidade, densidade e período analisado.

Risco metodológico: Tema declinante pode ser artefato de recorte temporal ou mudança terminológica.

Prompt 21.6.5 Narrativa da estrutura intelectual

Objetivo: Construir texto interpretativo sobre organização de uma área.

Contexto de uso: Discussão de mapas temáticos em artigo ou relatório.

Com base nos *clusters*, referências centrais e termos abaixo, escreva uma narrativa da estrutura intelectual da área.

Separe: núcleo conceitual, frentes recentes, tradições teóricas, aplicações, lacunas e hipóteses interpretativas.

Não apresente rótulos como fatos definitivos.

Dados: [DADOS].

Saída esperada: Narrativa interpretativa com cautelas e hipóteses.

Validação recomendada: Comparar com revisões de literatura e leitura de documentos centrais.

Risco metodológico: A narrativa pode impor coerência excessiva a um campo heterogêneo.

Prompt 21.6.6 Detecção de temas emergentes

Objetivo: Identificar temas recentes com evidência temporal.

Contexto de uso: Prospecção científica e análise temporal.

Analise a evolução dos *clusters* abaixo e identifique possíveis temas emergentes.

Para cada tema, indique evidência temporal, crescimento, documentos representativos, termos associados, riscos de artefato de busca e dados adicionais necessários.

Dados: [DADOS TEMPORAIS].

Saída esperada: Lista de temas emergentes com evidências e riscos.

Validação recomendada: Verificar se crescimento não decorre de mudança de indexação, terminologia ou base.

Risco metodológico: A IA pode confundir novidade terminológica com emergência científica substantiva.

Prompt 21.6.7 Validação de rótulos por especialistas

Objetivo: Planejar validação humana de nomes de *clusters*.

Contexto de uso: Validação de mapas temáticos em estudos robustos.

Crie um protocolo de validação de rótulos temáticos por especialistas.

Inclua critérios de seleção de especialistas, formulário de avaliação, escala de adequação, procedimento de consenso, tratamento de divergências e registro das decisões finais.

Saída esperada: Protocolo de validação especializado.

Validação recomendada: Aplicar piloto com pequeno número de *clusters* antes da validação completa.

Risco metodológico: Especialistas podem divergir por escola teórica, área ou tradição metodológica.

Prompt 21.6.8 Auditoria de coerência temática

Objetivo: Revisar se *clusters* e rótulos são consistentes.

Contexto de uso: Refinamento final de mapas temáticos.

Audite os *clusters* e rótulos abaixo. Identifique rótulos genéricos demais, termos incompatíveis, *clusters* heterogêneos, sobreposição entre grupos e necessidade de divisão ou fusão.

Sugira ajustes e indique grau de confiança.

Dados: [*Clusters* E RÓTULOS].

Saída esperada: Parecer de coerência temática e ajustes recomendados.

Validação recomendada: Recalcular mapas ou revisar parâmetros se houver inconsistência estrutural.

Risco metodológic: A IA pode propor fusões conceituais que alteram indevidamente a estrutura empírica.

21.7 Prompts para interpretação

A interpretação cientométrica é a passagem mais delicada entre resultado e argumento. Dados bibliométricos descrevem produção, citação, colaboração, circulação e estrutura relacional, mas não explicam sozinhos os processos científicos, sociais e institucionais que os produziram. A IA pode atuar como interlocutor crítico, organizando achados, propondo hipóteses, listando limitações e apontando extrapolações. O uso adequado depende de *prompts* que distingam descrição objetiva, interpretação provável e hipótese que exige evidência adicional. Esse tripé reduz o risco de que a IA produza narrativas fortes demais para dados fracos (HICKS *et al.*, 2015; BENDER *et al.*, 2021; JI *et al.*, 2023).

Prompt 21.7.1 Interpretação em três níveis

Objetivo: Separar fato observado, interpretação e hipótese.

Contexto de uso: Discussão de resultados cientométricos.

A partir dos resultados abaixo, escreva uma interpretação cientométrica em três níveis: descrição objetiva, interpretação provável e hipóteses que exigem evidência adicional.

Aponte também possíveis vieses de base de dados, recorte temporal e estratégia de busca.

Resultados: [RESULTADOS].

Saída esperada": Interpretação estratificada por grau de evidência.

Validação recomendada: Conferir se cada afirmação está vinculada a dado, literatura ou hipótese explícita.

Risco metodológico: A IA pode preencher lacunas explicativas com conjecturas plausíveis, mas não verificadas.

Prompt 21.7.2 Interpretação contextualizada de indicadores

Objetivo: Interpretar indicadores considerando campo e base.

Contexto de uso: Relatórios avaliativos e diagnósticos institucionais.

Interprete os indicadores abaixo considerando área, período, base de dados, idioma, tipo documental e cultura de citação.

Evite juízos absolutos de qualidade. Indique quais comparações são legítimas e quais seriam metodologicamente frágeis.

Indicadores: [DADOS].

Saída esperada: Interpretação contextual com limites comparativos.

Validação recomendada: Validar diferenças de cultura de publicação e citação da área.

Risco metodológico: Indicadores podem ser naturalizados como medidas neutras de desempenho.

Prompt 21.7.3 Formulação de hipóteses prudentes

Objetivo: Gerar hipóteses sem apresentar explicações como fatos.

Contexto de uso: Discussão, planejamento de estudos complementares e política científica.

Com base nos achados [ACHADOS], formule hipóteses explicativas prudentes.

Para cada hipótese, indique evidência que a sugere, evidência faltante, como testá-la e nível de confiança.

Não transforme hipótese em conclusão.

Saída esperada: Lista de hipóteses com evidências e testes possíveis.

Validação recomendada: Definir quais hipóteses serão investigadas por dados adicionais ou literatura substantiva.

Risco metodológico: A IA pode multiplicar hipóteses sem priorização analítica.

Prompt 21.7.4 Triangulação de evidências

Objetivo: Integrar métricas, literatura e conhecimento especializado.

Contexto de uso: Interpretação robusta e redação de discussão.

Construa uma matriz de triangulação para os achados abaixo.

Para cada achado, indique evidência bibliométrica, evidência textual, literatura substantiva, conhecimento especializado necessário, possíveis contradições e conclusão permitida.

Achados: [ACHADOS].

Saída esperada: Matriz de triangulação interpretativa.

Validação recomendada: Preencher com evidências reais e revisar contradições.

Risco metodológico: A IA pode sugerir literatura inexistente ou genérica se não receber fontes.

Prompt 21.7.5 Identificação de vieses interpretativos

Objetivo: Reconhecer vieses em resultados e narrativa.

Contexto de uso': Revisão de discussão, conclusão e sumário executivo.

Revise a interpretação abaixo e identifique vieses de confirmação, viés de disponibilidade, viés de base de dados, viés linguístico, viés disciplinar e viés institucional.

Para cada viés, proponha reescrita ou verificação.

Texto: [INTERPRETAÇÃO].

Saída esperada: Lista de vieses e sugestões de correção.

Validação recomendada: Aplicar revisão por pares internos ou especialistas externos.

Risco metodológico: A IA pode identificar viés de forma genérica sem evidência específica.

Prompt 21.7.6 - Interpretação para diferentes públicos

Objetivo: Adequar interpretação sem distorcer evidências.

Contexto de uso: Comunicação científica, relatórios e briefings.

Reescreva a interpretação abaixo para três públicos: pesquisadores especialistas, gestores de pesquisa e público não técnico.

Preserve a precisão metodológica, indique limitações essenciais e evite simplificações que alterem o sentido dos dados.

Texto: [TEXTO].

Saída esperada: Três versões da interpretação com linguagem adaptada.

Validação recomendada: Conferir se a versão não técnica não omite limitações relevantes.

Risco metodológico: Simplificação pode reduzir incerteza e produzir conclusões indevidas.

Prompt 21.7.7 Revisão de causalidade

Objetivo: Evitar extrapolação causal em análise cientométrica.

Contexto de uso: Auditoria de artigos, relatórios e pareceres.

Analise as conclusões abaixo e marque afirmações causais, correlacionais, descritivas e especulativas.

Reescreva conclusões causais não sustentadas como hipóteses ou associações.

Conclusões: [TEXTO].

Saída esperada: Classificação de afirmações e reescrita prudente.

Validação recomendada: Conferir desenho do estudo e evidência longitudinal ou qualitativa adicional.

Risco metodológico: A IA pode suavizar conclusões legítimas se a evidência não for fornecida no *prompt*.

Prompt 21.7.8 Construção da seção de limitações

Objetivo: Transformar riscos metodológicos em seção de limitações.

Contexto de uso: Redação final de artigo, relatório ou tese.

Escreva uma seção de limitações para o estudo descrito abaixo. Inclua limitações de base de dados, estratégia de busca, metadados, indicadores, normalização, visualização, interpretação por IA e generalização. Use tom científico e não defensivo. Estudo: [DESCRIÇÃO].

Saída esperada: Seção de limitações completa e transparente.

Validação recomendada: Conferir se as limitações correspondem a problemas reais do estudo.

Risco metodológico: A IA pode incluir limitações genéricas que não se aplicam ao caso.

21.8 Prompts para redação

A IA pode apoiar a redação de relatórios, artigos, pareceres, sumários executivos e materiais de comunicação. Esse apoio é útil quando o pesquisador já possui dados, tabelas, figuras, resultados e decisões metodológicas. O risco aparece quando a IA preenche lacunas com fórmulas genéricas, inventa procedimentos, suaviza limitações ou transforma resultados em linguagem avaliativa sem suporte. *Prompts* de redação devem fornecer material de entrada claro e instruções explícitas sobre tom, estrutura, público, restrições, citações e limites interpretativos. Em textos cientométricos, a redação deve preservar a diferença entre resultados, discussão, recomendação e decisão política (BENDER *et al.*, 2021; JI *et al.*, 2023)

Prompt 21.8.1 Redação de metodologia

Objetivo: Escrever metodologia a partir de dados reais do protocolo.

Contexto de uso: Artigos, relatórios técnicos e dissertações.

Redija a seção de metodologia de um estudo cientométrico usando apenas as informações abaixo.

Inclua desenho do estudo, bases, data de busca, *string*, critérios, tipos documentais, exportação, limpeza, indicadores, ferramentas, parâmetros, uso de IA e limitações.

Não invente procedimentos.

Informações: [PROTOCOLO].

Saída esperada: Seção metodológica completa e verificável.

Validação recomendada: Comparar o texto com logs, scripts, dados e decisões reais.

Risco metodológico: A IA pode completar lacunas metodológicas com práticas não realizadas.

Prompt 21.8.2 Resultados narrativos

Objetivo: Transformar dados em narrativa descritiva.

Contexto de uso: Redação de resultados quantitativos.

Escreva a seção de resultados com base na tabela e nas figuras abaixo.

Descreva padrões observáveis, diferenças relevantes e tendências, sem explicar causas.

Use linguagem objetiva e indique quando uma diferença exige cautela.

Dados: [TABELA/FIGURAS].

Saída esperada: Texto de resultados claro e sem causalidade indevida.

Validação recomendada: Checar todos os números contra as tabelas originais.

Risco metodológico: A IA pode inventar percentuais, tendências ou rankings se os dados estiverem incompletos.

Prompt 21.8.3 Discussão interpretativa

Objetivo: Escrever discussão com triangulação e limites.

Contexto de uso: Artigos e relatórios analíticos.

Com base nos resultados [RESULTADOS] e na literatura [LITERATURA], escreva uma seção de discussão.

Diferencie achados, interpretações, hipóteses, implicações e limitações.

Não apresente recomendações normativas sem evidência.

Cite apenas as referências fornecidas.

Saída esperada: Discussão cientométrica com suporte e cautela.

Validação recomendada: Conferir aderência às referências e ausência de citações inventadas.

Risco metodológico: A IA pode generalizar além do *corpus* ou citar literatura não fornecida.

Prompt 21.8.4 Sumário executivo

Objetivo: Sintetizar resultados para tomada de decisão.

Contexto de uso: Comunicação para gestores e formuladores de políticas.

Produza um sumário executivo de até [N] palavras para gestores.

Inclua objetivo, *corpus*, principais achados, implicações, recomendações prudentes, limitações e próximos passos.

Evite jargão técnico e não transforme indicadores em rankings de qualidade.

Resultados: [RESULTADOS].

Saída esperada: Sumário executivo objetivo e metodologicamente prudente.

Validação recomendada: Validar se as recomendações são proporcionais às evidências.

Risco metodológico: Pode simplificar demais limitações ou produzir recomendações genéricas.

Prompt 21.8.5 Legenda analítica de figura

Objetivo: Escrever legenda que explique gráfico sem exagero.

Contexto de uso: Visualizações em relatórios, artigos e *dashboards*.

Escreva uma legenda analítica para a figura [DESCRIÇÃO].

A legenda deve explicar o que é mostrado, unidade de análise, período, métrica, interpretação principal e limitação.

Não inclua conclusão que não esteja visível na figura.

Saída esperada: Legenda completa e interpretativa.

Validação recomendada: Comparar legenda com figura e dados originais.

Risco metodológico: A IA pode fazer a legenda parecer evidência independente.

Prompt 21.8.6 - Parecer técnico

Objetivo: Produzir parecer a partir de evidências cientométricas.

Contexto de uso: Avaliação institucional e apoio a políticas científicas.

Redija um parecer técnico com base nos achados abaixo.

Estruture em contexto, método, evidências, interpretação, riscos, recomendações e limites. Diferencie recomendação técnica de decisão política.

Achados: [ACHADOS].

Saída esperada: Parecer técnico com recomendações proporcionais.

Validação recomendada: Revisar por especialista metodológico e gestor demandante.

Risco metodológico: A IA pode transformar diagnóstico em prescrição sem considerar contexto político.

Prompt 21.8.7 Adaptação para público não técnico

Objetivo: Comunicar cientometria sem perda de rigor.

Contexto de uso: Divulgação institucional, boletins e apresentações.

Reescreva o texto abaixo para público não técnico.

Preserve os principais resultados, explique termos como citação, colaboração, *cluster* e indicador, mantenha limitações essenciais e evite linguagem promocional.

Texto: [TEXTO].

Saída esperada: Versão acessível e precisa.

Validação recomendada: Submeter a leitor não especialista e verificar compreensão.

Risco metodológico: Adaptação pode eliminar nuances importantes.

Prompt 21.8.8 Revisão de estilo científico

Objetivo: Aprimorar clareza e coerência sem alterar conteúdo.

Contexto de uso: Revisão final de relatórios e manuscritos.

Revise o texto abaixo para melhorar clareza, coesão, formalidade científica e precisão terminológica.

Não altere números, não acrescente referências, não invente resultados e não remova limitações. Destaque mudanças substantivas sugeridas separadamente.

Texto: [TEXTO].

Saída esperada: Texto revisado e lista de alterações substantivas.

Validação recomendada: Comparar versão revisada com original para evitar mudança de sentido.

Risco metodológico: A IA pode suavizar críticas ou fortalecer conclusões inadvertidamente.

21.9 Prompts para revisão crítica

A revisão crítica é uma das funções mais importantes da IA em cientometria quando usada como interlocutor, e não como autoridade. Um modelo pode ser instruído a identificar inconsistências, lacunas, fragilidades de argumento, indicadores inadequados, conclusões fortes demais, ausência de limitações, termos ambíguos e riscos de viés. Essa função é valiosa porque desloca a IA da posição de produtora de texto para a posição de examinadora de qualidade. Ainda assim, a revisão crítica automatizada precisa ser conferida por humanos, pois a IA pode priorizar problemas genéricos, ignorar especificidades disciplinares ou sugerir padrões normativos inadequados (BENDER *et al.*, 2021; JI *et al.*, 2023; HICKS *et al.*, 2015).

Prompt 21.9.1 Revisão metodológica geral

Objetivo: Auditar coerência entre pergunta, dados, método e conclusão.

Contexto de uso: Revisão interna antes de submissão ou entrega.

Revise criticamente o estudo descrito abaixo.

Avalie coerência entre pergunta, *corpus*, base de dados, estratégia de busca, indicadores, métodos de análise, visualizações, interpretação e conclusões.

Identifique fragilidades, lacunas e melhorias prioritárias.

Estudo: [DESCRIÇÃO].

Saída esperada: Parecer metodológico estruturado.

Validação recomendada: Conferir cada crítica com o protocolo real do estudo.

Risco metodológico: A IA pode aplicar critérios de revisão sistemática a estudo cientométrico exploratório de forma inadequada.

Prompt 21.9.2 Avaliação de conclusões fortes demais

Objetivo: Detectar extrapolações e linguagem excessiva.

Contexto de uso: Revisão de discussão, resumo e sumário executivo.

Avalie se as conclusões abaixo são fortes demais para os dados apresentados.

Classifique cada conclusão como sustentada, parcialmente sustentada, especulativa ou não sustentada.

Proponha reescrita prudente.

Dados: [DADOS].

Conclusões: [CONCLUSÕES].

Saída esperada: Classificação de conclusões e reescritas.

Validação recomendada: Revisar com equipe e manter apenas conclusões proporcionais.

Risco metodológico: A IA pode não reconhecer evidências externas não incluídas no *prompt*.

Prompt 21.9.3 Revisão de vieses de base

Objetivo: Identificar vieses associados à fonte de dados.

Contexto de uso: Seção de limitações e escolha de fonte de dados.

Analise o estudo abaixo e identifique vieses decorrentes da base de dados usada.

Considere idioma, cobertura regional, áreas sub-representadas, tipos documentais, periódicos indexados, citações ausentes e mudanças de cobertura ao longo do tempo.

Estudo: [DESCRIÇÃO].

Saída esperada: Lista de vieses de base e impactos interpretativos.

Validação recomendada: Conferir documentação da base e estudos comparativos de cobertura.

Risco metodológico: A IA pode apresentar vieses típicos sem verificar a base específica.

Prompt 21.9.4 Revisão de coerência visual

Objetivo: Avaliar se gráficos e mapas sustentam o argumento.

Contexto de uso: Revisão de relatório visual, artigo ou *dashboard*.

Revise a descrição das visualizações abaixo.

Para cada gráfico ou mapa, avalie legibilidade, adequação ao dado, coerência com a pergunta, risco de distorção visual e clareza da legenda.

Sugira melhorias.

Visualizações: [DESCRIÇÕES].

Saída esperada: Parecer sobre visualizações e recomendações de ajuste.

Validação recomendada: Inspecionar figuras reais e verificar dados subjacentes.

Risco metodológico: A IA não vê detalhes de escala ou cores se a figura não for fornecida adequadamente.

Prompt 21.9.5 Revisão de uso da IA

Objetivo: Auditar transparência e responsabilidade no uso de *prompts*.

Contexto de uso: Seção metodológica, apêndice de *prompts* e declaração de uso de IA.

Avalie a descrição abaixo sobre uso de IA no estudo.

Verifique se há transparência sobre modelo, versão, data, *prompts*, dados de entrada, finalidade, validação humana, limitações e responsabilidade autoral.

Sugira melhorias para relato ético e reprodutível.

Descrição: [TEXTO].

Saída esperada: Parecer sobre transparência no uso de IA.

Validação recomendada: Comparar com políticas da instituição, periódico ou agência.

Risco metodológico: A IA pode sugerir padrões de relato ainda não exigidos pelo contexto específico.

Prompt 21.9.6 Revisão de referências

Objetivo: Identificar risco de referências inexistentes ou inadequadas.

Contexto de uso: Revisão bibliográfica e auditoria de citação.

Revise a lista de referências abaixo.

Marque itens que parecem incompletos, potencialmente inexistentes, incompatíveis com o tema, sem DOI/link ou fora do tipo solicitado.

Não invente correções; indique como verificar cada referência.

Lista: [REFERÊNCIAS].

Saída esperada: Lista de referências suspeitas e procedimentos de verificação.

Validação recomendada: Checar DOI, periódico, autores e ano em bases confiáveis.

Risco metodológico: A IA pode não detectar referência fabricada se ela parecer plausível.

Prompt 21.9.7 Revisão de linguagem avaliativa

Objetivo: Evitar juízos normativos indevidos.

Contexto de uso: Relatórios institucionais, rankings e *dashboards*.

Identifique no texto abaixo expressões avaliativas como "excelente", "fraco", "líder", "atrasado", "superior" ou "baixo desempenho".

Avalie se cada expressão é sustentada por critério explícito.

Proponha linguagem descritiva quando necessário.

Texto: [TEXTO].

Saída esperada: Lista de expressões avaliativas e reescritas neutras.

Validação recomendada: Validar com objetivos do relatório e critérios de avaliação definidos.

Risco metodológico: A IA pode eliminar linguagem legítima em pareceres deliberadamente avaliativos.

Prompt 21.9.8 Revisão por persona crítica

Objetivo: Simular revisão por diferentes perfis de avaliador.

Contexto de uso: Preparação para banca, submissão, reunião institucional ou *peer review*.

Revise o estudo abaixo sob três perspectivas: especialista em bibliometria, especialista do domínio e gestor de pesquisa.

Para cada perspectiva, identifique preocupações, perguntas, pontos fortes, fragilidades e recomendações.

Estudo: [DESCRIÇÃO].

Saída esperada: Revisão multiperspectiva.

Validação recomendada: Usar como preparação para revisão humana real, não como substituto.

Risco metodológico: Personas podem simplificar posições profissionais complexas.

21.10 Prompts para auditoria e reprodutibilidade

A reprodutibilidade em cientometria depende da preservação de dados brutos, estratégias de busca, datas de coleta, versões de bases, critérios de limpeza, scripts, parâmetros de ferramentas, *prompts* usados, saídas geradas e decisões humanas. A IA pode ajudar a transformar processos dispersos em protocolos estruturados, criar checklists de auditoria, apontar lacunas de documentação e preparar apêndices. Entretanto, a documentação gerada por IA deve ser baseada em registros reais. O princípio fundamental é simples: se uma etapa não foi registrada, não deve ser apresentada como se tivesse sido executada. *Prompts* de auditoria devem, portanto, exigir distinção entre informação fornecida, informação ausente e inferência não permitida (PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021).

Prompt 21.10.1 Checklist de reprodutibilidade

Objetivo: Verificar completude da documentação do estudo.

Contexto de uso: Auditoria interna e preparação de apêndices.

Crie um checklist de reprodutibilidade para o estudo cientométrico abaixo.

Inclua busca, coleta, exportação, limpeza, indicadores, redes, visualizações, *prompts*, validação humana, scripts, versões de ferramentas, dados brutos, dados tratados e limitações.

Estudo: [DESCRIÇÃO].

Saída esperada: Checklist detalhado de itens documentais.

Validação recomendada: Aplicar checklist antes da entrega final e anexar evidências.

Risco metodológico: A IA pode incluir itens irrelevantes e omitir especificidades do estudo.

Prompt 21.10.2 Registro de prompt

Objetivo: Documentar uso de IA de forma auditável.

Contexto de uso: Apêndice metodológico e trilha de auditoria.

Transforme as informações abaixo em uma ficha de documentação de *prompt*.

Inclua identificador, data, modelo, versão se disponível, objetivo, etapa do estudo, dados de entrada, *prompt* integral, saída resumida, decisão humana, validação realizada, limitações e responsável.

Informações: [DADOS].

Saída esperada: Ficha padronizada de documentação de *prompt*.

Validação recomendada: Salvar *prompt* integral e saída original em repositório do projeto.

Risco metodológico: A IA pode resumir *prompt* de modo a perder detalhes necessários à reprodução.

Prompt 21.10.3 Detecção de lacunas no protocolo

Objetivo: Identificar informações ausentes em protocolo cientométrico.

Contexto de uso: Antes de publicação, depósito de dados ou entrega institucional.

Audite o protocolo abaixo e liste informações ausentes ou insuficientes para reprodução.

Separe lacunas críticas, lacunas moderadas e melhorias desejáveis.

Para cada lacuna, explique consequência e como corrigir.

Protocolo: [PROTOCOLO].

Saída esperada: Relatório de lacunas do protocolo.

Validação recomendada: Completar lacunas com registros reais, não com reconstruções posteriores sem evidência.

Risco metodológico: A IA pode sugerir preenchimentos plausíveis que não correspondem ao que foi feito.

Prompt 21.10.4 Auditoria de dados e scripts

Objetivo: Relacionar dados, scripts e produtos analíticos.

Contexto de uso: Organização de repositório e auditoria de projeto.

Com base nos arquivos descritos abaixo, construa uma matriz que relacione dados brutos, dados tratados, scripts, parâmetros, tabelas, figuras e seções do relatório.

Indique dependências e pontos de verificação.

Arquivos: [LISTA].

Saída esperada: Matriz de rastreabilidade entre arquivos e produtos.

Validação recomendada: Verificar existência e integridade dos arquivos citados.

Risco metodológico: A IA pode assumir dependências não existentes se os nomes forem ambíguos.

Prompt 21.10.5 Declaração de uso de IA

Objetivo: Redigir declaração transparente e responsável.

Contexto de uso: Manuscritos, relatórios, teses e apêndices.

Redija uma declaração de uso de IA para o estudo abaixo.

Informe finalidades, etapas apoiadas, tipo de modelo, validação humana, restrições, ausência de uso para gerar dados inexistentes e responsabilidade dos autores.

Use tom compatível com artigo científico.

Estudo: [DESCRIÇÃO].

Saída esperada: Declaração de uso de IA clara e ética.

Validação recomendada: Adequar às normas do periódico, instituição ou agência.

Risco metodológico: A IA pode minimizar seu próprio papel ou omitir limitações importantes.

Prompt 21.10.6 - Plano de compartilhamento de dados

Objetivo: Planejar abertura ou preservação responsável dos dados.

Contexto de uso: Ciência aberta, reprodutibilidade e gestão de dados.

Elabore um plano de compartilhamento de dados para o estudo cientométrico [DESCRIÇÃO].

Inclua dados que podem ser compartilhados, restrições de licença das bases, formato, metadados, scripts, dicionários de limpeza, *prompts*, documentação e repositório recomendado.

Saída esperada: Plano de compartilhamento e preservação.

Validação recomendada: Verificar licenças das bases comerciais e permissões de redistribuição

Risco metodológico: Pode sugerir compartilhamento de dados protegidos por contrato.

Prompt 21.10.7 Replicação por terceiro

Objetivo: Simular revisão de replicabilidade por outro pesquisador.

Contexto de uso: Revisão final de protocolo e repositório.

Assuma o papel de um pesquisador externo tentando replicar o estudo.

Com base no protocolo abaixo, liste dúvidas, passos impossíveis de reproduzir, parâmetros ausentes, decisões ambíguas e solicitações de informação adicional.

Protocolo: [PROTOCOLO].

Saída esperada: Lista de obstáculos à replicação.

Validação recomendada: Usar como teste antes de disponibilizar o protocolo final.

Risco metodológico: A IA pode ser exigente demais em estudos exploratórios, mas isso ajuda a documentação.

Prompt 21.10.8 Auditoria final do relatório

Objetivo: Verificar consistência entre método, resultados e conclusões.

Contexto de uso: Etapa final antes de submissão, entrega ou publicação.

Audite o relatório final abaixo.

Verifique se método, resultados, figuras, tabelas, discussão, limitações, referências, uso de IA e conclusões estão consistentes.

Liste problemas críticos, moderados e menores.

Não reescreva o relatório; produza diagnóstico de qualidade.

Relatório: [TEXTO].

Saída esperada: Diagnóstico final de qualidade e consistência.

Validação recomendada: Aplicar revisão humana após a auditoria automatizada.

Risco metodológico: A IA pode deixar passar inconsistências numéricas se o texto for longo ou fragmentado.

21.11 Modelo integrado de cadeia de prompts

Embora cada *prompt* da biblioteca possa ser utilizado isoladamente, a maior contribuição metodológica surge quando os *prompts* são encadeados. Uma cadeia de *prompts* é uma sequência de interações em que a saída de uma etapa alimenta a etapa seguinte. Em cientometria, esse encadeamento deve seguir a lógica do fluxo de pesquisa: pergunta, busca, coleta, limpeza, análise, visualização, interpretação, redação e auditoria. A vantagem da cadeia é que ela torna explícitas as decisões intermediárias. A desvantagem é o risco de propagação de erros: uma busca mal formulada pode contaminar limpeza, indicadores, mapas e conclusões. Por isso, cada etapa deve incluir um ponto de validação antes de avançar.

Para cada cadeia, recomenda-se criar um identificador, por exemplo: CP-AREA-001, CP-INSTITUCIONAL-002 ou CP-TENDENCIAS-003. O identificador permite rastrear qual conjunto de *prompts* foi aplicado a qual *corpus*. Também é recomendável manter uma planilha

de controle com as colunas: etapa, *prompt* usado, data, modelo, entrada, saída, decisão humana, validação e observações. Essa planilha torna o processo mais reprodutível e facilita a preparação de apêndices metodológicos.

Quadro 21.1 - Cadeia mínima de *prompts* para estudo de área científica

Etapa	Ação	Prompt de referência
1	Delimitar o problema	*Prompt* 21.1.1
2	Escolher a pergunta principal	*Prompt* 21.1.3
3	Expandir o vocabulário	*Prompt* 21.2.1
4	Construir a estratégia booleana	*Prompt* 21.2.2
5	Documentar a busca	*Prompt* 21.2.7
6	Verificar metadados	*Prompt* 21.3.7
7	Normalizar termos e instituições	*Prompts* 21.3.1 e 21.3.3
8	Selecionar indicadores	*Prompt* 21.4.1
9	Interpretar em três níveis	*Prompt* 21.7.1
10	Auditar o relatório final	*Prompt* 21.10.8

Fonte: Elaborado pelo autor (2026).

21.12 Ficha padrão de documentação de prompt

A seguir apresenta-se um modelo textual de ficha que pode ser usado em relatórios, apêndices, repositórios ou cadernos de pesquisa. A ficha deve ser preenchida com informações reais, sem reconstruir etapas inexistentes. Quando uma informação não estiver disponível, deve-se registrar explicitamente "não registrado" ou "não aplicável", em vez de produzir uma versão artificialmente completa.

Quadro 21.2 - Campos da ficha de documentação de *prompt*

Campo	Conteúdo
Identificador do *prompt*	código único da etapa.
Data e horário de execução	momento em que o *prompt* foi utilizado.
Modelo ou ferramenta	nome do sistema de IA utilizado, quando permitido.
Etapa do estudo	formulação, busca, limpeza, indicadores, redes, interpretação, redação ou auditoria.
Objetivo do *prompt*	tarefa metodológica que se pretendia realizar.
Dados de entrada	descrição dos dados fornecidos ao modelo.

Campo	Conteúdo
Prompt integral	texto completo da instrução enviada.
Saída obtida	resumo da resposta ou arquivo associado.
Decisão humana	o que foi aceito, alterado, rejeitado ou enviado para validação.
Validação realizada	fontes, especialistas, scripts, amostras ou verificações usados.
Limitações	problemas, incertezas ou riscos identificados.
Responsável	pesquisador que executou e validou a etapa.

Fonte: Elaborado pelo autor (2026).

21.13 Critérios de qualidade e matriz de riscos

Um bom *prompt* cientométrico deve cumprir cinco critérios. O primeiro é especificidade: deve indicar a tarefa e o contexto, evitando comandos genéricos como "analise estes dados". O segundo é rastreabilidade: deve permitir entender quais dados foram fornecidos e quais decisões foram solicitadas. O terceiro é prudência interpretativa: deve impedir inferências fortes sem evidência. O quarto é formato de saída: deve solicitar estrutura adequada para revisão, como tabelas textuais, listas de riscos, graus de confiança ou matrizes. O quinto é validação: deve pedir explicitamente como conferir a resposta (BROWN *et al.*, 2020; LIU *et al.*, 2023; WEI *et al.*, 2022; WANG *et al.*, 2023; YAO *et al.*, 2023).

Técnicas como *few-shot prompting*, *chain-of-thought*, *self-consistency* e instruções estruturadas podem melhorar o desempenho em tarefas específicas. Em cientometria, contudo, a qualidade deve ser aferida pela verificabilidade da resposta em relação a um *corpus* documentado, e não apenas pela fluência ou pela correção aparente (BROWN *et al.*, 2020; WEI *et al.*, 2022; WANG *et al.*, 2023).

A qualidade do *prompt* também depende da qualidade dos dados de entrada. Um *prompt* bem construído não corrige metadados incompletos, estratégias de busca enviesadas, duplicatas não tratadas ou indicadores mal escolhidos. Ele pode identificar problemas, mas não substitui trabalho de curadoria. Esse ponto é coerente com princípios de métricas responsáveis: indicadores devem apoiar julgamento especializado e não substituí-lo (HICKS *et al.*, 2015). Da mesma forma, *prompts* devem apoiar análise científica e não deslocar a responsabilidade interpretativa para o sistema de IA.

Quadro 21.3 - Critérios de qualidade para *prompts* cientométricos

Critério	Requisito
Especificidade	Explicita tarefa, contexto, dados e restrições.
Rastreabilidade	Permite identificar entradas, decisões e transformações.
Prudência interpretativa	Distingue dados, inferências, hipóteses e limites.
Formato de saída	Solicita estrutura que facilite revisão e comparação.
Validação	Indica como a resposta será conferida por fontes, dados ou especialistas.

Fonte: Elaborado pelo autor (2026).

A matriz abaixo sintetiza riscos recorrentes no uso de *prompts* em cientometria e formas de mitigação. Ela deve ser utilizada como instrumento de revisão antes de incorporar saídas de IA em produtos finais.

Quadro 21.4 - Matriz de riscos no uso de *prompts*

Risco	Exemplo	Mitigação
Alucinação	Referência, autor ou base inexistente.	Verificar toda afirmação factual em fonte externa.
Extrapolação	Transformar correlação em causalidade.	Exigir separação entre dado, inferência e hipótese.
Viés de base	Generalizar resultados de uma base para toda a ciência.	Relatar cobertura, idioma, áreas e tipos documentais.
Rótulo sedutor	Nomear *cluster* com termo elegante, mas impreciso.	Validar rótulos com documentos representativos e especialistas.
Perda de rastreabilidade	Usar resposta sem registrar *prompt* e dados.	Manter ficha de documentação de *prompt*.
Simplificação	Reduzir limitações em sumário executivo.	Manter limitações essenciais em todas as versões.

Fonte: Elaborado pelo autor (2026).

21.14 Estudo de caso integrador: da demanda institucional ao relatório reprodutível

Para demonstrar o uso articulado da biblioteca, imagine-se uma instituição universitária interessada em compreender sua posição em pesquisas sobre sustentabilidade urbana. A demanda inicial é vaga: "queremos saber como estamos nesse tema". Uma resposta imediata baseada apenas em contagem de publicações seria insuficiente, pois a pergunta institucional contém dimensões implícitas: quais subtemas são mais fortes, com quem a instituição colabora, em que periódicos publica, quais áreas participam do tema, que países ou instituições aparecem como parceiros e que oportunidades estratégicas podem ser identificadas. O primeiro passo, portanto, é transformar a demanda em pergunta cientométrica operacional. O *prompt* 21.1.1 permite formular alternativas, e o *prompt* 21.1.5 adapta a pergunta para a linguagem de gestão sem reduzir a análise a ranking.

Em seguida, a equipe precisa construir o *corpus*. O *prompt* 21.2.1 expande o vocabulário, incluindo termos como "*urban sustainability*", "*sustainable cities*", "*urban resilience*", "*smart sustainable cities*", "*urban transitions*" e equivalentes em português e espanhol. O *prompt* 21.2.2 transforma o vocabulário em *string* booleana. O *prompt* 21.2.5 adapta a estratégia para bases diferentes, respeitando campos e operadores. Nessa etapa, a IA não decide a estratégia final. Ela oferece uma versão inicial que deve ser testada. A equipe executa buscas piloto, examina amostras de títulos e resumos, identifica ruídos e lacunas, e então usa o *prompt* 21.2.4 para revisar a *string*. A data, o número de resultados e a *string* final são registrados para reprodução.

Depois da coleta, o desafio é preparar os metadados. A instituição pode aparecer com várias formas de afiliação, como nome em português, inglês, sigla, campus, laboratório ou unidade acadêmica. O *prompt* 21.3.3 auxilia a organizar variações, mas a validação deve usar fontes oficiais, identificadores institucionais e revisão humana. Palavras-chave também

exigem normalização: "*urban resilience*" e "*resilient cities*" podem pertencer ao mesmo grupo conceitual em alguns contextos, mas não em todos. O *prompt* 21.3.1 permite agrupar termos com grau de confiança e marcar casos ambíguos. O *prompt* 21.3.5 transforma decisões em dicionário de normalização, preservando rastreabilidade.

A seleção de indicadores deve responder à pergunta institucional, e não apenas reproduzir um painel padrão. Se o objetivo é diagnóstico de presença científica, indicadores de produção anual, áreas, periódicos e autores podem bastar. Se o objetivo é estratégia de colaboração, redes de coautoria, colaboração internacional e interinstitucional se tornam centrais. Se a pergunta envolve emergência temática, mapas temporais e detecção de crescimento de termos são mais adequados. O *prompt* 21.4.1 organiza essa relação entre pergunta e métrica, enquanto o *prompt* 21.4.8 ajuda a auditar indicadores inadequados. A equipe deve evitar afirmar que uma área é "melhor" apenas porque tem mais publicações, ou que um grupo é "mais relevante" apenas por centralidade em rede.

Na etapa interpretativa, a biblioteca atua como proteção contra narrativas excessivas. O *prompt* 21.7.1 separa descrição objetiva, interpretação provável e hipótese. Por exemplo, se há crescimento de publicações após 2018, a descrição objetiva é o aumento observado; uma interpretação provável pode relacioná-lo à ampliação da agenda internacional de sustentabilidade urbana; uma hipótese a testar seria a influência de políticas locais, editais ou parcerias internacionais. O *prompt* 21.7.4 permite triangulação com literatura e conhecimento especializado. O *prompt* 21.7.8 transforma limitações em seção explícita, incluindo cobertura da base, recorte temporal e possível sub-representação de produção local não indexada.

Finalmente, o relatório é redigido e auditado. O *prompt* 21.8.4 gera sumário executivo para gestores, mas deve manter limitações. O *prompt* 21.8.5 escreve legendas analíticas de figuras, evitando que mapas visuais sejam interpretados como evidência causal. O *prompt* 21.10.8 audita consistência entre método, resultados e conclusões. O produto final pode incluir um relatório técnico, um *dashboard*, um apêndice metodológico, um conjunto de arquivos de dados tratados, dicionários de limpeza, *prompts* utilizados e uma declaração de uso de IA. Esse estudo de caso mostra que a biblioteca não é um conjunto decorativo de comandos, mas um arcabouço para tornar o uso da IA mais científico, explícito e responsável.

21.15 Adaptação dos prompts por área, idioma e maturidade do campo

Nenhum *prompt* desta biblioteca deve ser usado sem adaptação ao campo científico. Áreas biomédicas, engenharias, ciências sociais, humanidades, ciência da computação, educação, artes e saúde coletiva possuem regimes distintos de publicação, autoria, citação, indexação e circulação. Em alguns campos, artigos de periódicos concentram a produção reconhecida; em outros, livros, capítulos, anais de conferência, documentos técnicos, normas, relatórios e *preprints* desempenham papel fundamental. Um *prompt* que pede "artigos indexados" pode ser adequado para uma área e reducionista para outra. Assim, a primeira adaptação consiste em explicitar a cultura documental do campo antes de solicitar indicadores ou mapas.

A adaptação linguística também é crucial. A ciência publicada em inglês tende a ser mais visível em bases internacionais, mas muitos campos possuem produção substantiva em português, espanhol, francês, alemão ou outros idiomas. Em estudos sobre políticas públicas, educação, saúde coletiva, direito, história ou desenvolvimento regional, excluir literatura local pode gerar diagnóstico enviesado. *Prompts* de busca devem solicitar variantes linguísticas e também perguntar se a tradução é equivalente, parcial ou problemática. Do mesmo modo, *prompts* de interpretação devem considerar que baixa citação internacional não implica

irrelevância social ou nacional (MONGEON; PAUL-HUS, 2016; MARTÍN-MARTÍN *et al.*, 2021; LARIVIÈRE *et al.*, 2013).

A maturidade do campo altera a escolha dos *prompts*. Em campos consolidados, redes de cocitação e indicadores de impacto podem revelar estrutura intelectual acumulada. Em temas emergentes, acoplamento bibliográfico, palavras-chave, resumos e *preprints* podem ser mais informativos do que citações, pois o tempo de acumulação citacional é curto. Em áreas ainda difusas, perguntas exploratórias e análise terminológica são prioritárias. Em áreas institucionalizadas, perguntas comparativas, indicadores normalizados e avaliação de colaboração podem ser mais úteis. A biblioteca deve ser lida como repertório flexível, não como sequência obrigatória.

Também é necessário adaptar a biblioteca ao objetivo do estudo. Um artigo acadêmico exige metodologia detalhada, discussão teórica e referências robustas. Um relatório de gestão exige síntese, visualizações claras e recomendações prudentes. Um parecer de política científica precisa distinguir evidência, incerteza e decisão. Um diagnóstico institucional exige cuidado com efeitos reputacionais e incentivos perversos. Um estudo de organização do conhecimento demanda atenção a vocabulário, categorias e curadoria de acervos. O mesmo achado pode ser apresentado de maneira diferente em cada produto, mas sua base empírica e suas limitações não devem ser alteradas para atender expectativas retóricas.

Por fim, a adaptação deve considerar o modelo de IA utilizado. Modelos distintos podem variar em precisão, tendência à alucinação, sensibilidade à formulação, capacidade de seguir formatos e habilidade em lidar com textos longos. Quando um *prompt* é crítico para o estudo, recomenda-se testar variações, comparar saídas, registrar diferenças e submeter decisões a validação humana. Essa prática aproxima a engenharia de *prompts* de um procedimento experimental: não se assume que uma formulação é universalmente adequada; testa-se, documenta-se e justifica-se a escolha.

21.16 Banco de meta-prompts para calibrar a biblioteca

Além dos *prompts* substantivos apresentados nas seções anteriores, estudos mais robustos podem usar meta-*prompts*, isto é, *prompts* que revisam, melhoram ou testam outros *prompts*. Eles são úteis quando a equipe deseja padronizar procedimentos, treinar pesquisadores ou comparar versões de instruções. Abaixo apresentam-se seis meta-*prompts* que podem ser acoplados a qualquer etapa da biblioteca.

Meta-prompt 21.A – Diagnóstico de clareza

Analise o *prompt* abaixo e identifique ambiguidades, termos vagos, ausência de contexto, falta de dados de entrada, formato de saída indefinido e riscos de interpretação.

Proponha uma versão revisada, mantendo o objetivo original. *Prompt*: [*PROMPT*].

Uso recomendado. aplicar antes de consolidar *prompts* que serão usados em etapas críticas do estudo, como busca, limpeza, rotulagem de *clusters* e interpretação. A validação deve comparar respostas obtidas por diferentes versões e registrar a escolha final.

Meta-prompt 21.B – Transformação em formato auditável

Reescreva o *prompt* abaixo para que a resposta seja auditável.

Inclua solicitações de justificativa, grau de confiança, evidência necessária, limitações e itens que devem ser verificados por humanos.

Prompt: [*PROMPT*].

Uso recomendado. aplicar antes de consolidar *prompts* que serão usados em etapas críticas do estudo, como busca, limpeza, rotulagem de *clusters* e interpretação. A validação deve comparar respostas obtidas por diferentes versões e registrar a escolha final.

Meta-prompt 21.C – Comparação de versões

Compare as versões A e B de um *prompt* cientométrico.

Avalie especificidade, risco de alucinação, controle de extrapolação, clareza do formato de saída, adequação à pergunta e facilidade de validação.

Recomende a melhor versão ou proponha uma terceira.

Versão A: [A]. Versão B: [B].

Uso recomendado. aplicar antes de consolidar *prompts* que serão usados em etapas críticas do estudo, como busca, limpeza, rotulagem de *clusters* e interpretação. A validação deve comparar respostas obtidas por diferentes versões e registrar a escolha final.

Meta-prompt 21.D – Redução de viés

Revise o *prompt* abaixo para reduzir vieses linguísticos, disciplinares, geográficos, institucionais e de base de dados.

Explique quais alterações foram feitas e por quê. *Prompt*: [*PROMPT*].

Uso recomendado. aplicar antes de consolidar *prompts* que serão usados em etapas críticas do estudo, como busca, limpeza, rotulagem de *clusters* e interpretação. A validação deve comparar respostas obtidas por diferentes versões e registrar a escolha final.

Meta-prompt 21.E – Preparação para replicação

Transforme o *prompt* abaixo em uma instrução adequada para um protocolo reprodutível.

Acrescente campos de entrada, critérios de decisão, forma de registrar saída e procedimento de validação.

Prompt: [*PROMPT*].

Uso recomendado. aplicar antes de consolidar *prompts* que serão usados em etapas críticas do estudo, como busca, limpeza, rotulagem de *clusters* e interpretação. A validação deve comparar respostas obtidas por diferentes versões e registrar a escolha final.

Meta-prompt 21.F – Teste de robustez

Proponha três variações controladas do *prompt* abaixo para testar sensibilidade da resposta.

Explique que diferença cada variação pretende avaliar e como comparar as saídas.

Prompt: [*PROMPT*].

Uso recomendado. aplicar antes de consolidar *prompts* que serão usados em etapas críticas do estudo, como busca, limpeza, rotulagem de *clusters* e interpretação. A validação deve comparar respostas obtidas por diferentes versões e registrar a escolha final.

21.17 Boas práticas para equipes de pesquisa

Quando a biblioteca é utilizada por uma equipe, e não por pesquisador individual, recomenda-se definir papéis. Um pesquisador pode ser responsável pela formulação dos *prompts*; outro pela verificação das saídas; outro pela conferência de referências; outro pela validação de metadados; e outro pela redação final. Essa divisão evita que a mesma pessoa gere, aceite e publique uma saída sem controle independente. Em estudos com impacto institucional, a dupla revisão é especialmente importante.

A equipe deve adotar convenções de versionamento. Um *prompt* pode existir como versão 1.0, 1.1 ou 2.0, conforme seja ajustado. Mudanças pequenas, como inclusão de formato de saída, podem ser registradas como revisão menor. Mudanças que alteram critérios de decisão, escopo ou interpretação devem ser registradas como revisão maior. O mesmo se aplica a dicionários de limpeza, *strings* de busca e scripts. A lógica de versionamento aproxima o trabalho cientométrico de boas práticas de ciência aberta e engenharia de dados.

Outra prática útil é manter um diário de decisões. Nem toda decisão será automatizada em script, mas toda decisão relevante deve ser documentada: por que um termo foi excluído, por que uma base foi priorizada, por que um *cluster* recebeu determinado nome, por que uma métrica foi abandonada, por que um período foi ajustado. A IA pode ajudar a redigir esse diário, mas as justificativas devem vir da equipe. Um relatório cientométrico robusto não é apenas o conjunto de seus resultados; é também a história documentada de suas escolhas metodológicas.

No treinamento de estudantes e analistas, a biblioteca pode ser usada como material pedagógico. Em vez de pedir que estudantes apenas "usem IA", pode-se solicitar que comparem três *prompts*, identifiquem riscos, validem respostas e documentem decisões. Essa prática ensina que a interação com IA é parte do método e não atalho para contornar o método. Em contextos institucionais, a biblioteca pode compor manuais internos de boas práticas para

relatórios bibliométricos, com adaptações às bases assinadas pela instituição e aos padrões de governança de dados.

Por fim, equipes devem manter cautela com dados sensíveis. Embora metadados bibliográficos sejam frequentemente públicos, relatórios institucionais podem conter informações estratégicas, avaliações internas, nomes de pesquisadores, desempenho de programas ou diagnósticos preliminares. Antes de inserir dados em sistemas de IA, é necessário verificar políticas de privacidade, contratos institucionais e restrições de confidencialidade. A biblioteca pressupõe uso responsável, mas cada instituição deve definir regras próprias para proteção de dados e governança da IA.

21.18 Integração com ferramentas cientométricas

A biblioteca de *prompts* não substitui ferramentas como bibliometrix, VOSviewer, CiteSpace, SciMAT, Gephi, Pajek, planilhas, linguagens de programação ou APIs bibliográficas. Ela atua antes, durante e depois do uso dessas ferramentas. Antes, ajuda a formular a pergunta, desenhar a coleta e escolher indicadores. Durante, ajuda a revisar parâmetros, interpretar saídas e identificar problemas. Depois, ajuda a redigir resultados, revisar limitações e preparar documentação. Essa integração é fundamental para evitar a ilusão de que a IA, sozinha, executa análise cientométrica (ARIA; CUCCURULLO, 2017; CHEN, 2006; COBO *et al.*, 2012; VAN ECK; WALTMAN, 2010).

Em bibliometrix, por exemplo, *prompts* podem apoiar a escolha de análises descritivas, redes, mapas temáticos e relatórios Biblioshiny. Em VOSviewer, podem ajudar a interpretar mapas de coocorrência, cocitação e acoplamento, desde que a IA receba uma descrição adequada dos *clusters* e métricas. Em CiteSpace, podem apoiar a leitura de *burst detection*, linhas do tempo e nós estruturantes. Em SciMAT, podem ajudar a interpretar evolução temática e matrizes estratégicas. Em todos os casos, a IA deve receber saídas verificáveis e não apenas imagens sem contexto.

Ao trabalhar com APIs, *prompts* podem ajudar a gerar pseudocódigo, listas de campos, checklists de paginação, tratamentos de erro e documentação. Entretanto, código gerado por IA deve ser revisado, testado e comentado. Parâmetros como *endpoint*, filtros, limites de requisição, chaves de API e licenças não devem ser presumidos. A documentação oficial das bases continua sendo fonte primária. O papel do *prompt* é organizar o raciocínio e acelerar rascunhos, não substituir consulta técnica.

A integração mais produtiva ocorre quando cada saída de ferramenta é acompanhada de um *prompt* de interpretação e outro de auditoria. Por exemplo, após gerar um mapa de coocorrência, a equipe pode usar um *prompt* para nomear *clusters* e outro para auditar rótulos. Após calcular indicadores, pode usar um *prompt* para interpretar e outro para detectar extrapolações. Após redigir o relatório, pode usar um *prompt* para revisar conclusões e outro para verificar reprodutibilidade. Essa dupla função, produtiva e crítica, reduz riscos e aumenta a qualidade do estudo.

21.19 Limites da biblioteca e agenda de atualização

Uma biblioteca de *prompts* não é um artefato estático. Modelos de linguagem mudam, interfaces de bases bibliográficas mudam, políticas editoriais mudam, padrões de avaliação científica mudam e vocabulários de pesquisa se transformam. Por essa razão, os *prompts* reunidos neste capítulo devem ser revisados periodicamente. O que hoje é uma boa instrução para obter uma matriz de análise pode tornar-se insuficiente diante de modelos com novas

capacidades, novas restrições ou novos padrões de segurança. A atualização deve preservar a lógica metodológica: clareza, rastreabilidade, validação e responsabilidade.

Outro limite é que *prompts* não resolvem conflitos epistemológicos. Se uma área discorda sobre a definição de um conceito, o *prompt* pode mapear as definições concorrentes, mas não decidir qual é verdadeira. Se uma instituição deseja usar indicadores para decisões de financiamento, o *prompt* pode indicar riscos, mas não eliminar disputas políticas. Se uma base sub-representa produção periférica, o *prompt* pode alertar para o viés, mas não reconstruir automaticamente a literatura ausente. A biblioteca amplia a capacidade de formular boas perguntas; ela não dissolve as tensões da avaliação científica.

Também é necessário reconhecer que a própria interação com IA pode introduzir vieses. *Prompts* em inglês podem produzir respostas mais alinhadas à literatura anglófona; *prompts* excessivamente genéricos podem favorecer categorias dominantes; *prompts* que pedem "principais autores" podem reproduzir hierarquias de visibilidade; *prompts* que solicitam "melhores periódicos" podem reforçar dependência de métricas de prestígio. Uma biblioteca responsável deve, portanto, conter não apenas *prompts* produtivos, mas também *prompts* críticos capazes de questionar os pressupostos dos primeiros.

A agenda de atualização da biblioteca deve incluir testes comparativos entre modelos, avaliação de consistência de respostas, revisão de referências, incorporação de novas diretrizes de transparência no uso de IA, adaptação para diferentes áreas do conhecimento e tradução cuidadosa para outras línguas. Também é recomendável que grupos de pesquisa registrem casos de falha: *prompts* que geraram classificações equivocadas, rótulos inadequados, citações inexistentes ou conclusões exageradas. Esses casos são pedagogicamente valiosos porque mostram que a competência em IA não consiste apenas em produzir boas respostas, mas em reconhecer respostas ruins.

Em síntese, a biblioteca deve ser compreendida como infraestrutura metodológica viva. Seu valor depende menos da quantidade de *prompts* e mais da qualidade das práticas que ela induz. Quando usada sem validação, ela pode acelerar erros. Quando usada com documentação, crítica e curadoria, pode elevar o padrão de transparência dos estudos cientométricos assistidos por inteligência artificial.

21.20 Considerações finais

A biblioteca de *prompts* cientométricos apresentada neste capítulo organiza a interação com modelos de linguagem como prática metodológica. Seu propósito não é oferecer fórmulas fixas, mas padrões de raciocínio aplicáveis a diferentes etapas de um estudo: formular perguntas, buscar literatura, limpar dados, selecionar indicadores, interpretar redes, nomear *clusters*, redigir resultados, revisar criticamente e documentar a reprodutibilidade. Em todos os casos, a IA é tratada como instrumento de apoio e interlocução, não como fonte autônoma de autoridade científica.

A principal contribuição dessa biblioteca é inserir no próprio *prompt* os mecanismos de controle que a cientometria exige: delimitação de unidade de análise, explicitação de bases, indicação de metadados, grau de confiança, validação humana, separação entre dados e hipóteses, registro de limitações e prevenção de extrapolações. Quando esses mecanismos são ignorados, a IA tende a produzir respostas fluentes, mas metodologicamente frágeis. Quando são incorporados, o *prompt* passa a funcionar como protocolo mínimo de interação.

O uso responsável da biblioteca exige adaptação. Cada campo científico possui vocabulários, práticas de publicação, padrões de autoria, culturas de citação e bases mais

ou menos adequadas. Um *prompt* que funciona bem para literatura biomédica pode ser inadequado para humanidades; um indicador útil para física pode ser injusto para artes; uma base forte em periódicos internacionais pode invisibilizar produção regional. Portanto, os *prompts* devem ser ajustados por área, língua, base, objetivo e público.

Por fim, a biblioteca reforça uma tese central deste livro: a cientometria com IA não é a substituição do pesquisador por um sistema automatizado, mas a construção de uma prática analítica aumentada, documentada e crítica. O valor científico não está apenas em obter respostas rápidas, mas em tornar visíveis as decisões que produzem o *corpus*, os indicadores, os mapas e as interpretações. *Prompts* bem documentados ampliam a capacidade de análise; *prompts* não documentados obscurecem o método. A diferença entre esses dois usos define a qualidade epistemológica da cientometria assistida por inteligência artificial.

Capítulo 22

Conclusão: cientometria com inteligência artificial como prática metodológica responsável

22.1 Tese de encerramento

A tese central da obra é que a cientometria com inteligência artificial não deve ser compreendida como simples automação de tarefas bibliométricas, mas como uma reorganização metodológica do modo de formular perguntas, construir o *corpus*, calcular indicadores, interpretar redes, comunicar resultados e documentar decisões. A IA é uma camada de apoio à análise da ciência, e não substituta da teoria cientométrica, da curadoria de dados, da interpretação especializada ou da responsabilidade autoral. Essa posição é coerente com a literatura de métodos bibliométricos, que enfatiza a articulação entre pergunta, dados, técnicas e interpretação, e com as diretrizes de métricas responsáveis, que subordinam indicadores ao julgamento contextualizado (ZUPIC; ČATER, 2015; DONTHU *et al.*, 2021; HICKS *et al.*, 2015).

A engenharia de *prompts* assume, nesse quadro, relevância operacional e epistemológica. Um *prompt* cientométrico robusto define o papel atribuído ao modelo, delimita a tarefa, explicita as evidências disponíveis, impõe restrições interpretativas, solicita um formato de saída e estabelece procedimentos de validação. Quando registrado e versionado, passa a integrar o protocolo do estudo de modo semelhante a uma estratégia de busca, um script de análise, uma matriz de codificação ou uma regra de normalização. A literatura sobre métodos de *prompting* confirma que a formulação da instrução, os exemplos, a decomposição da tarefa e os mecanismos de consistência influenciam as respostas dos modelos (LIU *et al.*, 2023; WEI *et al.*, 2022; WANG *et al.*, 2023; WHITE *et al.*, 2023).

A cientometria, por sua vez, não se reduz à contagem de publicações, citações ou palavras-chave. Indicadores, redes e mapas são construções mediadas por bases, critérios de normalização, algoritmos, limiares, formas de visualização e interpretações humanas. A presença de modelos generativos torna ainda mais urgente registrar o que foi solicitado, quais dados foram fornecidos, que respostas foram produzidas e quais saídas foram aceitas, rejeitadas ou revistas. O uso responsável da IA depende, portanto, de uma cultura metodológica de rastreabilidade.

Mensagem central: A IA pode ampliar a inteligência coletiva empregada na análise da ciência, mas não deve ocupar o lugar da deliberação científica. Seu valor depende da qualidade dos dados, da clareza do *prompt*, da validação das saídas e da responsabilidade humana pelas conclusões.

22.2 Síntese do percurso metodológico da obra

O percurso do livro foi estruturado como um fluxo de decisões interdependentes, da formulação do problema à auditoria do produto final. O Quadro 21.1 consolida esse itinerário e explicita, para cada etapa, a contribuição possível da IA e o controle metodológico indispensável.

Quadro 22.1 - Síntese do ciclo cientométrico assistido por IA

Etapa	Apoio possível da IA	Controle indispensável
Pergunta e desenho	Decompor objetivos, sugerir unidades de análise, recortes e produtos.	Validar relevância, viabilidade e alinhamento entre pergunta, teoria e dado.
Busca e coleta	Expandir vocabulários, comparar termos, estruturar *strings* e planejar metadados.	Testar nas bases reais, registrar data, campos, filtros e cobertura.
Limpeza e normalização	Sugerir deduplicação, variantes autorais, instituições e termos preferenciais.	Manter dicionários, graus de confiança e revisão humana dos casos ambíguos.
Indicadores	Relacionar métricas à pergunta e explicar limites para diferentes públicos.	Evitar equivalência entre quantidade, impacto e qualidade científica.
Redes e mapas	Apoiar leitura estrutural, rotulagem de *clusters* e comparação de comunidades.	Documentar parâmetros e não converter proximidade ou associação em causalidade.
Evolução temporal	Organizar séries, mudanças terminológicas e sinais de emergência.	Testar artefatos de indexação, busca, recência e variação de cobertura.
Leitura e classificação	Sumarizar resumos, extrair categorias e classificar documentos.	Aplicar amostragem, concordância, justificativas textuais e auditoria dos erros.
Interpretação e comunicação	Estruturar resultados, discussões, legendas, briefings e relatórios.	Separar descrição, inferência, hipótese e recomendação; explicitar limitações.

Fonte: Elaborado pelo autor (2026).

A qualidade do *corpus* condiciona todas as etapas posteriores. Bases bibliográficas possuem diferenças de cobertura disciplinar, regional, linguística e documental; por isso, a escolha da fonte deve ser justificada e, quando pertinente, comparada com alternativas (MONGEON; PAUL-HUS, 2016; VISSER; VAN ECK; WALTMAN, 2021). A busca precisa ser relatada com precisão, incluindo estratégia, data, filtros e critérios de elegibilidade, em consonância com princípios de transparência aplicados às revisões e às pesquisas bibliográficas (PAGE *et al.*, 2021; RETHLEFSEN *et al.*, 2021).

Também se reafirma que indicadores medem dimensões distintas. O índice *h*, por exemplo, combina produtividade e citação, mas depende do tempo de carreira, da área e da base considerada (HIRSCH, 2005). Indicadores de impacto citacional apresentam pressupostos e limitações próprios (WALTMAN, 2016), enquanto interdisciplinaridade, colaboração e acesso aberto exigem medidas específicas e contextualizadas (PORTER *et al.*, 2007; PIWOWAR *et al.*, 2018).

No mapeamento científico, coautoria, cocitação, acoplamento bibliográfico e coocorrência de palavras representam relações diferentes. A indexação de citações e as relações entre documentos constituem fundamentos históricos do campo (GARFIELD, 1955).

Cocitação e acoplamento têm bases conceituais próprias (SMALL, 1973; KESSLER, 1963), enquanto a análise de copalavras e os métodos contemporâneos de mapeamento permitem representar estruturas conceituais e temáticas (CALLON; COURTIAL; LAVILLE, 1991; VAN ECK; WALTMAN, 2010; COBO *et al.*, 2011; COBO *et al.*, 2012). A IA pode apoiar a interpretação dessas estruturas, mas os rótulos devem permanecer provisórios até serem confrontados com termos, documentos e conhecimento especializado.

22.3 Engenharia de prompts como infraestrutura metodológica

A engenharia de *prompts* foi tratada como infraestrutura metodológica porque conecta o problema de pesquisa, o *corpus*, o modelo computacional e a interpretação humana. Um comando genérico tende a produzir respostas genéricas; uma instrução estruturada reduz ambiguidades e facilita a auditoria. O Quadro 21.2 apresenta os componentes mínimos recomendados.

Todo estudo cientométrico que utilize IA deve manter um registro de *prompts*, contendo identificador, versão, data, modelo, dados de entrada, resposta, decisão humana e validação. Em projetos complexos, esse registro pode assumir a forma de diário metodológico. Sem ele, torna-se difícil distinguir o que deriva dos dados, da interpretação humana ou de uma sugestão algorítmica.

Cadeias de *prompts* podem organizar sequências como pergunta, busca, coleta, limpeza, análise, interpretação e redação. Contudo, cadeias longas propagam erros. Por isso, precisam incluir pontos de verificação nos quais a equipe revise amostras, confronte resultados e decida se a etapa seguinte pode começar. Técnicas como raciocínio passo a passo e autoconsistência podem melhorar certas tarefas, mas não transformam a resposta em evidência e não dispensam avaliação humana (WEI *et al.*, 2022; WANG *et al.*, 2023).

A variabilidade das respostas também deve ser tratada como informação metodológica. Se execuções independentes geram classificações ou interpretações incompatíveis, a tarefa pode estar subespecificada, os dados podem ser insuficientes ou o fenômeno pode ser genuinamente ambíguo. Nesses casos, a resposta adequada não é escolher automaticamente a formulação mais convincente, mas revisar o *prompt*, ampliar as evidências e registrar a incerteza.

Quadro 22.2 - Componentes de um *prompt* cientométrico auditável

Componente	Função metodológica	Pergunta de controle
Contexto	Situa domínio, finalidade, etapa, público e escopo.	O modelo sabe em que estudo e etapa está atuando?
Papel	Define a perspectiva funcional esperada.	A resposta deve assumir visão de cientometrista, bibliotecário, revisor ou gestor?
Tarefa	Transforma a intenção em operação delimitada.	O que deve ser classificado, comparado, revisado, explicado ou redigido?
Evidências de entrada	Delimita a base factual disponível.	Quais dados, tabelas, textos ou metadados sustentam a resposta?
Restrições	Proíbe invenções, causalidade indevida e extrapolação.	O que o modelo não pode concluir ou completar?
Saída e validação	Define estrutura, confiança e verificações.	Como a resposta será revisada, comparada e incorporada ao protocolo?

Fonte: Elaborado pelo autor (2026).

22.4 IA como interlocutor crítico, não como autoridade epistêmica

Modelos de linguagem produzem respostas fluentemente formuladas mesmo quando a base factual é incompleta. Essa propriedade pode criar aparência de competência e autoridade. Entre os riscos documentados estão fabricação de informações, reprodução de vieses e dificuldade de vincular cada enunciado a evidências verificáveis (BENDER *et al.*, 2021; JI *et al.*, 2023). Em cientometria, esses problemas podem aparecer como referências inexistentes, métricas inadequadas, rótulos sedutores para *clusters*, normalizações indevidas e conclusões mais fortes do que os dados permitem.

A posição metodologicamente produtiva é utilizar a IA como interlocutora crítica: solicitar interpretações alternativas, identificação de vieses, revisão de conclusões, explicitação de incertezas e indicação de evidências faltantes. Esse uso amplia a capacidade de autoavaliação da equipe sem delegar a decisão. A pergunta decisiva não é se o modelo "sabe" a resposta, mas se sua saída ajuda a testar a coerência entre dado, método e argumento.

Modelos especializados em textos científicos podem melhorar tarefas de extração, classificação e representação semântica, como demonstram arquiteturas de transformadores e modelos treinados em literatura acadêmica (VASWANI *et al.*, 2017; DEVLIN *et al.*, 2019; BELTAGY; LO; COHAN, 2019). A especialização, entretanto, não elimina a necessidade de validação: ela altera o perfil dos erros e exige avaliação em corpora e tarefas cientométricas concretas.

A IA também pode traduzir resultados para pesquisadores, gestores e públicos não especializados. Essa adaptação é útil desde que preserve definições, limitações e condições de validade. Simplificar a linguagem não significa simplificar a evidência nem ocultar incertezas. A responsabilidade pelo texto, pelas citações e pelas consequências institucionais permanece humana.

22.5 Limites e riscos persistentes

Cinco riscos permanecem centrais mesmo com a evolução dos modelos. Eles não constituem razões para rejeitar a tecnologia, mas exigem controles proporcionais ao impacto do estudo.

Os vieses das bases antecedem o uso da IA e podem ser amplificados por ela. Cobertura desigual, preferência por determinados tipos documentais e concentração da comunicação científica em circuitos específicos afetam o que se torna visível (MONGEON; PAUL-HUS, 2016; VISSER; VAN ECK; WALTMAN, 2021). Da mesma forma, métricas e rankings podem induzir usos avaliativos descontextualizados; por isso, sua interpretação deve observar finalidade, campo, escala e missão institucional (HICKS *et al.*, 2015; WALTMAN, 2016).

Quadro 22.3 - Riscos persistentes e medidas de controle

Risco	Manifestação no estudo	Controle recomendado
Alucinação	Referências, metadados, números ou explicações sem suporte.	Conferir toda afirmação factual em fontes e arquivos do projeto.
Viés combinado	Assimetrias do modelo reforçam vieses linguísticos, geográficos e disciplinares das bases.	Comparar cobertura, incluir fontes complementares e declarar ausências.

Risco	Manifestação no estudo	Controle recomendado
Perda de rastreabilidade	Classificações e textos são incorporados sem registro da interação.	Versionar *prompts*, entradas, saídas, decisões e validações.
Falsa neutralidade	Mapas, rankings e *dashboards* parecem objetivos apesar das escolhas analíticas.	Expor parâmetros, limiares, unidades, fontes e interpretações alternativas.
Aceleração acrítica	A velocidade de produção reduz o tempo de reflexão e revisão.	Definir pontos de parada, dupla revisão e critérios mínimos de qualidade.

Fonte: Elaborado pelo autor (2026).

22.6 Perspectivas futuras

As perspectivas futuras não devem ser lidas como previsões determinísticas, mas como direções de desenvolvimento que exigem investigação, governança e validação. O Quadro 22.4 reúne oito frentes destacadas ao longo da conclusão.

A integração com ferramentas consolidadas deverá preservar a distinção entre cálculo e interpretação. Bibliometrix, CiteSpace, SciMAT e VOSviewer exemplificam diferentes abordagens para análise e visualização bibliométrica (ARIA; CUCCURULLO, 2017; CHEN, 2006; COBO *et al.*, 2012; VAN ECK; WALTMAN, 2010). Módulos de IA podem explicar parâmetros, sugerir verificações e oferecer rótulos candidatos, mas não devem ocultar os procedimentos que geraram as redes e os mapas.

Repositórios de *prompts* e benchmarks poderão aproximar a engenharia de *prompts* de uma prática cumulativa, permitindo comparar versões, tarefas e modelos. O valor desses recursos dependerá da preservação do contexto, dos dados de entrada, dos critérios de validação e dos casos de falha. Bibliotecas sem documentação podem acelerar a repetição de erros; bibliotecas auditáveis podem elevar o padrão metodológico da área.

Quadro 22.4 - Direções futuras para a cientometria com IA

Direção	Possibilidade	Desafio metodológico
Grafos de conhecimento	Vincular respostas a entidades, registros, citações e identificadores verificáveis.	Garantir proveniência, atualização e explicabilidade das consultas.
Modelos científicos especializados	Aprimorar classificação, extração de métodos e análise semântica.	Avaliar desempenho por domínio e reconhecer erros específicos.
Fluxos de agentes	Executar sequências controladas de busca, limpeza, análise e relatório.	Definir permissões, logs, limites de ação e interrupção humana.
Benchmarks cientométricos	Comparar modelos em desambiguação, rotulagem, classificação e auditoria.	Construir conjuntos de referência públicos e critérios de qualidade.
Repositórios de prompts	Compartilhar versões, exemplos, riscos e validações por tarefa.	Evitar receitas universais e manter contexto de aplicação.
Integração com softwares	Incorporar apoio interpretativo em ferramentas de mapeamento e análise.	Apresentar alternativas e incertezas, não rótulos automáticos definitivos.

Direção	Possibilidade	Desafio metodológico
Visualização interativa	Permitir consultas em linguagem natural sobre redes e indicadores.	Exibir filtros, métricas, fontes e limites junto às respostas.
Formação metodológica	Integrar IA, métricas, busca, ética e auditoria à capacitação.	Ensinar avaliação crítica, e não mera memorização de comandos.

Fonte: Elaborado pelo autor (2026).

22.7 Princípios para uma cientometria aumentada por IA

A síntese da obra permite formular dez princípios operacionais. Eles devem orientar tanto estudos acadêmicos quanto diagnósticos institucionais e relatórios de política científica.

Quadro 22.5 - Princípios operacionais para estudos cientométricos com IA

Princípio	Exigência operacional	Função do prompt
Primazia da pergunta	Selecionar métodos e ferramentas somente após delimitar o problema.	Explicitar objetivo, unidade de análise e produto esperado.
Suficiência dos dados	Avaliar cobertura, completude e adequação do *corpus*.	Indicar dados ausentes e impedir preenchimentos inventados.
Rastreabilidade	Registrar fontes, filtros, parâmetros, versões e decisões.	Solicitar entradas, saídas, justificativas e grau de confiança.
Separação analítica	Distinguir dado, inferência, hipótese e recomendação.	Estruturar a resposta por níveis de evidência.
Triangulação	Confrontar métricas, conteúdo, literatura e conhecimento especializado.	Pedir convergências, divergências e evidências adicionais.
Proporcionalidade	Limitar a força das conclusões à qualidade das evidências.	Sinalizar extrapolações e causalidade não demonstrada.
Validação humana	Rever amostras, ambiguidades, rótulos e conclusões.	Marcar casos incertos e indicar procedimentos de conferência.
Transparência	Declarar como, onde e com que finalidade a IA foi usada.	Produzir ficha de documentação e seção de limitações.
Responsabilidade	Manter decisões e consequências sob responsabilidade da equipe.	Evitar linguagem de autoridade ou decisão automática.
Abertura à revisão	Atualizar *prompts*, categorias e interpretações diante de novos dados.	Registrar versões, críticas e mudanças justificadas.

Fonte: Elaborado pelo autor (2026).

22.8 Implicações para pesquisa, gestão e política científica

Para pesquisadores, a IA amplia a capacidade de explorar grandes *corpora*, organizar literatura e testar interpretações, mas exige competências de curadoria, avaliação de modelos e auditoria. Formular bons *prompts* passa a integrar o repertório metodológico, sem substituir o domínio substantivo do campo.

Para bibliotecários e profissionais da informação, surgem funções relacionadas à construção de vocabulários, validação de metadados, documentação de estratégias, curadoria de *prompts* e treinamento de equipes. A automação não reduz a importância dessa expertise; ao contrário, aumenta a necessidade de mediação qualificada entre sistemas, dados e usuários.

Para gestores e formuladores de políticas, a cientometria com IA pode acelerar diagnósticos sobre capacidades, colaboração, internacionalização e temas emergentes. A facilidade de gerar relatórios, contudo, não deve converter métricas em comandos automáticos de decisão. Evidências cientométricas devem integrar processos deliberativos que considerem missão pública, desigualdades regionais, pluralidade disciplinar e impactos sociais.

Para programas de pós-graduação e instituições avaliadoras, permanece a advertência de que métricas apoiam diagnósticos, mas não substituem avaliação qualitativa. O uso de indicadores deve ser multidimensional, transparente e sensível às diferenças de área e de missão, conforme preconizam as abordagens responsáveis de avaliação da pesquisa (HICKS *et al.*, 2015; DONTHU *et al.*, 2021).

22.9 Agenda de pesquisa

A consolidação da cientometria com IA exige uma agenda própria, capaz de avaliar não apenas o desempenho técnico dos modelos, mas também os efeitos epistemológicos e institucionais de sua incorporação.

Quadro 22.6 - Eixos prioritários da agenda de pesquisa

Eixo	Questões prioritárias
Metodológico	Comparar modelos e estratégias de *prompt* em normalização, rotulagem, classificação, interpretação e auditoria, com conjuntos de referência e revisão humana.
Epistemológico	Investigar como modelos influenciam perguntas, categorias, lacunas e narrativas sobre campos científicos.
Técnico	Desenvolver sistemas que vinculem respostas a bases verificáveis, identificadores, scripts e visualizações reprodutíveis.
Educacional	Formar pesquisadores para formular instruções, avaliar respostas, reconhecer vieses e documentar fluxos.
Normativo	Definir padrões mínimos para declarar modelos, versões, *prompts*, dados de entrada, validações e intervenções humanas.

Fonte: Elaborado pelo autor (2026).

Essa agenda deverá combinar ciência da informação, cientometria, sociologia da ciência, estudos sociais da tecnologia, ciência de dados e ética em IA. A comparação entre modelos precisa considerar não apenas acurácia, mas estabilidade, rastreabilidade, sensibilidade a idioma e domínio, capacidade de expressar incerteza e impacto das respostas sobre decisões reais.

22.10 Considerações finais e fechamento prospectivo

A conclusão da obra é orientada por uma dupla exigência: ampliar capacidades analíticas e preservar responsabilidade científica. A IA permite lidar com grandes volumes de literatura, produzir classificações preliminares, sintetizar padrões, apoiar visualizações, revisar argumentos e adaptar a comunicação. A engenharia de *prompts* fornece um meio para dirigir e documentar essas capacidades, mas nenhuma delas elimina a necessidade de teoria, método, curadoria, validação e julgamento humano.

O futuro desejável situa-se entre o entusiasmo acrítico e a rejeição absoluta. A primeira postura transforma a IA em solução automática; a segunda ignora ganhos reais de exploração, eficiência e comunicação. A posição defendida neste livro é metodologicamente responsável: utilizar a IA como instrumento de ampliação cognitiva, com cada etapa registrada, revista e contextualizada.

A cientometria assistida por IA deverá evoluir para ecossistemas nos quais dados, modelos, *prompts*, visualizações e decisões humanas estejam conectados em ambientes transparentes. A maturidade desses ecossistemas não será medida apenas pela automação, mas pela capacidade de combinar inovação com cautela, poder computacional com sensibilidade epistemológica e rapidez com mecanismos efetivos de revisão.

A engenharia de *prompts* pode ser entendida como linguagem de mediação científica: traduz objetivos em instruções operacionais e torna explícitos os limites dessas instruções. Ao registrar versões, parâmetros, entradas, saídas e decisões, a equipe cria memória metodológica, favorece a repetição crítica e permite revisar resultados diante de novas evidências.

Fechamento

A cientometria com IA será tão científica quanto forem verificáveis seus dados, transparentes suas mediações, rigorosas suas validações e responsáveis suas interpretações. Modelos podem sugerir, organizar, comparar e redigir; a comunidade humana continua responsável por definir o problema, avaliar as evidências e responder pelas consequências.

Referências Bibliográficas

ABRAMO, G.; D'ANGELO, C. A. Evaluating research: from informed peer review to bibliometrics. **Scientometrics**, v. 87, n. 3, p. 499-514, 2011. DOI: https://doi.org/10.1007/s11192-011-0352-7.

ADAM, G. P. *et al.* Literature search sandbox: a large language model that generates search queries for systematic reviews. **JAMIA Open**, v. 7, n. 3, ooae098, 2024. DOI: https://doi.org/10.1093/jamiaopen/ooae098.

AKSNES, D. W.; LANGFELDT, L.; WOUTERS, P. Citations, citation indicators, and research quality: an overview of basic concepts and theories. **SAGE Open**, v. 9, n. 1, 2019. DOI: https://doi.org/10.1177/2158244019829575.

ALPERIN, J. P. *et al.* An analysis of the suitability of OpenAlex for bibliometric analyses. **arXiv**, arXiv:2404.17663, 2024. Disponível em: https://arxiv.org/abs/2404.17663. Acesso em: 25 jul. 2026.

AMERICAN SOCIETY FOR CELL BIOLOGY. **San Francisco Declaration on Research Assessment (DORA)**. San Francisco, 2012. Disponível em: https://sfdora.org/read/. Acesso em: 25 jul. 2026.

ARCHAMBAULT, É. *et al.* Comparing bibliometric statistics obtained from the Web of Science and Scopus. **Journal of the American Society for Information Science and Technology**, v. 60, n. 7, p. 1320-1326, 2009. DOI: https://doi.org/10.1002/asi.21062.

ARIA, M.; CUCCURULLO, C. bibliometrix: an R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959-975, 2017. DOI: https://doi.org/10.1016/j.joi.2017.08.007.

BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. **Proceedings of the International AAAI Conference on Web and Social Media**, v. 3, n. 1, p. 361-362, 2009. DOI: https://doi.org/10.1609/icwsm.v3i1.13937.

BELTAGY, I.; LO, K.; COHAN, A. SciBERT: a pretrained language model for scientific text. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING (EMNLP-IJCNLP), 2019. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2019. p. 3615-3620. DOI: https://doi.org/10.18653/v1/D19-1371.

BENDER, E. M. *et al.* On the dangers of stochastic parrots: can language models be too big? In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2021. **Proceedings** [...]. New York: ACM, 2021. p. 610-623. DOI: https://doi.org/10.1145/3442188.3445922.

BIRKLE, C. *et al.* Web of Science as a data source for research on scientific and scholarly activity. **Quantitative Science Studies**, v. 1, n. 1, p. 363-376, 2020. DOI: https://doi.org/10.1162/qss_a_00018.

BLEI, D. M.; LAFFERTY, J. D. Dynamic topic models. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 23., 2006, Pittsburgh. **Proceedings** [...]. New York: ACM, 2006. p. 113-120. DOI: https://doi.org/10.1145/1143844.1143859.

BLEI, D. M.; NG, A. Y.; JORDAN, M. I. Latent Dirichlet Allocation. **Journal of** Machine Learning **Research**, v. 3, p. 993-1022, 2003. Disponível em: https://www.jmlr.org/papers/v3/blei03a.html. Acesso em: 25 jul. 2026.

BOLAÑOS, F. *et al.* Artificial intelligence for literature reviews: opportunities and challenges. Artificial Intelligence **Review**, v. 57, art. 259, 2024. DOI: https://doi.org/10.1007/s10462-024-10902-3.

BÖRNER, K.; CHEN, C.; BOYACK, K. W. Visualizing knowledge domains. **Annual Review of Information Science and Technology**, v. 37, n. 1, p. 179-255, 2003. DOI: https://doi.org/10.1002/aris.1440370106.

BORNMANN, L. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. **Journal of Informetrics**, v. 8, n. 4, p. 895-903, 2014. DOI: https://doi.org/10.1016/j.joi.2014.09.005.

BORNMANN, L. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics. **Scientometrics**, v. 103, p. 1123-1144, 2015. DOI: https://doi.org/10.1007/s11192-015-1565-y.

BORNMANN, L.; DANIEL, H.-D. What do citation counts measure? A review of studies on citing behavior. **Journal of Documentation**, v. 64, n. 1, p. 45-80, 2008. DOI: https://doi.org/10.1108/00220410810844150.

BORNMANN, L.; HAUNSCHILD, R. Alternative article-level metrics: the use of alternative metrics in research evaluation. **EMBO Reports**, v. 19, n. 12, e47260, 2018. DOI: https://doi.org/10.15252/embr.201847260.

BORNMANN, L.; HAUNSCHILD, R.; MUTZ, R. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. **Humanities and Social Sciences Communications**, v. 8, art. 224, 2021. DOI: https://doi.org/10.1057/s41599-021-00903-w.

BORNMANN, L.; LEYDESDORFF, L. Scientometrics in a changing research landscape. **EMBO Reports**, v. 15, n. 12, p. 1228-1232, 2014. DOI: https://doi.org/10.15252/embr.201439608.

BORNMANN, L.; LEYDESDORFF, L. Skewness of citation impact data and covariates of citation distributions: a large-scale empirical analysis based on Web of Science data. **Journal of Informetrics**, v. 11, n. 1, p. 164-175, 2017. DOI: https://doi.org/10.1016/j.joi.2016.12.001.

BORNMANN, L.; MARX, W. The proposal of a broadening of perspective in evaluative bibliometrics by complementing the times cited with a cited reference analysis. **Journal of Informetrics**, v. 7, n. 1, p. 84-88, 2013. DOI: https://doi.org/10.1016/j.joi.2012.09.003.

BORNMANN, L.; MARX, W. How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons. **Scientometrics**, v. 98, p. 211-219, 2014. DOI: https://doi.org/10.1007/s11192-013-1020-x.

BORNMANN, L.; MUTZ, R. Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references. **Journal of the Association for Information Science and Technology**, v. 66, n. 11, p. 2215-2222, 2015. DOI: https://doi.org/10.1002/asi.23329.

BOSTOCK, M.; OGIEVETSKY, V.; HEER, J. D3: data-driven documents. **IEEE Transactions on Visualization and Computer Graphics**, v. 17, n. 12, p. 2301-2309, 2011. DOI: https://doi.org/10.1109/TVCG.2011.185.

BOYACK, K. W.; KLAVANS, R. Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? **Journal of the American Society for Information Science and Technology**, v. 61, n. 12, p. 2389-2404, 2010. DOI: https://doi.org/10.1002/asi.21419.

BRADFORD, S. C. Sources of information on specific subjects. **Engineering**, v. 137, p. 85-86, 1934. Reimpresso em: Journal of Information Science, v. 10, n. 4, p. 176-180, 1985. DOI: https://doi.org/10.1177/016555158501000407.

BRAMER, W. M. *et al.* A systematic approach to searching: an efficient and complete method to develop literature searches. **Journal of the Medical Library Association**, v. 106, n. 4, p. 531-541, 2018. DOI: https://doi.org/10.5195/jmla.2018.283.

BROWN, T. B. *et al.* Language models are few-shot learners. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 33., 2020. **Proceedings** [...]. [S. l.]: NeurIPS, 2020. p. 1877-1901. Disponível em: https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html. Acesso em: 25 jul. 2026.

CACHOLA, I. *et al.* TLDR: extreme summarization of scientific documents. In: FINDINGS OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS: EMNLP 2020, 2020. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2020. p. 4766-4777. DOI: https://doi.org/10.18653/v1/2020.findings-emnlp.428.

CALLON, M.; COURTIAL, J.-P.; LAVILLE, F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry. **Scientometrics**, v. 22, p. 155-205, 1991. DOI: https://doi.org/10.1007/BF02019280.

CASSI, L. *et al.* Analysing institutions interdisciplinarity by extensive use of Rao-Stirling diversity index. **PLOS ONE**, v. 12, n. 1, e0170296, 2017. DOI: https://doi.org/10.1371/journal.pone.0170296.

CHAVARRO, D.; RAFOLS, I.; TANG, P. To what extent is inclusion in the Web of Science an indicator of journal quality? **Research Evaluation**, v. 27, n. 2, p. 106-118, 2018. DOI: https://doi.org/10.1093/reseval/rvy001.

CHELLI, M. *et al.* Hallucination rates and reference accuracy of ChatGPT and Bard for systematic reviews: comparative analysis. **Journal of Medical Internet Research**, v. 26, e53164, 2024. DOI: https://doi.org/10.2196/53164.

CHEN, B. *et al.* Unleashing the potential of prompt engineering for large language models. **Patterns**, v. 6, n. 6, art. 101260, 2025. DOI: https://doi.org/10.1016/j.patter.2025.101260.

CHEN, C. Searching for intellectual turning points: progressive knowledge domain visualization. **Proceedings of the National Academy of Sciences**, v. 101, supl. 1, p. 5303-5310, 2004. DOI: https://doi.org/10.1073/pnas.0307513100.

CHEN, C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. **Journal of the American Society for Information Science and Technology**, v. 57, n. 3, p. 359-377, 2006. DOI: https://doi.org/10.1002/asi.20317.

CHEN, C.; IBEKWE-SANJUAN, F.; HOU, J. The structure and dynamics of co-citation clusters: a multiple-perspective co-citation analysis. **Journal of the American Society for Information Science and Technology**, v. 61, n. 7, p. 1386-1409, 2010. DOI: https://doi.org/10.1002/asi.21309.

CHEN, C.; SONG, M. Visualizing a field of research: a methodology of systematic scientometric reviews. **PLOS ONE**, v. 14, n. 10, e0223994, 2019. DOI: https://doi.org/10.1371/journal.pone.0223994.

CLEVELAND, W. S.; MCGILL, R. Graphical perception: theory, experimentation, and application to the development of graphical methods. **Journal of the American Statistical Association**, v. 79, n. 387, p. 531-554, 1984. DOI: https://doi.org/10.1080/01621459.1984.10478080.

CLEVELAND, W. S.; MCGILL, R. Graphical perception and graphical methods for analyzing scientific data. **Science**, v. 229, n. 4716, p. 828-833, 1985. DOI: https://doi.org/10.1126/science.229.4716.828.

COBO, M. J. *et al.* An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the Fuzzy Sets Theory field. **Journal of Informetrics**, v. 5, n. 1, p. 146-166, 2011. DOI: https://doi.org/10.1016/j.joi.2010.10.002.

COBO, M. J. *et al.* Science mapping software tools: review, analysis, and cooperative study among tools. **Journal of the American Society for Information Science and Technology**, v. 62, n. 7, p. 1382-1402, 2011. DOI: https://doi.org/10.1002/asi.21525.

COBO, M. J. *et al.* SciMAT: a new science mapping analysis software tool. **Journal of the American Society for Information Science and Technology**, v. 63, n. 8, p. 1609-1630, 2012. DOI: https://doi.org/10.1002/asi.22688.

COHAN, A. *et al.* SPECTER: document-level representation learning using citation-informed transformers. In: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 58., 2020. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2020. p. 2270-2282. DOI: https://doi.org/10.18653/v1/2020.acl-main.207.

COHEN, A. M. *et al.* Reducing workload in systematic review preparation using automated citation classification. **Journal of the American Medical Informatics Association**, v. 13, n. 2, p. 206-219, 2006. DOI: https://doi.org/10.1197/jamia.M1929.

COMINS, J. A.; LEYDESDORFF, L. RPYS i/o: a web-based tool for the historiography and visualization of citation classics, sleeping beauties and research fronts. **Scientometrics**, v. 107, p. 1509-1517, 2016. DOI: https://doi.org/10.1007/s11192-016-1903-z.

COZZENS, S. E. What do citations count? The rhetoric-first model. **Scientometrics**, v. 15, n. 5-6, p. 437-447, 1989. DOI: https://doi.org/10.1007/BF02017064.

CULBERT, J. H. *et al.* Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. **Scientometrics**, v. 130, n. 4, p. 2475-2492, 2025. DOI: https://doi.org/10.1007/s11192-025-05293-3.

DECLARATION ON RESEARCH ASSESSMENT. The San Francisco Declaration on Research Assessment. **Journal of Experimental Biology**, v. 216, n. 12, p. 2163-2164, 2013. DOI: https://doi.org/10.1242/jeb.093020.

DEERWESTER, S. *et al.* Indexing by latent semantic analysis. **Journal of the American Society for Information Science**, v. 41, n. 6, p. 391-407, 1990. DOI: https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9.

DELGADO-QUIRÓS, L.; ORTEGA, J. L. Completeness degree of publication metadata in eight free-access scholarly databases. **Quantitative Science Studies**, v. 5, n. 1, p. 31-49, 2024. DOI: https://doi.org/10.1162/qss_a_00286.

DE RIJCKE, S. *et al.* Evaluation practices and effects of indicator use: a literature review. **Research Evaluation**, v. 25, n. 2, p. 161-169, 2016. DOI: https://doi.org/10.1093/reseval/rvv038.

DEVLIN, J. *et al.* BERT: pre-training of deep bidirectional transformers for language understanding. In: CONFERENCE OF THE NORTH AMERICAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (NAACL-HLT), 2019. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2019. p. 4171-4186. DOI: https://doi.org/10.18653/v1/N19-1423.

DIENG, A. B.; RUIZ, F. J. R.; BLEI, D. M. The Dynamic Embedded Topic Model. **arXiv**, arXiv:1907.05545, 2019. Disponível em: https://arxiv.org/abs/1907.05545. Acesso em: 25 jul. 2026.

DONTHU, N. *et al.* How to conduct a bibliometric analysis: an overview and guidelines. **Journal of Business Research**, v. 133, p. 285-296, 2021. DOI: https://doi.org/10.1016/j.jbusres.2021.04.070.

DURAN-SILVA, N. *et al.* AffilGood: building reliable institution name disambiguation tools to improve scientific literature analysis. In: WORKSHOP ON SCHOLARLY DOCUMENT PROCESSING, 4., 2024, Bangkok. **Proceedings** [...]. Bangkok: Association for Computational Linguistics, 2024. p. 135-144. Disponível em: https://aclanthology.org/2024.sdp-1.13/. Acesso em: 25 jul. 2026.

EGGHE, L. Theory and practise of the g-index. **Scientometrics**, v. 69, p. 131-152, 2006. DOI: https://doi.org/10.1007/s11192-006-0144-7.

ELMAGARMID, A. K.; IPEIROTIS, P. G.; VERYKIOS, V. S. Duplicate record detection: a survey. **IEEE Transactions on Knowledge and Data Engineering**, v. 19, n. 1, p. 1-16, 2007. DOI: https://doi.org/10.1109/TKDE.2007.250581.

ERDT, M. *et al.* Altmetrics: an analysis of the state-of-the-art in measuring research impact on social media. **Scientometrics**, v. 109, p. 1117-1166, 2016. DOI: https://doi.org/10.1007/s11192-016-2077-0.

FANELLI, D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. **PLOS ONE**, v. 4, n. 5, e5738, 2009. DOI: https://doi.org/10.1371/journal.pone.0005738.

FÄRBER, M. *et al.* SemOpenAlex: the scientific landscape in 26 billion RDF triples. **arXiv**, arXiv:2308.03671, 2023. Disponível em: https://arxiv.org/abs/2308.03671. Acesso em: 25 jul. 2026.

FARQUHAR, S. *et al.* Detecting hallucinations in large language models using semantic entropy. **Nature**, v. 630, p. 625-630, 2024. DOI: https://doi.org/10.1038/s41586-024-07421-0.

FIORINI, N. *et al.* Best Match: new relevance search for PubMed. **PLOS Biology**, v. 16, n. 8, e2005343, 2018. DOI: https://doi.org/10.1371/journal.pbio.2005343.

FIRE, M.; GUESTRIN, C. Over-optimization of academic publishing metrics: observing Goodhart's law in action. **GigaScience**, v. 8, n. 6, giz053, 2019. DOI: https://doi.org/10.1093/gigascience/giz053.

FLORIDI, L.; COWLS, J. A unified framework of five principles for AI in society. **Harvard Data Science Review**, v. 1, n. 1, 2019. DOI: https://doi.org/10.1162/99608f92.8cd550d1.

FORTUNATO, S. *et al.* Science of science. **Science**, v. 359, n. 6379, eaao0185, 2018. DOI: https://doi.org/10.1126/science.aao0185.

FRIENDLY, M. Milestones in the history of thematic cartography, statistical graphics, and data visualization. **Statistical Science**, v. 23, n. 4, p. 502-535, 2008. DOI: https://doi.org/10.1214/08-STS268.

GAO, C. A. *et al.* Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. **npj Digital Medicine**, v. 6, art. 75, 2023. DOI: https://doi.org/10.1038/s41746-023-00819-6.

GARFIELD, E. Citation indexes for science: a new dimension in documentation through association of ideas. **Science**, v. 122, n. 3159, p. 108-111, 1955. DOI: https://doi.org/10.1126/science.122.3159.108.

GARFIELD, E. Citation analysis as a tool in journal evaluation. **Science**, v. 178, n. 4060, p. 471-479, 1972. DOI: https://doi.org/10.1126/science.178.4060.471.

GARFIELD, E. The history and meaning of the journal impact factor. **JAMA**, v. 295, n. 1, p. 90-93, 2006. DOI: https://doi.org/10.1001/jama.295.1.90.

GIPP, B.; BEEL, J. Citation Proximity Analysis (CPA): a new approach for identifying related work based on co-citation analysis. **arXiv**, arXiv:0908.3535, 2009. Disponível em: https://arxiv.org/abs/0908.3535. Acesso em: 25 jul. 2026.

GLÄNZEL, W. National characteristics in international scientific co-authorship relations. **Scientometrics**, v. 51, p. 69-115, 2001. DOI: https://doi.org/10.1023/A:1010512628145.

GRAVEL, J.; D'AMOURS-GRAVEL, M.; OSMANLLIU, E. Learning to fake it: limited responses and fabricated references provided by ChatGPT for medical questions. **Mayo Clinic Proceedings: Digital Health**, v. 1, n. 3, p. 226-234, 2023. DOI: https://doi.org/10.1016/j.mcpdig.2023.05.004.

GRIFFITHS, T. L.; STEYVERS, M. Finding scientific topics. **Proceedings of the National Academy of Sciences**, v. 101, supl. 1, p. 5228-5235, 2004. DOI: https://doi.org/10.1073/pnas.0307752101.

GUERRERO-BOTE, V. P. *et al.* Comparative analysis of the bibliographic data sources Dimensions and Scopus: an approach at the country and institutional levels. **Frontiers in Research Metrics and Analytics**, v. 5, 593494, 2021. DOI: https://doi.org/10.3389/frma.2020.593494.

GUSENBAUER, M.; HADDAWAY, N. R. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. **Research Synthesis Methods**, v. 11, n. 2, p. 181-217, 2020. DOI: https://doi.org/10.1002/jrsm.1378.

HAAS, P. M. Introduction: epistemic communities and international policy coordination. **International Organization**, v. 46, n. 1, p. 1-35, 1992. DOI: https://doi.org/10.1017/S0020818300001442.

HAGHANI, M. What makes an informative and publication-worthy scientometric analysis of literature: a guide for authors, reviewers and editors. **Transportation Research Interdisciplinary Perspectives**, v. 22, 100956, 2023. DOI: https://doi.org/10.1016/j.trip.2023.100956.

HAUSTEIN, S. *et al.* Coverage and adoption of altmetrics sources in the bibliometric community. **Scientometrics**, v. 101, p. 1145-1163, 2014. DOI: https://doi.org/10.1007/s11192-013-1221-3.

HAUSTEIN, S. *et al.* Tweeting biomedicine: an analysis of tweets and citations in the biomedical literature. **Journal of the Association for Information Science and Technology**, v. 65, n. 4, p. 656-669, 2014. DOI: https://doi.org/10.1002/asi.23101.

HEER, J.; BOSTOCK, M.; OGIEVETSKY, V. A tour through the visualization zoo. **Communications of the ACM**, v. 53, n. 6, p. 59-67, 2010. DOI: https://doi.org/10.1145/1743546.1743567.

HEIBI, I.; PERONI, S.; SHOTTON, D. Software review: COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. **Scientometrics**, v. 121, p. 1213-1228, 2019. DOI: https://doi.org/10.1007/s11192-019-03217-6.

HENDRICKS, G. *et al.* Crossref: the sustainable source of community-owned scholarly metadata. **Quantitative Science Studies**, v. 1, n. 1, p. 414-427, 2020. DOI: https://doi.org/10.1162/qss_a_00022.

HICKS, D. *et al.* Bibliometrics: the Leiden Manifesto for research metrics. **Nature**, v. 520, p. 429-431, 2015. DOI: https://doi.org/10.1038/520429a.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National Academy of Sciences**, v. 102, n. 46, p. 16569-16572, 2005. DOI: https://doi.org/10.1073/pnas.0507655102.

HOOK, D. W.; PORTER, S. J.; HERZOG, C. Dimensions: building context for search and evaluation. **Frontiers in Research Metrics and Analytics**, v. 3, art. 23, 2018. DOI: https://doi.org/10.3389/frma.2018.00023.

HUANG, L. *et al.* A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. **ACM Transactions on Information Systems**, v. 43, n. 2, p. 1-55, 2025. DOI: https://doi.org/10.1145/3703155.

IOANNIDIS, J. P. A. Why most published research findings are false. **PLOS Medicine**, v. 2, n. 8, e124, 2005. DOI: https://doi.org/10.1371/journal.pmed.0020124.

JI, Z. *et al.* Survey of hallucination in natural language generation. **ACM Computing Surveys**, v. 55, n. 12, art. 248, p. 1-38, 2023. DOI: https://doi.org/10.1145/3571730.

JOBIN, A.; IENCA, M.; VAYENA, E. The global landscape of AI ethics guidelines. **Nature Machine Intelligence,** v. 1, p. 389-399, 2019. DOI: https://doi.org/10.1038/s42256-019-0088-2.

KASNECI, E. *et al.* ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences**, v. 103, art. 102274, 2023. DOI: https://doi.org/10.1016/j.lindif.2023.102274.

KATZ, J. S.; MARTIN, B. R. What is research collaboration? **Research Policy**, v. 26, n. 1, p. 1-18, 1997. DOI: https://doi.org/10.1016/S0048-7333(96)00917-1.

KAYSER, V.; BLIND, K. Extending the knowledge base of foresight: the contribution of text mining. **Technological Forecasting and Social Change**, v. 116, p. 208-215, 2017. DOI: https://doi.org/10.1016/j.techfore.2016.07.017.

KEIM, D. *et al.* Visual analytics: definition, process, and challenges. In: KERREN, A. *et al.* (ed.). **Information visualization**: human-centered issues and perspectives. Berlin: Springer, 2008. p. 154-175. DOI: https://doi.org/10.1007/978-3-540-70956-5_7.

KESSLER, M. M. Bibliographic coupling between scientific papers. **American Documentation**, v. 14, n. 1, p. 10-25, 1963. DOI: https://doi.org/10.1002/asi.5090140103.

KLAVANS, R.; BOYACK, K. W. Research portfolio analysis and topic prominence. **Journal of Informetrics**, v. 11, n. 4, p. 1158-1174, 2017. DOI: https://doi.org/10.1016/j.joi.2017.10.002.

KLEINBERG, J. Bursty and hierarchical structure in streams. **Data Mining and Knowledge Discovery**, v. 7, p. 373-397, 2003. DOI: https://doi.org/10.1023/A:1024940629314.

KOFFEL, J. B.; RETHLEFSEN, M. L. Reproducibility of search strategies is poor in systematic reviews published in high-impact pediatrics, cardiology and surgery journals: a cross-sectional study. **PLOS ONE**, v. 11, n. 9, e0163309, 2016. DOI: https://doi.org/10.1371/journal.pone.0163309.

KOJIMA, T. *et al.* Large language models are zero-shot reasoners. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 35., 2022. **Proceedings** [...]. [S. l.]: NeurIPS, 2022. p. 22199-22213. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2022/hash/8bb0d291acd4acf06ef112099c16f326-Abstract-Conference.html. Acesso em: 25 jul. 2026.

LARIVIÈRE, V. *et al.* Bibliometrics: global gender disparities in science. **Nature**, v. 504, p. 211-213, 2013. DOI: https://doi.org/10.1038/504211a.

LARIVIÈRE, V.; GINGRAS, Y. The impact factor's Matthew Effect: a natural experiment in bibliometrics. **Journal of the American Society for Information Science and Technology**, v. 61, n. 2, p. 424-427, 2010. DOI: https://doi.org/10.1002/asi.21232.

LEE, J. *et al.* BioBERT: a pre-trained biomedical language representation model for biomedical text mining. **Bioinformatics**, v. 36, n. 4, p. 1234-1240, 2020. DOI: https://doi.org/10.1093/bioinformatics/btz682.

LEWIS, P. *et al.* Retrieval-augmented generation for knowledge-intensive NLP tasks. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 33., 2020. **Proceedings** [...]. [S. l.]: NeurIPS, 2020. p. 9459-9474. Disponível em: https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html. Acesso em: 25 jul. 2026.

LEYDESDORFF, L. *et al.* Citations: indicators of quality? The impact fallacy. **Frontiers in Research Metrics and Analytics**, v. 1, art. 1, 2016. DOI: https://doi.org/10.3389/frma.2016.00001.

LEYDESDORFF, L.; MILOJEVIĆ, S. Scientometrics. In: WRIGHT, J. D. (ed.). **International Encyclopedia of the Social & Behavioral Sciences**. 2. ed. Amsterdam: Elsevier, 2015. DOI: https://doi.org/10.1016/B978-0-08-097086-8.85030-8.

LEYDESDORFF, L.; RAFOLS, I. A global map of science based on the ISI subject categories. **Journal of the American Society for Information Science and Technology**, v. 60, n. 2, p. 348-362, 2009. DOI: https://doi.org/10.1002/asi.20967.

LEYDESDORFF, L.; RAFOLS, I. Interactive overlays: a new method for generating global journal maps from Web of Science data. **Journal of Informetrics**, v. 6, n. 2, p. 318-332, 2012. DOI: https://doi.org/10.1016/j.joi.2011.11.003.

LEYDESDORFF, L.; WAGNER, C. S.; BORNMANN, L. Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. **Journal of Informetrics**, v. 13, n. 1, p. 255-269, 2019. DOI: https://doi.org/10.1016/j.joi.2018.12.006.

LIEBRENZ, M. *et al.* Generating scholarly content with ChatGPT: ethical challenges for medical publishing. **The Lancet Digital Health**, v. 5, n. 3, p. e105-e106, 2023. DOI: https://doi.org/10.1016/S2589-7500(23)00019-5.

LIM, W. M.; KUMAR, S.; DONTHU, N. How to combine and clean bibliometric data and use bibliometric tools synergistically: guidelines using metaverse research. **Journal of Business Research**, v. 182, 114760, 2024. DOI: https://doi.org/10.1016/j.jbusres.2024.114760.

LIN, C.-Y. ROUGE: a package for automatic evaluation of summaries. In: TEXT SUMMARIZATION BRANCHES OUT: ACL WORKSHOP, 2004, Barcelona. Proceedings [...]. Barcelona: Association for Computational Linguistics, 2004. p. 74-81. Disponível em: https://aclanthology.org/W04-1013/. Acesso em: 25 jul. 2026.

LIN, J.; FENNER, M. An introduction to the Open Researcher and Contributor ID. **Learned Publishing**, v. 26, n. 4, p. 270-272, 2013. DOI: https://doi.org/10.1087/20130404.

LIU, P. *et al.* Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. **ACM Computing Surveys**, v. 55, n. 9, art. 195, p. 1-35, 2023. DOI: https://doi.org/10.1145/3560815.

LOTKA, A. J. The frequency distribution of scientific productivity. **Journal of the Washington Academy of Sciences**, v. 16, n. 12, p. 317-323, 1926. Disponível em: https://www.jstor.org/stable/24529203. Acesso em: 25 jul. 2026.

LUAN, Y. *et al.* Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING (EMNLP), 2018. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2018. p. 3219-3232. DOI: https://doi.org/10.18653/v1/D18-1360.

LUHN, H. P. The automatic creation of literature abstracts. **IBM Journal of Research and Development**, v. 2, n. 2, p. 159-165, 1958. DOI: https://doi.org/10.1147/rd.22.0159.

LUND, B. D.; WANG, T. ChatGPT and a new academic reality: artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. **Journal of the Association for Information Science and Technology**, v. 74, n. 5, p. 570-581, 2023. DOI: https://doi.org/10.1002/asi.24750.

MACFARLANE, A.; RUSSELL-ROSE, T.; SHOKRANEH, F. Search strategy formulation for systematic reviews: issues, challenges and opportunities. **Intelligent Systems with Applications**, v. 15, 200091, 2022. DOI: https://doi.org/10.1016/j.iswa.2022.200091.

MACROBERTS, M. H.; MACROBERTS, B. R. Problems of citation analysis: a critical review. **Journal of the American Society for Information Science**, v. 40, n. 5, p. 342-349, 1989. DOI: https://doi.org/10.1002/(SICI)1097-4571(198909)40:5<342::AID-ASI7>3.0.CO;2-U.

MANE, K. K.; BÖRNER, K. Mapping topics and topic bursts in PNAS. **Proceedings of the National Academy of Sciences**, v. 101, supl. 1, p. 5287-5290, 2004. DOI: https://doi.org/10.1073/pnas.0307626100.

MARSHALL, I. J.; WALLACE, B. C. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. **Systematic Reviews**, v. 8, art. 163, 2019. DOI: https://doi.org/10.1186/s13643-019-1074-9.

MARTÍN-MARTÍN, A. *et al.* Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. **Journal of Informetrics**, v. 12, n. 4, p. 1160-1177, 2018. DOI: https://doi.org/10.1016/j.joi.2018.09.002.

MARTÍN-MARTÍN, A. *et al.* Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. **Scientometrics**, v. 126, p. 871-906, 2021. DOI: https://doi.org/10.1007/s11192-020-03690-4.

MARTINS, P. *et al.* Performance and scalability of data cleaning and preprocessing tools: a benchmark on large real-world datasets. **Data**, v. 10, n. 5, 68, 2025. DOI: https://doi.org/10.3390/data10050068.

MAYNEZ, J. *et al.* On faithfulness and factuality in abstractive summarization. In: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 58., 2020. **Proceedings**

[...]. Stroudsburg: Association for Computational Linguistics, 2020. p. 1906-1919. DOI: https://doi.org/10.18653/v1/2020.acl-main.173.

MCGOWAN, J. *et al.* PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. **Journal of Clinical Epidemiology**, v. 75, p. 40-46, 2016. DOI: https://doi.org/10.1016/j.jclinepi.2016.01.021.

MERTON, R. K. The Matthew Effect in science. **Science**, v. 159, n. 3810, p. 56-63, 1968. DOI: https://doi.org/10.1126/science.159.3810.56.

MIHALCEA, R.; TARAU, P. TextRank: bringing order into text. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING (EMNLP), 2004, Barcelona. **Proceedings** [...]. Barcelona: Association for Computational Linguistics, 2004. p. 404-411. Disponível em: https://aclanthology.org/W04-3252/. Acesso em: 25 jul. 2026.

MILOJEVIĆ, S. Quantifying the cognitive extent of science. **Journal of Informetrics**, v. 9, n. 4, p. 962-973, 2015. DOI: https://doi.org/10.1016/j.joi.2015.10.005.

MINGERS, J.; LEYDESDORFF, L. A review of theory and practice in scientometrics. **European Journal of Operational Research**, v. 246, n. 1, p. 1-19, 2015. DOI: https://doi.org/10.1016/j.ejor.2015.04.002.

MOED, H. F. UK research assessment exercises: informed judgments on research quality or quantity? **Scientometrics**, v. 74, p. 153-161, 2008. DOI: https://doi.org/10.1007/s11192-008-0108-1.

MOED, H. F. New developments in the use of citation analysis in research evaluation. **Archivum Immunologiae et Therapiae Experimentalis**, v. 57, p. 13-18, 2009. DOI: https://doi.org/10.1007/s00005-009-0001-5.

MOED, H. F.; HALEVI, G. Multidimensional assessment of scholarly research impact. **Journal of the Association for Information Science and Technology**, v. 66, n. 10, p. 1988-2002, 2015. DOI: https://doi.org/10.1002/asi.23314.

MOHER, D. *et al.* Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. **PLOS Medicine**, v. 6, n. 7, e1000097, 2009. DOI: https://doi.org/10.1371/journal.pmed.1000097.

MONGEON, P.; PAUL-HUS, A. The journal coverage of Web of Science and Scopus: a comparative analysis. **Scientometrics**, v. 106, p. 213-228, 2016. DOI: https://doi.org/10.1007/s11192-015-1765-5.

MORAL-MUÑOZ, J. A. *et al.* Software tools for conducting bibliometric analysis in science: an up-to-date review. **El Profesional de la Información**, v. 29, n. 1, e290103, 2020. DOI: https://doi.org/10.3145/epi.2020.ene.03.

MÜLLER, M.-C.; REITZ, F.; ROY, N. Data sets for author name disambiguation: an empirical analysis and a new resource. **Scientometrics**, v. 111, p. 1467-1500, 2017. DOI: https://doi.org/10.1007/s11192-017-2363-5.

MULLER, R.; DE RIJCKE, S. Thinking with indicators: exploring the epistemic impacts of academic performance indicators in the life sciences. **Research Evaluation,** v. 26, n. 3, p. 157-168, 2017. DOI: https://doi.org/10.1093/reseval/rvx023.

NARIN, F.; HAMILTON, K. S.; OLIVASTRO, D. The increasing linkage between U.S. technology and public science. **Research Policy**, v. 26, n. 3, p. 317-330, 1997. DOI: https://doi.org/10.1016/S0048-7333(97)00013-9.

NEWMAN, M. E. J. Scientific collaboration networks. I. Network construction and fundamental results. **Physical Review E**, v. 64, art. 016131, 2001. DOI: https://doi.org/10.1103/PhysRevE.64.016131.

NEWMAN, M. E. J. The structure of scientific collaboration networks. **Proceedings of the National Academy of Sciences**, v. 98, n. 2, p. 404-409, 2001. DOI: https://doi.org/10.1073/pnas.021544898.

NEWMAN, M. E. J. Coauthorship networks and patterns of scientific collaboration. **Proceedings of the National Academy of Sciences**, v. 101, supl. 1, p. 5200-5205, 2004. DOI: https://doi.org/10.1073/pnas.0307545100.

NEWMAN, M. E. J. Modularity and community structure in networks. **Proceedings of the National Academy of Sciences**, v. 103, n. 23, p. 8577-8582, 2006. DOI: https://doi.org/10.1073/pnas.0601602103.

NG, J. Y. *et al.* Guidance for the reporting of bibliometric analyses: a scoping review. **Quantitative Science Studies**, v. 6, p. 988-1001, 2025. DOI: https://doi.org/10.1162/qss.a.12.

NOWAKOWSKA, M. A comprehensive approach to preprocessing data for bibliometric analysis. **Scientometrics**, v. 130, n. 9, p. 5191-5225, 2025. DOI: https://doi.org/10.1007/s11192-025-05415-x.

NYE, B. *et al.* A corpus with multi-level annotations of patients, interventions and outcomes to support language processing for medical literature. In: WORKSHOP ON BIOMEDICAL NATURAL LANGUAGE PROCESSING, 2018, Melbourne. **Proceedings** [...]. Melbourne: Association for Computational Linguistics, 2018. p. 197-207. Disponível em: https://aclanthology.org/W18-5612/. Acesso em: 25 jul. 2026.

O'MARA-EVES, A. *et al.* Using text mining for study identification in systematic reviews: a systematic review of current approaches. **Systematic Reviews**, v. 4, art. 5, 2015. DOI: https://doi.org/10.1186/2046-4053-4-5.

OUYANG, L. *et al.* Training language models to follow instructions with human feedback. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 35., 2022. **Proceedings** [...]. [S. l.]: NeurIPS, 2022. p. 27730-27744. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2022/hash/b1efde53be364a73914f58805a001731-Abstract-Conference.html. Acesso em: 25 jul. 2026.

PAGE, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, n. 71, 2021. DOI: https://doi.org/10.1136/bmj.n71.

PARK, M.; LEAHEY, E.; FUNK, R. J. Papers and patents are becoming less disruptive over time. **Nature**, v. 613, p. 138-144, 2023. DOI: https://doi.org/10.1038/s41586-022-05543-x.

PERIANES-RODRÍGUEZ, A.; WALTMAN, L.; VAN ECK, N. J. Constructing bibliometric networks: a comparison between full and fractional counting. **Journal of Informetrics**, v. 10, n. 4, p. 1178-1195, 2016. DOI: https://doi.org/10.1016/j.joi.2016.10.006.

PIWOWAR, H. Altmetrics: value all research products. **Nature**, v. 493, p. 159, 2013. DOI: https://doi.org/10.1038/493159a.

PIWOWAR, H. *et al.* The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. **PeerJ**, v. 6, e4375, 2018. DOI: https://doi.org/10.7717/peerj.4375.

PORTER, A. L. *et al.* Measuring researcher interdisciplinarity. **Scientometrics**, v. 72, p. 117-147, 2007. DOI: https://doi.org/10.1007/s11192-007-1700-y.

PORTER, A. L.; DETAMPEL, M. J. Technology opportunities analysis. **Technological Forecasting and Social Change**, v. 49, n. 3, p. 237-255, 1995. DOI: https://doi.org/10.1016/0040-1625(95)00022-3.

PORTER, A. L.; RAFOLS, I. Is science becoming more interdisciplinary? Measuring and *mapping* six research fields over time. **Scientometrics**, v. 81, p. 719-745, 2009. DOI: https://doi.org/10.1007/s11192-008-2197-2.

PRICE, D. J. de Solla. Networks of scientific papers. **Science**, v. 149, n. 3683, p. 510-515, 1965. DOI: https://doi.org/10.1126/science.149.3683.510.

PRIEM, J.; PIWOWAR, H.; ORR, R. OpenAlex: a fully-open index of scholarly works, authors, venues, institutions, and concepts. **arXiv**, arXiv:2205.01833, 2022. DOI: https://doi.org/10.48550/arXiv.2205.01833.

PRITCHARD, A. Statistical bibliography or bibliometrics. **Journal of Documentation**, v. 25, n. 4, p. 348-349, 1969. Disponível em: https://cir.nii.ac.jp/crid/1572543025335888640?lang=en. Acesso em: 25 jul. 2026.

RAFOLS, I. *et al.* How journal rankings can suppress interdisciplinary research: a comparison between Innovation Studies and Business & Management. **Research Policy**, v. 41, n. 7, p. 1262-1282, 2012. DOI: https://doi.org/10.1016/j.respol.2012.03.015.

RAFOLS, I.; PORTER, A. L.; LEYDESDORFF, L. Science overlay maps: a new tool for research policy and library management. **Journal of the American Society for Information Science and Technology**, v. 61, n. 9, p. 1871-1887, 2010. DOI: https://doi.org/10.1002/asi.21368.

RAMAN, R. *et al.* Unveiling the dynamics of AI applications: a review of reviews using scientometrics and BERTopic modeling. **Journal of Innovation & Knowledge**, v. 9, n. 3, 100517, 2024. DOI: https://doi.org/10.1016/j.jik.2024.100517.

REIMERS, N.; GUREVYCH, I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING (EMNLP-IJCNLP), 2019. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2019. p. 3982-3992. DOI: https://doi.org/10.18653/v1/D19-1410.

RENZE, M.; GUVEN, E. The effect of sampling temperature on problem solving in large language models. In: FINDINGS OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS: EMNLP 2024, 2024. **Proceedings** [...]. Stroudsburg: Association for Computational Linguistics, 2024. p. 7346-7356. Disponível em: https://aclanthology.org/2024.findings-emnlp.432/. Acesso em: 25 jul. 2026.

RETHLEFSEN, M. L. *et al.* PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. **Systematic Reviews**, v. 10, art. 39, 2021. DOI: https://doi.org/10.1186/s13643-020-01542-z.

ROSVALL, M.; BERGSTROM, C. T. Mapping change in large networks. **PLOS ONE**, v. 5, n. 1, e8694, 2010. DOI: https://doi.org/10.1371/journal.pone.0008694.

ROTOLO, D.; HICKS, D.; MARTIN, B. R. What is an emerging technology? **Research Policy**, v. 44, n. 10, p. 1827-1843, 2015. DOI: https://doi.org/10.1016/j.respol.2015.06.006.

RUSHFORTH, A. D.; HAMMARFELT, B. The rise of responsible metrics as a professional reform movement: a collective action frames account. **Quantitative Science Studies**, v. 4, n. 4, p. 879-897, 2023. DOI: https://doi.org/10.1162/qss_a_00280.

SAHOO, P. *et al.* A systematic survey of prompt engineering in large language models: techniques and applications. **arXiv**, arXiv:2402.07927, 2024. Disponível em: https://arxiv.org/abs/2402.07927. Acesso em: 25 jul. 2026.

SAMPSON, M. *et al.* No consensus exists on search reporting methods for systematic reviews. **Journal of Clinical Epidemiology**, v. 61, n. 8, p. 748-754, 2008. DOI: https://doi.org/10.1016/j.jclinepi.2008.05.001.

SARIKAYA, A. *et al.* What do we talk about when we talk about dashboards? **IEEE Transactions on Visualization and Computer Graphics**, v. 25, n. 1, p. 682-692, 2019. DOI: https://doi.org/10.1109/TVCG.2018.2864903.

SCHMIDT, M. The Sankey diagram in energy and material flow management: part I: history. **Journal of Industrial Ecology**, v. 12, n. 1, p. 82-94, 2008. DOI: https://doi.org/10.1111/j.1530-9290.2008.00004.x.

SCHMIDT, M. The Sankey diagram in energy and material flow management: part II: methodology and current applications. **Journal of Industrial Ecology**, v. 12, n. 2, p. 173-185, 2008. DOI: https://doi.org/10.1111/j.1530-9290.2008.00015.x.

SCHULHOFF, S. *et al.* The Prompt Report: a systematic survey of prompting techniques. **arXiv**, arXiv:2406.06608, 2024. DOI: https://doi.org/10.48550/arXiv.2406.06608.

SEGEL, E.; HEER, J. Narrative visualization: telling stories with data. **IEEE Transactions on Visualization and Computer Graphics**, v. 16, n. 6, p. 1139-1148, 2010. DOI: https://doi.org/10.1109/TVCG.2010.179.

SEGLEN, P. O. Why the impact factor of journals should not be used for evaluating research. **BMJ**, v. 314, n. 7079, p. 497-502, 1997. DOI: https://doi.org/10.1136/bmj.314.7079.497.

SELBST, A. D. *et al.* Fairness and abstraction in sociotechnical systems. In: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2019. **Proceedings** [...]. New York: ACM, 2019. p. 59-68. DOI: https://doi.org/10.1145/3287560.3287598.

SHNEIDERMAN, B. The eyes have it: a task by data type taxonomy for information visualizations. In: IEEE SYMPOSIUM ON VISUAL LANGUAGES, 1996. **Proceedings** [...]. [S. l.]: IEEE, 1996. p. 336-343. DOI: https://doi.org/10.1109/VL.1996.545307.

SIMERA, I. *et al.* Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. **BMC Medicine**, v. 8, art. 24, 2010. DOI: https://doi.org/10.1186/1741-7015-8-24.

SINGH, V. K. *et al.* The journal coverage of Web of Science, Scopus and Dimensions: a comparative analysis. **Scientometrics**, v. 126, p. 5113-5142, 2021. DOI: https://doi.org/10.1007/s11192-021-03948-5.

SMALL, H. Co-citation in the scientific literature: a new measure of the relationship between two documents. **Journal of the American Society for Information Science**, v. 24, n. 4, p. 265-269, 1973. DOI: https://doi.org/10.1002/asi.4630240406.

SMALL, H. A co-citation model of a scientific specialty: a longitudinal study of collagen research. **Social Studies of Science**, v. 7, n. 2, p. 139-166, 1977. DOI: https://doi.org/10.1177/030631277700700202.

SMALL, H. Tracking and predicting growth areas in science. **Scientometrics**, v. 68, p. 595-610, 2006. DOI: https://doi.org/10.1007/s11192-006-0132-y.

SOLLACI, L. B.; PEREIRA, M. G. The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. **Journal of the Medical Library Association**, v. 92, n. 3, p. 364-367, 2004. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC442179/. Acesso em: 25 jul. 2026.

TAHAMTAN, I.; BORNMANN, L. What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018. **Scientometrics**, v. 121, p. 1635-1684, 2019. DOI: https://doi.org/10.1007/s11192-019-03243-4.

TAHAMTAN, I.; SAFIPOUR AFSHAR, A.; AHAMADZADEH, K. Factors affecting number of citations: a comprehensive review of the literature. **Scientometrics**, v. 107, p. 1195-1225, 2016. DOI: https://doi.org/10.1007/s11192-016-1889-2.

TEKLES, A.; BORNMANN, L. Author name disambiguation of bibliometric data: a comparison of several unsupervised approaches. **Quantitative Science Studies**, v. 1, n. 4, p. 1510-1528, 2020. DOI: https://doi.org/10.1162/qss_a_00081.

THELWALL, M. Bibliometrics to webometrics. **Journal of Information Science**, v. 34, n. 4, p. 605-621, 2008. DOI: https://doi.org/10.1177/0165551507087238.

THELWALL, M. Microsoft Academic: a multidisciplinary comparison of citation counts with Scopus and Mendeley for 29 journals. **Journal of Informetrics**, v. 12, n. 4, p. 1201-1212, 2018. DOI: https://doi.org/10.1016/j.joi.2018.10.006.

THELWALL, M. *et al.* Do altmetrics work? Twitter and ten other social web services. **PLOS ONE**, v. 8, n. 5, e64841, 2013. DOI: https://doi.org/10.1371/journal.pone.0064841.

THELWALL, M.; VAUGHAN, L.; BJÖRNEBORN, L. Webometrics. **Annual Review of Information Science and Technology**, v. 39, n. 1, p. 81-135, 2005. DOI: https://doi.org/10.1002/aris.1440390110.

THOMAS, J.; MCNAUGHT, J.; ANANIADOU, S. Applications of text mining within systematic reviews. **Research Synthesis Methods**, v. 2, n. 1, p. 1-14, 2011. DOI: https://doi.org/10.1002/jrsm.27.

TSAFNAT, G. *et al.* Systematic review automation technologies. **Systematic Reviews**, v. 3, art. 74, 2014. DOI: https://doi.org/10.1186/2046-4053-3-74.

VAN ECK, N. J.; WALTMAN, L. Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics**, v. 84, p. 523-538, 2010. DOI: https://doi.org/10.1007/s11192-009-0146-3.

VAN ECK, N. J.; WALTMAN, L. Visualizing bibliometric networks. In: DING, Y.; ROUSSEAU, R.; WOLFRAM, D. (ed.). **Measuring scholarly impact**. Cham: Springer, 2014. p. 285-320. DOI: https://doi.org/10.1007/978-3-319-10377-8_13.

VAN ECK, N. J.; WALTMAN, L. Citation-based clustering of publications using CitNetExplorer and VOSviewer. **Scientometrics**, v. 111, p. 1053-1070, 2017. DOI: https://doi.org/10.1007/s11192-017-2300-7.

VAN ECK, N. J. *et al.* A comparison of two techniques for bibliometric mapping: multidimensional scaling and VOS. **Journal of the American Society for Information Science and Technology**, v. 61, n. 12, p. 2405-2416, 2010. DOI: https://doi.org/10.1002/asi.21421.

VAN LEEUWEN, T. N. *et al.* Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. **Scientometrics**, v. 51, p. 335-346, 2001. DOI: https://doi.org/10.1023/A:1010549719484.

VASWANI, A. *et al.* Attention is all you need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017. **Proceedings** [...]. [S. l.]: NeurIPS, 2017. p. 5998-6008. Disponível em: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html. Acesso em: 25 jul. 2026.

VISSER, M.; VAN ECK, N. J.; WALTMAN, L. Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. **Quantitative Science Studies**, v. 2, n. 1, p. 20-41, 2021. DOI: https://doi.org/10.1162/qss_a_00112.

WAGNER, C. S.; LEYDESDORFF, L. Network structure, self-organization, and the growth of international collaboration in science. **Research Policy**, v. 34, n. 10, p. 1608-1618, 2005. DOI: https://doi.org/10.1016/j.respol.2005.08.002.

WALLACE, B. C. *et al.* Semi-automated screening of biomedical citations for systematic reviews. **BMC Bioinformatics**, v. 11, art. 55, 2010. DOI: https://doi.org/10.1186/1471-2105-11-55.

WALTERS, W. H.; WILDER, E. I. Fabrication and errors in the bibliographic citations generated by ChatGPT. **Scientific Reports**, v. 13, 14045, 2023. DOI: https://doi.org/10.1038/s41598-023-41032-5.

WALTMAN, L. A review of the literature on citation impact indicators. **Journal of Informetrics**, v. 10, n. 2, p. 365-391, 2016. DOI: https://doi.org/10.1016/j.joi.2016.02.007.

WALTMAN, L.; VAN ECK, N. J. A new methodology for constructing a publication-level classification system of science. **Journal of the American Society for Information Science and Technology**, v. 63, n. 12, p. 2378-2392, 2012. DOI: https://doi.org/10.1002/asi.22748.

WALTMAN, L.; VAN ECK, N. J. Source normalized indicators of citation impact: an overview of different approaches and an empirical comparison. **Scientometrics**, v. 96, p. 699-716, 2013. DOI: https://doi.org/10.1007/s11192-012-0913-4.

WALTMAN, L.; VAN ECK, N. J.; NOYONS, E. C. M. A unified approach to mapping and clustering of bibliometric networks. **Journal of Informetrics**, v. 4, n. 4, p. 629-635, 2010. DOI: https://doi.org/10.1016/j.joi.2010.07.002.

WALTMAN, L. *et al.* Towards a new crown indicator: some theoretical considerations. **Journal of Informetrics**, v. 5, n. 1, p. 37-47, 2011. DOI: https://doi.org/10.1016/j.joi.2010.08.001.

WANG, C.; BLEI, D.; HECKERMAN, D. Continuous time dynamic topic models. In: CONFERENCE ON UNCERTAINTY IN ARTIFICIAL INTELLIGENCE, 28., 2012. **Proceedings** [...]. [S. l.]: AUAI Press, 2012. p. 579-586. Disponível em: https://arxiv.org/abs/1206.3298. Acesso em: 25 jul. 2026.

WANG, X. *et al.* Self-consistency improves chain of thought reasoning in language models. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 2023. **Proceedings** [...]. [S. l.]: ICLR, 2023. Disponível em: https://openreview.net/forum?id=1PL1NIMMrw. Acesso em: 25 jul. 2026.

WEI, J. *et al.* Chain-of-thought prompting elicits reasoning in large language models. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 35., 2022. **Proceedings** [...]. [S. l.]: NeurIPS, 2022. p. 24824-24837. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2022/hash/9d5609613524ecf4f15af0f7b31abca4-Abstract-Conference.html. Acesso em: 25 jul. 2026.

WHITE, H. D.; GRIFFITH, B. C. Author cocitation: a literature measure of intellectual structure. **Journal of the American Society for Information Science**, v. 32, n. 3, p. 163-171, 1981. DOI: https://doi.org/10.1002/asi.4630320302.

WHITE, J. *et al.* A prompt pattern catalog to enhance prompt engineering with ChatGPT. In: CONFERENCE ON PATTERN LANGUAGES OF PROGRAMS, 30., 2023. **Proceedings** [...]. [S. l.]: PLoP, 2023. Disponível em: https://arxiv.org/abs/2302.11382. Acesso em: 25 jul. 2026.

WOUTERS, P. *et al.* Rethinking impact factors: better ways to judge a journal. **Nature**, v. 569, p. 621-623, 2019. DOI: https://doi.org/10.1038/d41586-019-01643-3.

WU, S. *et al.* Visual analysis of LLM-based entity resolution from scientific papers. **Visual Informatics**, v. 9, n. 2, 100236, 2025. DOI: https://doi.org/10.1016/j.visinf.2025.100236.

YAO, S. *et al.* ReAct: synergizing reasoning and acting in language models. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 2023. **Proceedings** [...]. [S. l.]: ICLR, 2023. Disponível em: https://openreview.net/forum?id=WE_vluYUL-X. Acesso em: 25 jul. 2026.

ZUPIC, I.; ČATER, T. Bibliometric methods in management and organization. **Organizational Research Methods**, v. 18, n. 3, p. 429-472, 2015. DOI: https://doi.org/10.1177/1094428114562629

unesco
chair

Cátedra UNESCO de
Juventude, Educação
e Sociedade