



**UNIVERSIDADE DE BRASÍLIA – UnB**  
**FACULDADE DE CIÊNCIA DA INFORMAÇÃO - FCI**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – PPGCINF**

**LILIANE SIMÕES DOS SANTOS**

**ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA DE ONTOLOGIA EM  
CIÊNCIA DA INFORMAÇÃO NO CENÁRIO INTERNACIONAL**

**BRASÍLIA**

**2025**

**LILIANE SIMÕES DOS SANTOS**

**ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA DE ONTOLOGIA EM  
CIÊNCIA DA INFORMAÇÃO NO CENÁRIO INTERNACIONAL**

Dissertação de mestrado apresentada à banca examinadora como requisito para obtenção do título de Mestre em Ciência da Informação pela Faculdade de Ciência da Informação da Universidade de Brasília.

Orientador: Prof. Dr. Cláudio Gottschalg Duque

**BRASÍLIA**

**2025**

Ficha catalográfica elaborada automaticamente,  
com os dados fornecidos pelo(a) autor(a)

SS237a SIMÕES DOS SANTOS, LILIANE  
ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA DE ONTOLOGIA  
EM CIÊNCIA DA INFORMAÇÃO NO CENÁRIO INTERNACIONAL / LILIANE  
SIMÕES DOS SANTOS; orientador Cláudio Gottschalg Duque.  
Brasília, 2025.  
195 p.

Dissertação (Mestrado em Ciência da Informação)  
Universidade de Brasília, 2025.

1. Ontologia. 2. Ciência da Informação. 3. Bibliometria.  
4. Mapeamento científico. 5. Indicadores bibliométricos. I.  
Gottschalg Duque, Cláudio, orient. II. Título.

**UNIVERSIDADE DE BRASÍLIA**  
**PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

**Ata Nº: 98**

Aos vinte e oito dias do mês de novembro do ano de dois mil vinte e cinco, instalou-se a banca examinadora de Dissertação de Mestrado da aluna Liliane Simões dos Santos, matrícula 220004170. A banca examinadora foi composta pelos professores Dra. Ivette Kafure Muñoz / membro titular interno à UnB / PPGCINF, Dr. Maurício Barcellos Almeida / membro titular externo à UnB / UFMG, Dr. Tomás Roberto Cotta Orlandi / membro suplente / ECT, (Suplente) e Dr. Claudio Gottschalg Duque / presidente / PPGCINF, orientador(a)/presidente. A discente apresentou o trabalho intitulado "Análise Bibliométrica da Produção Científica de Ontologia em Ciência da Informação no Cenário Internacional".

Concluída a exposição, procedeu-se a arguição do(a) candidato(a), e após as considerações dos examinadores o resultado da avaliação do trabalho foi:

( X ) Pela aprovação do trabalho;

( ) Pela aprovação do trabalho, com revisão de forma, indicando o prazo de até 30 (trinta) dias para apresentação definitiva do trabalho revisado;

( ) Pela reformulação do trabalho, indicando o prazo de **(Nº DE MESES)** para nova versão;

( ) Pela reprovação do trabalho, conforme as normas vigentes na Universidade de Brasília.

Conforme os Artigos 34, 39 e 40 da Resolução 0080/2021 - CEPE, o(a) candidato(a) não terá o título se não cumprir as exigências acima.

Dr. Claudio Gottschalg Duque, PPGCINF / UnB  
(Presidente)

Dra. Ivette Kafure Muñoz, PPGCINF / UnB  
(Membro Titular Interno)

Dr. Maurício Barcellos Almeida, UFMG  
(Membro Titular Externo)

Dr. Tomás Roberto Cotta Orlandi, ECT  
(Membro Suplente)

Liliane Simões dos Santos  
(Mestranda)



Documento assinado eletronicamente por **Claudio Gottschalg Duque, Pesquisador(a) Colaborador(a) Pleno(a) da Faculdade de Ciência da Informação**, em 09/12/2025, às 10:53, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Ivette Kafure Munoz, Membro do Colegiado da Pós-Graduação da Faculdade de Ciência da Informação**, em 09/12/2025, às 18:19, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Maurício Barcellos Almeida, Usuário Externo**, em 10/12/2025, às 10:44, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Liliane Simões dos Santos, Usuário Externo**, em 10/12/2025, às 13:30, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Ana Lucia de Abreu Gomes, Vice-Coordenador(a) da Pós-Graduação da Faculdade de Ciência da Informação**, em 10/12/2025, às 17:02, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



A autenticidade deste documento pode ser conferida no site [http://sei.unb.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.unb.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **13338395** e o código CRC **A60B7A08**.

*Dedico este trabalho a Deus por sua infinita bondade, graça e misericórdia ao me conceder força e sabedoria nos momentos de dificuldade além de sustentar a minha fé ao longo de toda a minha trajetória acadêmica. Certamente, sem a sua infinita bondade, graça e misericórdia a concretização deste sonho não seria possível. Portanto, a Ele, toda honra e glória por esta conquista tão significativa e especial pra mim.*

## AGRADECIMENTOS

Em primeiro lugar, a Deus, por guiar e conduzir a minha vida com propósito.

Aos meus pais, Olga e Lenildo, e familiares que acolheram meus sonhos e apoiaram as escolhas realizadas ao longo da minha trajetória de vida. O amor incondicional deles é o que me impulsiona a crescer e evoluir a cada dia.

À Universidade de Brasília, em especial, à Faculdade de Ciência da Informação e ao Programa de Pós-Graduação em Ciência da Informação por proporcionar um ambiente de aprendizado, desconstrução e reformulação do conhecimento científico.

Aos eventos científicos que participei durante o mestrado, em especial, o Seminário de Pesquisa em Ontologias no Brasil e o Encontro Brasileiro de Bibliometria e Cientometria que trouxeram reflexões pertinentes que contribuíram para o aperfeiçoamento da pesquisa científica.

Ao orientador, Prof. Dr. Cláudio Gottschalg Duque e ao grupo de pesquisa *“Research Expert Group for Intelligent Information in Multimodal Environment using Natural language Technologies and Ontologies”* por todos os ensinamentos compartilhados durante a trajetória acadêmica.

Aos membros da Banca de Qualificação e de Defesa por aceitarem o convite em participar da banca examinadora e contribuir com o projeto de pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior por apoiar o desenvolvimento do estudo e financiar a pesquisa com a bolsa de estudo.

Aos amigos, em especial, a Raffaella Dayane Afonso e a rede de apoio da psicóloga Michele Santos Baleeiro por me acolherem com palavras de encorajamento além de um olhar sensível e uma escuta ativa em momentos de vulnerabilidade emocional.

Aos meus animais de estimação (Neru, Garu e Yuki) por cada gesto de amor e companheirismo durante as diversas madrugadas de solidão acadêmica.

Por fim, a Liliane, que se permitiu sonhar... Que acreditou naquela ideia iniciada em 2010 enquanto fazia o bacharelado em Letras (LEAMSI) quando despertou o interesse pela Arquivologia e a área da Ciência da Informação... Que não desistiu apesar de tantas circunstâncias desafiadoras... Que insistiu e resistiu com dedicação e persistência até o fim!

“Na era da informação, a invisibilidade é equivalente à morte” Zygmunt Bauman (2003).

## RESUMO

A internacionalização da produção científica impulsionada pela publicação em acesso aberto tem promovido a colaboração global e a democratização do conhecimento. Sabe-se que o movimento de acesso aberto visa democratizar o acesso à informação, remover barreiras de acessibilidade, incentivar redes de colaboração e impulsionar o desenvolvimento científico e tecnológico. Ao considerar que ontologia no contexto da Ciência da Informação são representações formais e estruturadas de um domínio especializado do conhecimento ela é definida como uma estrutura que visa melhorar a recuperação da informação, padronizar a terminologia e representar o conhecimento em ambiente digital. O objetivo do estudo é mapear a produção científica de ontologia em Ciência da Informação visando identificar padrões e tendências do campo científico a fim de descrever o panorama da pesquisa no cenário internacional. A estratégia de pesquisa adotada no estudo é a análise bibliométrica uma vez que ao analisar quantitativamente a produção científica é possível identificar padrões, tendências e as principais contribuições à comunidade científica. A coleta de dados foi realizada na base *Web of Science*. Os dados foram exportados em formato .txt para corrigir erros de formatação, duplicações e inconsistências. A análise dos dados foi realizada com o software Bibliometrix que é um pacote R recomendado para realizar análises bibliométricas. Os resultados apresentam o panorama do campo científico que descreve os indicadores bibliométricos da produção científica de ontologia em Ciência da Informação no cenário internacional.

**Palavras-chave:** Ontologia. Ciência da Informação. Bibliometria. Mapeamento científico. Indicadores bibliométricos.

## ABSTRACT

The internationalization of scientific production, driven by open access publishing, has promoted global collaboration and the democratization of knowledge. It is known that the open access movement aims to democratize access to information, remove accessibility barriers, encourage collaborative networks and boost scientific and technological development. Considering that ontology in the context of information science refers to formal and structured representations of a specialized domain of knowledge it is defined as a framework that aims to improve information retrieval, standardize terminology and represent knowledge in a digital environment. The objective of this study is to map the scientific production of ontology in information science, aiming to identify patterns and trends in the scientific field in order to describe the research landscape in the international context. Bibliometric analysis is the research strategy adopted in this study because by quantitatively analyzing scientific production it is possible to identify patterns, trends and the main contributions to the scientific community. Data collection was carried out using the Web of Science database. The data were exported in .txt format to correct formatting errors, duplications, and inconsistencies. Data analysis was performed using Bibliometrix software, an R package recommended for bibliometric analysis. The results present an overview of the scientific field, describing the bibliometric indicators of scientific production of ontology in information science in the international context.

**Keywords:** Ontology. Information Science. Bibliometrics. Scientific mapping. Bibliometric indicators.

## LISTA DE FIGURAS

Figura 1 - Fluxo de mapeamento científico .....	81
Figura 2 - Validação dos metadados do corpus de referência .....	85
Figura 3 - Coleta de dados na <i>Web of Science</i> .....	88
Figura 4 - Documentos na edição " <i>Social Sciences Citation Index - SSCI</i> " .....	89
Figura 5 - Artigos indexados na edição "SSCI" .....	90
Figura 6 - Artigos em acesso aberto na edição SSCI.....	91
Figura 7 - Artigos em acesso aberto na Ciência da Informação.....	91
Figura 8 - Visão geral do <i>corpus</i> de referência.....	99
Figura 9 - Estimativa da vida útil da produção científica .....	113
Figura 10 - Perfil global Almeida, M. B .....	131
Figura 11 - Perfil local Almeida, M. B .....	133
Figura 12 - Perfil global Marcondes, C. H .....	134
Figura 13 - Perfil local Marcondes, C. H.....	136
Figura 14 - Evolução temporal da produção científica dos autores .....	140
Figura 15 - Nuvem de palavras .....	164
Figura 16 - Agrupamento por acoplamento bibliográfico .....	170
Figura 17 - Rede de co-ocorrência.....	172
Figura 18 - Rede de co-citação de documentos.....	174
Figura 19 - Mapa de colaboração científica global .....	176

## LISTA DE GRÁFICOS

Gráfico 1 - Indicador da produtividade científica anual.....	108
Gráfico 2 - Indicador da média de citações por ano .....	109
Gráfico 3 - Ciclo de vida da produção científica .....	111
Gráfico 4 - Análise Referências → Autores → Palavras-chave.....	116
Gráfico 5 - Análise Fontes → Autores → Países.....	118
Gráfico 6 - Análise Palavras-chave → Autores → Referências citadas .....	119
Gráfico 7 - Análise Países → Autores → Palavras-chave.....	121
Gráfico 8 - Evolução temporal das fontes de publicações científicas.....	128
Gráfico 9 - Indicador da produtividade do autor através da Lei de Lotka .....	142
Gráfico 10 - Evolução temporal da produção de afiliações .....	147
Gráfico 11 - Evolução temporal da produção científica dos países.....	153
Gráfico 12 - Indicador de frequência das palavras .....	166
Gráfico 13 - Indicador de tópicos em alta .....	168

## LISTA DE QUADROS

Quadro 1 - Interpretações conceituais de ontologia .....	30
Quadro 2 - Distinção conceitual entre os termos de ontologia .....	31
Quadro 3 - Estrutura basilar do modelo ontológico .....	32
Quadro 4 - Classificação das abordagens ontológicas .....	33
Quadro 5 - Vertentes conceituais de ontologia.....	35
Quadro 6 - Interdisciplinaridade das abordagens conceituais de ontologia .....	38
Quadro 7 - Princípios ontológicos .....	45
Quadro 8 - Leis e princípios bibliométricos.....	57
Quadro 9 - Termos e conceitos relacionados à comunicação científica.....	62
Quadro 10 - Princípios e recomendações para a comunicação de dados científicos.....	65
Quadro 11 - Fatores de influência na comunicação dos dados de pesquisa .....	66
Quadro 12 - Recomendações da política brasileira de ciência aberta e adoção de boas práticas de avaliação científica .....	70
Quadro 13 - Objetivos do Programa Institucional de Internacionalização (PrInt) .....	73
Quadro 14 - Ações de inserção internacional.....	74
Quadro 15 - Análise comparativa do fluxo de mapeamento científico.....	80
Quadro 16 - Análise comparativa de bases de dados bibliográficas .....	83
Quadro 17 - Critérios de escolha da Base de Dados <i>Web of Science</i> .....	86
Quadro 18 - Etapas de coleta de dados na <i>Web of Science</i> .....	92
Quadro 19 - Critérios de normalização dos dados de pesquisa .....	93
Quadro 20 - Indicadores adotados na análise bibliométrica.....	103
Quadro 21 - Modo de visualização dos indicadores bibliométricos .....	106
Quadro 22 - Estratégias de refinamento da análise bibliométrica .....	107

## LISTA DE TABELAS

Tabela 1 - Informações dos dados de pesquisa.....	101
Tabela 2 - Top 10: fontes de publicações científicas mais relevantes.....	123
Tabela 3 - Top 10: fontes de publicações científicas locais mais citadas.....	124
Tabela 4 - Top 10: fontes de publicações científicas - Lei de Bradford.....	125
Tabela 5 - Top 10: métricas de impacto local das fontes de publicações científica.....	126
Tabela 6 - Top 10: autores mais relevantes.....	137
Tabela 7 - Top 10: autores locais mais citados.....	139
Tabela 8 - Métricas de impacto local dos autores.....	143
Tabela 9 - Top 10: afiliações mais relevantes.....	145
Tabela 10 - Top 10: países do autor correspondente.....	149
Tabela 11 - Top 10: produção científica dos países.....	151
Tabela 12 - Top 10: países mais citados.....	154
Tabela 13 - Top 10: documentos mais citados globalmente.....	157
Tabela 14 - Top 10: documentos locais mais citados.....	159
Tabela 15 - Top 10: referências locais mais citadas.....	161
Tabela 16 - Top 10: palavras mais frequentes.....	163
Tabela 17 - Colaboração Brasil - Austria - Colombia – Portugal.....	177

## LISTA DE ABREVIATURAS E SIGLAS

AB	Abstract
APCs	Taxas de processamento de artigos
A&HCI	Arts & Humanities Citation Index
BRAPCI	Base de Dados em Ciência da Informação
C1	Afiliação institucional
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CR	Referências citadas
CT & I	Ciência, Tecnologia e Inovação
DE	Palavras-chave dos autores (Descriptors)
DI	Digital Object Identifier
DT	Tipo de documento
ESCI	Emerging Sources Citation Index
EUA	Estados Unidos da América
IA	Inteligência Artificial
ID	Palavras-chave Plus
IN	Instrução Normativa
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
ISI	Institute for Scientific Information
JAMIA	Journal of the American Medical Informatics Association
JASIST	Journal of the American Society for Information Science and Technology
LEAMSI	Línguas Estrangeiras Aplicadas ao Multilinguismo e à Sociedade da Informação
LNCS	Lecture Notes in Computer Science
LISTA	Library, Information Science & Technology Abstracts
MCA	Análise de Correspondência Múltipla
MDS	Escalonamento Multidimensional
MCTI	Ministério da Ciência, Tecnologia e Inovação
OpenAlex	OpenAlex Database
OWL	Web Ontology Language
PCA	Análise de Componentes Principais

PPGCINF	Programa de Pós-Graduação em Ciência da Informação
PrInt	Programa Institucional de Internacionalização
PY	Publication Year
R	Linguagem de programação R
RP	Autor correspondente
SCIE	Science Citation Index Expanded
SSCI	Social Sciences Citation Index
TC	Total de citações
TI	Título
TXT	Documento de texto
UnB	Universidade de Brasília
UNESCO	Nações Unidas para a Educação, a Ciência e a Cultura
WoS	Web of Science
WoS CC	Web of Science Core Collection

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>18</b>
1.1	EVIDÊNCIAS CIENTÍFICAS E TRABALHOS RELACIONADOS .....	19
1.2	LACUNAS DE PESQUISA .....	21
1.3	PROBLEMATIZAÇÃO DA PESQUISA .....	24
1.4	OBJETIVO GERAL DO ESTUDO .....	25
<b>1.4.1</b>	<b>Objetivos Específicos do Estudo .....</b>	<b>25</b>
1.5	JUSTIFICATIVA DO ESTUDO .....	25
<b>2</b>	<b>COMPREENDENDO O CAMPO CIENTÍFICO DE ONTOLOGIA .....</b>	<b>29</b>
2.1	MARCO TEÓRICO E DEFINIÇÃO CONCEITUAL DE ONTOLOGIA .....	29
2.2	TRADIÇÃO E LINHAS DE PESQUISA EM ONTOLOGIA .....	36
2.3	ONTOLOGIA NA PERSPECTIVA DA CIÊNCIA DA INFORMAÇÃO .....	39
<b>3</b>	<b>BIBLIOMETRIA E ESTUDO MÉTRICO DA INFORMAÇÃO .....</b>	<b>47</b>
3.1	MARCO TEÓRICO E DEFINIÇÃO CONCEITUAL DE BIBLIOMETRIA .....	47
3.2	BIBLIOMETRIA: MÉTODO DE MAPEAMENTO CIENTÍFICO .....	50
3.3	CONCEITOS ESTATÍSTICOS APLICADOS ÀS UNIDADES DE INFORMAÇÃO .....	53
3.4	INDICADORES BIBLIOMÉTRICOS DE ATIVIDADES CIENTÍFICAS .....	55
<b>4</b>	<b>BOAS PRÁTICAS CIENTÍFICAS EM CIÊNCIA ABERTA E INCLUSIVAS .....</b>	<b>60</b>
4.1	A COMUNICAÇÃO CIENTÍFICA E A DEMOCRATIZAÇÃO DO ACESSO À INFORMAÇÃO .....	60
4.2	O MOVIMENTO DE ACESSO ABERTO EM PUBLICAÇÕES CIENTÍFICAS .....	64
4.3	A INTERNACIONALIZAÇÃO DA CIÊNCIA .....	69
<b>5</b>	<b>DELINEAMENTO METODOLÓGICO .....</b>	<b>78</b>
5.1	FLUXO DE MAPEAMENTO CIENTÍFICO .....	79
5.2	BASE DE DADOS BIBLIOGRÁFICA: <i>WEB OF SCIENCE</i> .....	82
5.3	COLETA, ANÁLISE, VISUALIZAÇÃO E INTERPRETAÇÃO DE DADOS .....	88
<b>6</b>	<b>FERRAMENTAS PARA ANÁLISE BIBLIOMÉTRICA .....</b>	<b>95</b>
6.1	BIBEXCEL .....	95
6.2	RSTUDIO .....	96
6.3	BIBLIOMETRIX .....	96

6.4	BIBLIOSHINY .....	96
<b>7</b>	<b>ANÁLISE BIBLIOMÉTRICA .....</b>	<b>97</b>
7.1	PANORAMA DO CAMPO CIENTÍFICO .....	98
7.2	FONTES DE PUBLICAÇÕES CIENTÍFICAS .....	122
7.3	AUTORES .....	130
7.4	AFILIAÇÕES .....	144
7.5	PAÍSES .....	148
7.6	DOCUMENTOS.....	156
7.7	REFERÊNCIAS CITADAS .....	160
7.8	PALAVRAS .....	162
7.9	AGRUPAMENTO POR ACOPLAMENTO BIBLIOGRÁFICO .....	169
7.10	ESTRUTURA CONCEITUAL.....	171
7.11	ESTRUTURA INTELECTUAL .....	174
7.12	ESTRUTURA SOCIAL .....	175
<b>8</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>179</b>
8.1	CONTRIBUIÇÃO DO ESTUDO À CIÊNCIA DA INFORMAÇÃO.....	179
8.2	LIMITAÇÕES DA PESQUISA BIBLIOMÉTRICA.....	180
8.3	RECOMENDAÇÕES PARA INVESTIGAÇÕES FUTURAS .....	182
	<b>REFERÊNCIAS.....</b>	<b>184</b>
	<b>APÊNDICE A - CRONOGRAMA DE PESQUISA.....</b>	<b>193</b>
	<b>ANEXO A - DECLARAÇÃO DE FINANCIAMENTO.....</b>	<b>195</b>

## 1 INTRODUÇÃO

A ontologia constitui um campo de estudo amplamente abordado em diversas áreas de pesquisa e domínios do conhecimento (Almeida, 2014). Contudo, desde a década de 1990 observa-se um aumento expressivo na demanda por ontologias motivado pelo avanço da web semântica e a necessidade crescente de estruturar a informação em ambientes digitais. Nesse contexto, o interesse por ontologias passou a concentrar-se na promoção da interoperabilidade dos sistemas de informação e na viabilização do compartilhamento de dados estruturados, contribuindo para o aprimoramento da comunicação científica (Santos, 2014).

Sob a perspectiva da Ciência da Informação, a ontologia é compreendida como um instrumento de representação do conhecimento que atua no controle terminológico e semântico dos sistemas de informação (Santos, 2014). Nessa área, os estudos evidenciam que o uso de ontologia representa uma evolução em relação aos instrumentos tradicionais de organização e representação do conhecimento ao realizar inferências automáticas e estabelecer relações conceituais complexas nos sistemas informacionais (Almeida *et al.*, 2005).

Assim, as ontologias são concebidas como estruturas formais desenvolvidas por especialistas, destinadas a definir os conceitos e as relações que descrevem um domínio específico do conhecimento, assegurando a consistência e a precisão terminológica. Sob esse ponto de vista, considera-se que a finalidade das ontologias na Ciência da Informação é aprimorar os processos de organização e recuperação da informação, favorecendo a representação estruturada dos conteúdos informacionais em domínios científicos específicos (Almeida; Bax, 2003).

Essa concepção reforça o papel das ontologias para o avanço dos sistemas de informação, fortalecendo o processo de gestão da informação e contribuindo com a disseminação do conhecimento científico. No âmbito das ciências aplicadas, os estudos também apontam para o conceito de ontologia aplicada, entendido como um artefato representacional incorporado aos sistemas de informação com o propósito de prover uma estrutura de mundo às

máquinas, permitindo a inferência automática e o aprendizado computacional (Almeida, 2020).

Ao considerar que a ontologia se constitui como uma disciplina de base filosófica e como um artefato representacional do conhecimento observa-se que a sua construção se apoia em princípios metafísicos adaptados a contextos informacionais. Dessa forma, as ontologias, enquanto artefatos representacionais, estruturam e padronizam o conhecimento em sistemas de informação, viabilizando não apenas a interoperabilidade dos sistemas de informação, mas, também a integração dos ambientes digitais (Almeida, 2020). Tal abordagem evidencia o caráter teórico-prático das ontologias que se consolidam como instrumentos que atuam no desenvolvimento da inteligência artificial e na interoperabilidade dos sistemas de informação.

## 1.1 EVIDÊNCIAS CIENTÍFICAS E TRABALHOS RELACIONADOS

As evidências científicas e os trabalhos relacionados apresentados neste estudo possibilitam contextualizar o referencial teórico que fundamenta a análise bibliométrica da produção científica de ontologia em Ciência da Informação no cenário internacional. A literatura demonstra como essa temática se articula com diferentes vertentes de pesquisa e evidencia o caráter interdisciplinar que permeia o campo. Ademais, as investigações ressaltam a relevância do tema e apontam lacunas teóricas e metodológicas que ainda demandam aprofundamento, contribuindo para o fortalecimento das bases conceituais e práticas da área e para o avanço da pesquisa científica no domínio da Ciência da Informação. Dessa forma, a revisão de literatura sustenta a necessidade de uma investigação que busca compreender a evolução e o comportamento do campo científico de ontologia em Ciência da Informação no cenário internacional.

Entre os estudos que embasam esta pesquisa, destaca-se a análise de Santos (2014) que examina o processo de apropriação da ontologia pela Ciência da Informação no contexto brasileiro. A autora esclarece conceitos fundamentais e discute influências teóricas relacionadas ao uso, aos componentes e às metodologias empregadas na construção de ontologia. Seu estudo evidencia

uma tendência de consolidação teórica em torno das definições conceituais e das implicações práticas associadas ao uso de ontologia na Ciência da Informação.

De acordo com Santos (2014), a ontologia é concebida como um conceito multifacetado que atua nos processos de organização e recuperação da informação. A autora define ontologia como uma ferramenta de organização do conhecimento que potencializa a recuperação da informação por meio da estruturação e do relacionamento semântico dos dados. Desse modo, a apropriação da ontologia pela Ciência da Informação concentra-se na seleção e no mapeamento de termos, conceitos e inter-relações que possibilitam a precisão informacional.

Em consonância com essa abordagem, Marin Neto (2018) discute a representação do conhecimento por meio de ontologia, enfatizando seu papel na integração das fontes de informação em ambientes digitais e na realização das inferências automáticas. O autor identifica metodologias voltadas à criação e ao aprendizado de ontologia, destacando-as como ferramentas semânticas para representação do conhecimento e a consolidação dos estudos sobre organização e recuperação da informação.

De maneira complementar, Santos (2022) apresenta um panorama das políticas governamentais brasileiras relacionadas ao direito de acesso à informação, ressaltando a importância do acesso aberto como uma boa prática de comunicação científica. A autora defende a adoção dos princípios de ciência aberta e o fortalecimento das práticas científicas que abranjam os aspectos de gestão da informação orientados à transparência e à disseminação do conhecimento científico.

Nessa mesma direção, Souza (2023) realiza um mapeamento da produção científica dos programas de pós-graduação em Ciência da Informação, identificando indicadores bibliométricos que mensuram as atividades científicas na área de organização e representação da informação e do conhecimento. O estudo fornece uma visão abrangente da dinâmica produtiva e das tendências temáticas do campo.

Por fim, Costa (2024) desenvolve uma análise bibliométrica voltada à criação de ontologia, descrevendo o panorama das pesquisas científicas, as

redes de colaboração científica e as associações temáticas que configuram o campo. Sua investigação reforça a importância das abordagens quantitativas e sistemáticas para o mapeamento científico em um domínio do conhecimento.

Assim, observa-se que o conjunto de estudos apresentados sustenta a relevância da investigação proposta pois evidencia a necessidade de compreender de forma mais abrangente a configuração, evolução e conexões existentes no campo científico de ontologia em Ciência da Informação no âmbito internacional.

## 1.2 LACUNAS DE PESQUISA

A identificação das lacunas de pesquisa permite reconhecer os limites e as insuficiências presentes nas investigações científicas, configurando-se como um exercício para o avanço do conhecimento científico. No contexto da Ciência da Informação, a revisão de literatura revela avanços na compreensão conceitual e na aplicação prática da ontologia como um instrumento de representação do conhecimento. Contudo, observa-se que parte significativa das pesquisas concentra-se em estudos de natureza teórica ou em práticas empíricas voltadas ao desenvolvimento de projetos ontológicos, atribuindo menor ênfase à análise sistemática da produção científica enquanto método de investigação e de mensuração da maturidade do campo científico.

Dessa forma, as lacunas apresentadas neste estudo partem do interesse em realizar uma análise bibliométrica das publicações científicas em escala internacional com o objetivo de mapear a produção científica relacionada à ontologia na Ciência da Informação. A partir das indicações de trabalhos futuros, torna-se evidente a necessidade de ampliar a compreensão sobre o comportamento e a estrutura da produção científica, além de identificar tendências, interações e áreas emergentes no campo. Essa constatação reforça a importância de desenvolver uma pesquisa bibliométrica que além de sistematizar quantitativamente os indicadores bibliométricos também interprete as dinâmicas de construção, disseminação e consolidação do conhecimento científico de ontologia na Ciência da Informação.

Nesse sentido, o presente estudo busca preencher tal lacuna ao realizar uma análise bibliométrica que descreve o panorama da produção científica de ontologia na Ciência da Informação no contexto internacional, contribuindo para o fortalecimento teórico e metodológico da área e a ampliação das bases de conhecimento da comunidade científica.

Entre as evidências que corroboram com a existência de lacunas, Santos (2014) identifica uma desconexão entre o campo teórico e o prático, apontando limitações significativas em estudos que abordam os componentes e as metodologias de construção de ontologia. A autora observa que embora existam avanços em pesquisas aplicadas, as metodologias permanecem insuficientemente exploradas, devendo, portanto, ser priorizadas em investigações futuras a fim de impulsionar o desenvolvimento científico do campo.

Em decorrência dessa análise, Santos (2014) propõe como direções de pesquisa futura a consolidação dos conceitos ontológicos a fim de aprimorar a compreensão e aplicação prática das classificações ontológicas; o incentivo à colaboração interdisciplinar com vistas ao fortalecimento do desenvolvimento e da aplicação de ontologia; a ampliação das investigações empíricas em diferentes domínios e a incorporação de publicações internacionais para delinear um panorama abrangente do conhecimento científico.

De modo convergente, Marin Neto (2018) discute limitações associadas à implementação de ontologia e à representação formal do conhecimento, destacando o desafio de transformar o conhecimento informal em notações lógicas precisas. Ele salienta que erros de formalização e de linguagem comprometem o processo de inferência dos sistemas de informação e menciona o problema da indecidibilidade na representação lógica que ainda carece de soluções definitivas.

O autor sugere o desenvolvimento de linguagens ontológicas simplificadas, padronizadas e estruturadas com vistas a minimizar ambiguidades e inconsistências. Ele propõe que pesquisas futuras explorem a integração da ontologia com técnicas de inteligência artificial; a investigação da relação entre diferentes formas de representação do conhecimento e o aprimoramento das

metodologias de criação e gerenciamento ontológico, garantindo a sua relevância e eficácia no contexto dos sistemas de informação (Marin Neto, 2018).

No cenário contemporâneo, o avanço das tecnologias de informação e comunicação intensificou a produção e o compartilhamento de dados, criando ambientes informacionais complexos. Todavia, conforme apontam Castro e Simionato (2020), esse crescimento se deu de forma desordenada, resultando em sobrecarga informacional e na necessidade de repensar estratégias de armazenamento, representação e preservação digital.

Nesse contexto, a ontologia assume papel estratégico uma vez que a sua aplicação implica na reflexão sobre o compartilhamento, a interoperabilidade e o uso da informação. Os autores destacam ainda que o movimento de ciência aberta representa uma tendência global voltada à integração, gestão e reutilização da informação científica, fortalecendo práticas científicas que asseguram o acesso à informação e o reuso dos dados de pesquisa (Castro; Simionato, 2020).

Em consonância com essa perspectiva, Santos (2022) observa que o movimento de ciência aberta tem transformado a prática científica, promovendo a transparência, produtividade e reprodutibilidade das pesquisas. A autora ressalta que o compartilhamento dos dados de pesquisa representa uma estratégia metodológica necessária para ampliar o alcance e o impacto da produção científica, além de fomentar o protagonismo e a valorização do pesquisador. Contudo, adverte que a transição para a ciência aberta é um processo multidimensional que demanda mudanças culturais e estruturais no modo de produzir e gerir o conhecimento.

Diante desse diagnóstico, Santos (2022) propõe para pesquisas futuras: a implementação e validação de políticas de dados abertos; o estudo de estratégias de gerenciamento e compartilhamento de dados; a integração de planos de gestão de dados ao fluxo de trabalho científico e a avaliação do impacto de iniciativas de acesso aberto no avanço da comunicação científica.

Nessa mesma direção, Souza (2023) enfatiza a importância das revisões sistemáticas do estado da arte como uma forma de fortalecer a visibilidade e a epistemologia da Ciência da Informação. O autor recomenda que futuras investigações: explorem as distinções conceituais do campo; desenvolvam

vocabulários controlados; aprofundem o uso de dados vinculados e de aplicativos de dados abertos além de estabelecer padrões e diretrizes para a catalogação de recursos digitais de modo a otimizar a descrição e o acesso à informação por meio de metadados estruturados.

Complementarmente, Medeiros e Vitoriano (2015) destacam que os avanços tecnológicos favorecem a apropriação dos fundamentos da bibliometria pela comunidade científica, o que amplia o interesse por métodos quantitativos aplicados à análise da produção científica. Essa tendência é retomada por Costa (2024) que propõe um procedimento estruturado para realizar análises bibliométricas em diferentes áreas do conhecimento. O autor aplica métodos bibliométricos à investigação da produção científica, das redes de colaboração científica e das associações temáticas no campo científico.

O autor sugere para estudos futuros: a expansão das análises bibliométricas; a integração de estruturas ontológicas voltadas ao compartilhamento do conhecimento científico e a exploração dos padrões de colaboração científica e associação temática de modo a oferecer uma compreensão aprofundada sobre o desenvolvimento e o impacto das pesquisas científicas (Costa, 2024).

### 1.3 PROBLEMATIZAÇÃO DA PESQUISA

Diante do exposto, observa-se que a literatura evidencia a necessidade de abordagens sistemáticas e interdisciplinares capazes de articular os aspectos teórico, metodológico e aplicado da ontologia na Ciência da Informação. Nesse contexto, a presente investigação adota a abordagem de análise bibliométrica como método de mapeamento científico, promovendo uma compreensão abrangente sobre o desenvolvimento conceitual do campo, as dinâmicas de colaboração científica e as tendências temáticas das práticas científicas.

Assim, a presente pesquisa parte da seguinte questão norteadora:

- **Como a produção científica de ontologia está representada no domínio da Ciência da Informação no cenário internacional, considerando os padrões de autoria, as redes de colaboração científica e as associações temáticas do campo científico?**

## 1.4 OBJETIVO GERAL DO ESTUDO

O presente estudo tem como objetivo geral realizar uma análise bibliométrica da produção científica de ontologia em Ciência da Informação no cenário internacional a fim de compreender o comportamento e as tendências do campo.

### 1.4.1 Objetivos Específicos do Estudo

- Mapear as atividades científicas relacionadas à ontologia na Ciência da Informação, considerando o acesso aberto das publicações indexadas na base de dados *Web of Science*.
- Descrever o panorama do campo, evidenciando a estrutura de produção, autoria e relacionamento.
- Apresentar os indicadores bibliométricos da produção científica, considerando os padrões de autoria, as redes de colaboração científica e as associações temáticas do campo.

## 1.5 JUSTIFICATIVA DO ESTUDO

Considerando que a Ciência da Informação foi uma das áreas a incorporar tardiamente o conceito de ontologia, a literatura evidencia que o termo ainda não se encontra plenamente consolidado no campo. Assim, estudos dedicados ao tema contribuem para fortalecer essa consolidação ao demonstrar que as ontologias reduzem ambiguidades da linguagem natural por meio de padrões terminológicos, vocabulários controlados e atribuição semântica aos termos, cujos aspectos favorecem a recuperação da informação em ambientes digitais (Santos, 2014).

Por isso, a ontologia tem se consolidado como um instrumento para a representação e a organização do conhecimento no âmbito da Ciência da Informação. As evidências científicas indicam sua ampla utilização em diferentes contextos, especialmente em ambientes digitais que demandam estruturas conceituais capazes de aprimorar a organização, recuperação e

interoperabilidade dos sistemas de informação. Diversos autores reconhecem que a aplicação da ontologia na Ciência da Informação contribui tanto para o fortalecimento das práticas de representação do conhecimento quanto para o desenvolvimento de metodologias voltadas à organização da informação (Almeida; Bax, 2003; Guarino, 1998; Gruber, 1993).

Ao ser compreendida como um modelo conceitual, a ontologia garante a comunicação padronizada e livre de ambiguidades. No contexto da web semântica, a sua aplicação possibilita a construção de vocabulários controlados e linguagens estruturadas para representar termos e conceitos. Entre seus principais usos na Ciência da Informação, destacam-se a integração de sistemas, a recuperação da informação, o desenvolvimento de sistemas informacionais e a representação e organização do conhecimento (Almeida, 2021). Tais elementos evidenciam a relevância da ontologia na Ciência da Informação e reforçam a necessidade de estudos que aprofundem a sua contribuição para o avanço do campo científico.

Emygdio, Almeida e Teixeira (2021) destacam que a ontologia promove melhorias significativas na qualidade das representações ao adotar a lógica descritiva cuja sintaxe favorece a explicitação e a validação contínua do conhecimento por inferência. Além disso, ampliam a capacidade da elaboração de consultas inteligentes, aumentando a transparência das entidades ontológicas, favorecendo a interoperabilidade semântica, auxiliando na identificação das demandas organizacionais e estimulando as oportunidades de inovação em produtos, processos e serviços, fornecendo subsídios para a avaliação da qualidade dos dados de pesquisa e para a capacitação de profissionais da informação.

Embora a ontologia seja considerada um artefato representacional, a sua aplicação ainda carece de uma base conceitual robusta. O autor argumenta que a comunidade científica deve aprofundar as investigações empíricas nas áreas de organização do conhecimento, bibliometria, entre outras, a fim de integrar as boas práticas de projeto ontológico com a interdisciplinaridade e o compartilhamento da informação, os quais são pilares necessários para a consolidação do conhecimento científico (Almeida, 2020).

Entretanto, observa-se que grande parte desses estudos se concentram em abordagens teóricas, conceituais ou tecnológicas relacionadas à definição, estruturação e implementação de modelos ontológicos. Apesar dos avanços, ainda há um número limitado de investigações que realizam análises quantitativas e sistematizadas da produção científica sobre o uso e a evolução do conceito de ontologia no campo. Essa lacuna limita a compreensão sobre a consolidação da ontologia enquanto área de investigação na Ciência da Informação. A literatura ainda demonstra que a ontologia do social, compreendida como o estudo sistemático da natureza e da estrutura da realidade social, contribui para este trabalho ao considerar o contexto da Ciência da Informação enquanto uma ciência social e interdisciplinar voltada ao compartilhamento da informação e a representação do conhecimento (Almeida, 2020).

A relevância deste estudo fundamenta-se, portanto, na necessidade de compreender a evolução e a configuração da produção científica de ontologia no domínio da Ciência da Informação, especialmente diante do papel que essa estrutura desempenha nos processos de representação, organização e recuperação do conhecimento em contextos informacionais complexos onde o crescimento expressivo das pesquisas sobre o tema evidencia a importância da ontologia como um instrumento teórico e aplicado ao mesmo tempo em que suscita a demanda por análises sistematizadas que possibilitam compreender como o campo está representado, quem são seus principais atores e de que forma o conhecimento tem sido produzido e disseminado.

Sob o ponto de vista teórico, este estudo justifica-se por contribuir para a consolidação epistemológica da área ao oferecer um panorama quantitativo e relacional da produção científica internacional de ontologia na Ciência da Informação. Essa abordagem possibilita identificar tendências, redes de colaboração científica e temas emergentes, fortalecendo a compreensão sobre o desenvolvimento conceitual e metodológico do campo.

Do ponto de vista metodológico, a pesquisa se apoia na análise bibliométrica, reconhecida por sua capacidade de mensurar e interpretar a atividade científica por meio de indicadores bibliométricos. Essa metodologia oferece uma base sólida para identificar padrões de autoria e institucionalidade,

bem como para analisar dinâmicas de formação e consolidação de redes científicas.

Santos (2022) ressalta a importância da política de gerenciamento de dados de pesquisa e enfatiza que a disponibilidade dos dados de pesquisa favorece a transparência e a otimização da comunicação científica. Segundo a autora, o movimento de acesso aberto amplia as oportunidades de colaboração, fortalecendo a validação de políticas de dados abertos e aumentando o alcance e a qualidade da ciência ao torná-la mais acessível, responsiva e eficiente. Assim, o compartilhamento dos dados de pesquisa se apresenta como uma estratégia metodológica que visa potencializar a comunicação científica.

Em termos práticos, o estudo se mostra relevante por fornecer subsídios à gestão da informação e do conhecimento em contextos acadêmicos e institucionais, favorecendo a tomada de decisão sobre políticas científicas, formação de parcerias e identificação de lacunas temáticas. Além disso, o mapeamento científico contribui para ampliar a visibilidade da produção científica, promovendo uma visão integrada e estratégica do desenvolvimento do campo científico no cenário internacional.

Sabe-se que estudos bibliométricos em áreas correlatas demonstram o potencial da abordagem bibliométrica para identificar padrões de autoria, redes de colaboração científica, evolução temporal e núcleos temáticos de pesquisa. Contudo, são escassas as pesquisas que aplicam tais métodos especificamente à produção científica referente a ontologia na Ciência da Informação, o que reforça a necessidade de uma análise bibliométrica que permita visualizar tendências, articulações e inter-relações entre pesquisadores, instituições e temas emergentes.

Esta pesquisa se justifica pela articulação entre a relevância teórica, contribuição metodológica e aplicabilidade prática, configurando-se como um esforço necessário para fortalecer e compreender a dinâmica científica que envolve o uso e o estudo da ontologia no contexto informacional contemporâneo que reforçam a importância das técnicas de organização da informação e representação do conhecimento ao aprimorar o tratamento dos dados de pesquisa e atuar na seleção, processamento, recuperação e disseminação da informação (Almeida; Bax, 2003).

## 2 COMPREENDENDO O CAMPO CIENTÍFICO DE ONTOLOGIA

O referencial teórico que fundamenta este estudo apresenta inicialmente o marco teórico e a definição conceitual de ontologia cuja compreensão evoluiu de uma disciplina filosófica dedicada ao estudo das entidades e de sua existência para um artefato representacional formal aplicado em diferentes domínios científicos.

Essa evolução histórica sustenta as tradições e linhas de pesquisa que consolidaram a ontologia como um campo interdisciplinar, articulando os fundamentos metafísicos, lógicos e computacionais na construção de modelos conceituais voltados à representação da realidade.

No âmbito da Ciência da Informação, tais avanços permitiram a incorporação da ontologia como um instrumento de organização e representação do conhecimento, atuando no controle terminológico, na padronização semântica e no aprimoramento dos processos de recuperação da informação em ambientes digitais.

Assim, a seção apresenta as bases conceituais, históricas e interdisciplinares necessárias para compreender o papel da ontologia na Ciência da Informação.

### 2.1 MARCO TEÓRICO E DEFINIÇÃO CONCEITUAL DE ONTOLOGIA

Gruber (1993) define ontologia como um esquema conceitual que fornece uma descrição lógica e compartilhada dos dados, funcionando como uma especificação de uma conceitualização no âmbito dos sistemas de informação. A definição proposta por Gruber (1993) recebeu contribuições críticas de Guarino e Giaretta (1995) que ampliaram o debate ao diferenciar perspectivas extensionais e intensionalistas. Enquanto Gruber (1993) adota uma abordagem extensional, Guarino e Giaretta (1995) defendem a dimensão intensional na qual a ontologia é compreendida como uma teoria lógica capaz de expressar uma estrutura semântica que codifica regras implícitas e legitima a organização de uma parcela da realidade.

Com base nessa perspectiva, Guarino e Giaretta (1995) apresentam as interpretações de ontologia e propõem uma distinção conceitual entre “Ontologia” x “ontologia”.

O Quadro 1 apresenta uma síntese das principais interpretações conceituais do termo “ontologia”, evidenciando a polissemia e a evolução do conceito em diferentes tradições teóricas.

Quadro 1 - Interpretações conceituais de ontologia

TERMO	CONCEITO
“Ontologia” x “ontologia”	1. É uma disciplina filosófica
	2. É um sistema conceitual informal
	3. É um sistema baseado em semântica formal
	4. É uma especificação de uma conceitualização
	5. É uma representação de um sistema conceitual via teoria lógica
	6. É um vocabulário usado por uma teoria lógica
	7. É uma especificação meta-nível de uma teoria lógica

Fonte: adaptado de Guarino e Giaretta (1995).

As definições reunidas mostram que a ontologia pode ser compreendida tanto em sua acepção filosófica, enquanto disciplina dedicada ao estudo do ser, quanto em abordagens formais e aplicadas, especialmente no âmbito da Ciência da Informação e da Computação. Destacam-se interpretações que a definem como sistema conceitual informacional, especificação formal de uma conceitualização, representação lógica do conhecimento e vocabulário controlado fundamentado em teorias lógicas. O quadro também evidencia níveis distintos de abstração, incluindo ontologias como especificações em meta-nível de teorias lógicas. Dessa forma, a sistematização apresentada contribui para a compreensão da complexidade conceitual do termo e fundamenta sua utilização em contextos interdisciplinares, particularmente nos estudos sobre organização, representação e modelagem do conhecimento.

O Quadro 2 apresenta a distinção conceitual entre os termos “Ontologia” e “ontologia”, conforme proposto por Guarino e Giaretta (1995), evidenciando diferenças quanto à natureza e ao nível de abstração de cada termo.

Quadro 2 - Distinção conceitual entre os termos de ontologia

TERMO	CONCEITO
“Ontologia”	É uma disciplina filosófica
“ontologia”	É um artefato

Fonte: adaptado de Guarino e Giaretta (1995).

Enquanto “Ontologia”, grafada com inicial maiúscula, refere-se à disciplina filosófica dedicada à investigação do ser e das categorias fundamentais da realidade, “ontologia”, em minúsculo, é compreendida como um artefato, isto é, uma construção formal utilizada para representar e organizar conceitos em domínios específicos. Essa distinção é relevante nos contextos da Ciência da Informação e da Computação, pois delimita o uso do termo entre sua base filosófica e suas aplicações práticas na modelagem, representação e interoperabilidade do conhecimento. Assim, o quadro contribui para reduzir ambiguidades terminológicas e fornece um referencial conceitual consistente para a análise e aplicação de ontologias no campo científico investigado.

Uschold e Gruninger (1996) definem ontologia como uma visão de mundo associada a um determinado domínio, enfatizando seu papel na explicitação das pressuposições compartilhadas que orientam a representação do conhecimento. Em continuidade a essa perspectiva, Vickery (1997) aprofunda a discussão ao apresentar os níveis de conhecimento envolvidos nas ontologias, compreendendo-as como expressões verbais do conhecimento e como elementos simbólicos da linguagem ao incluir conceitos, termos e suas relações estruturantes.

Avançando nesse debate, Guarino (1998) revisita a noção de conceitualização a partir de um enfoque intensional, propondo uma interpretação segundo a qual a ontologia é entendida como um artefato composto por um vocabulário estruturado destinado a descrever uma dada realidade. O autor

também introduz o conceito de ontologia aplicada concebida como uma síntese entre a tradição filosófica e as demandas práticas da sociedade da informação, o que consolida o caráter interdisciplinar do campo.

Na sequência desse amadurecimento conceitual, Sowa (1999) descreve ontologia como um catálogo de tipos de entidades consideradas existentes em um domínio sob a perspectiva do sujeito que utiliza determinada linguagem. Smith (2002) reforça essa compreensão ao destacar a função da ontologia na descrição, organização e sistematização do conhecimento enquanto Cunha (2002) evidencia sua contribuição para o estabelecimento de um entendimento comum e compartilhado do domínio, favorecendo a interação entre seus componentes.

Retomando e ampliando as formulações anteriores, Gruber (2003) define ontologia como uma especificação formal e explícita de uma conceitualização compartilhada, complementada posteriormente por um conjunto de princípios representacionais voltados à modelagem do domínio (Gruber, 2008).

A partir dessa base conceitual, Almeida e Bax (2003) descrevem a conceitualização como um modelo abstrato de representação de fenômenos reais e propõem uma classificação dos componentes para construção de ontologias.

O Quadro 3 apresenta a estrutura basilar de um modelo ontológico, sistematizando os principais componentes que sustentam a representação formal do conhecimento em ontologias

Quadro 3 - Estrutura basilar do modelo ontológico

CLASSIFICAÇÃO	DESCRIÇÃO
Classes	As classes são organizadas em uma taxonomia
Relações	As relações representam o tipo de interação entre os conceitos do domínio
Axiomas	Os axiomas são usados para modelar as sentenças verdadeiras
Instâncias	As instâncias são utilizadas para representar elementos específicos

Fonte: adaptado de Almeida e Bax (2003).

O quadro evidencia que as classes constituem a base conceitual do domínio, organizadas em uma taxonomia que explicita relações hierárquicas; as

relações descrevem os diferentes tipos de interação entre os conceitos; os axiomas estabelecem restrições e regras lógicas que garantem a coerência semântica do modelo; e as instâncias representam os elementos concretos ou particulares do domínio modelado. Dessa forma, o quadro oferece uma visão sintética e estruturada dos elementos essenciais de uma ontologia, contribuindo para a compreensão de sua arquitetura conceitual e de seu papel na organização e representação do conhecimento.

Além de delinear a estrutura basilar do processo ontológico, os autores Almeida e Bax (2003) avançam na discussão ao propor uma classificação abrangente das diferentes abordagens ontológicas, considerando múltiplas dimensões que orientam tanto a elaboração quanto a aplicação desses artefatos em sistemas de informação. Essa categorização contempla critérios relacionados à função, ao grau de formalismo, ao propósito de aplicação, à estrutura e ao conteúdo, evidenciando a diversidade teórica-metodológica que caracteriza o campo de estudo da ontologia.

O Quadro 4 apresenta uma classificação sistematizada das principais abordagens ontológicas descritas na literatura, evidenciando a diversidade conceitual e metodológica associada ao desenvolvimento e ao uso de ontologias.

Quadro 4 - Classificação das abordagens ontológicas

ABORDAGEM ONTOLÓGICA	CLASSIFICAÇÃO	DESCRIÇÃO
Função	ontologias de domínio	São reutilizáveis no domínio e fornecem vocabulário sobre conceitos, relacionamentos, atividades e regras que as governam
	ontologias de tarefa	Fornecem um vocabulário sistematizado de termos, especificando tarefas que podem ou não estar no mesmo domínio
	ontologias gerais	Incluem um vocabulário relacionado a coisas, eventos, tempo, espaço, causalidade, comportamento, funções etc...
Grau de Formalismo	ontologias altamente informais	Expressa livremente em linguagem natural
	ontologias semi-informais	Expressa em linguagem natural de forma restrita e estruturada

ABORDAGEM ONTOLÓGICA	CLASSIFICAÇÃO	DESCRIÇÃO
	ontologias semi formais	Expressa em uma linguagem artificial definida formalmente
	ontologia rigorosamente formal	Os termos são definidos com semântica formal, teoremas e provas
Aplicação	ontologias de autoria neutra	Um aplicativo é escrito em uma única língua e depois convertido para uso em diversos sistemas, utilizando-se as informações
	ontologias como especificação	Cria-se uma ontologia para um domínio, a qual é usada para documentação e manutenção no desenvolvimento de <i>softwares</i>
	ontologias de acesso comum à informação	Quando o vocabulário é inacessível a ontologia torna a informação inteligível, proporcionando conhecimento compartilhado dos termos
Estrutura	ontologias de alto nível	Descrevem conceitos gerais relacionados a todos os elementos da ontologia (espaço, tempo, matéria, objeto, evento, ação, etc.), os quais são independentes do problema ou domínio
	ontologias de domínio	Descrevem o vocabulário relacionado a um domínio
	ontologias de tarefa	Descrevem uma tarefa ou atividade
Conteúdo	ontologias terminológicas	Especificam termos que serão usados para representar o conhecimento em um domínio
	ontologias de informação	Especificam a estrutura de registros de banco de dados
	ontologias de modelagem do conhecimento	Especificam conceitualizações do conhecimento pois têm uma estrutura interna semanticamente rica e são refinadas para uso no domínio do conhecimento que descrevem
	ontologias de aplicação	Contêm as definições necessárias para modelar o conhecimento em uma aplicação
	ontologias de domínio	Expressam conceitualizações que são específicas para um determinado domínio do conhecimento
	ontologias genéricas	Similares às ontologias de domínio, mas, os conceitos que as definem são considerados genérico e comuns a vários campos
	ontologias de representação	Explicam as conceitualizações que estão por trás dos formalismos de representação do conhecimento

Fonte: adaptado de Almeida e Bax (2003).

A organização do quadro contempla diferentes critérios analíticos — função, grau de formalismo, aplicação, estrutura e conteúdo — permitindo compreender como as ontologias podem variar desde representações altamente

informais, expressas em linguagem natural, até modelos rigorosamente formais baseados em semântica lógica. Essa tipologia evidencia o caráter multifacetado das ontologias, destacando seu papel tanto como instrumentos de representação e organização do conhecimento quanto como artefatos técnicos aplicados ao desenvolvimento de sistemas de informação, interoperabilidade e compartilhamento semântico em distintos domínios do conhecimento.

Almeida e Bax (2003) discutem sobre a diversidade das abordagens ontológicas que evidenciam a complexidade inerente ao campo e impõe desafios à formulação de metodologias de avaliação formalmente consolidadas. De acordo com os autores a variedade de perspectivas, os níveis de formalismo e as múltiplas finalidades das ontologias resultam em questionamentos sobre os mecanismos de interação, os formalismos de representação do conhecimento e os critérios técnicos empregados na avaliação desses artefatos. Esse cenário de multiplicidade conceitual é reforçado por Figueiredo e Almeida (2017) ao apontarem que o termo ontologia assume diferentes interpretações e aplicações em uma ampla gama de domínios, o que reafirma seu caráter interdisciplinar.

O Quadro 5 apresenta as principais vertentes conceituais de ontologia, sistematizando as diferentes interpretações do termo de acordo com seu propósito teórico e aplicado.

Quadro 5 - Vertentes conceituais de ontologia

INTERPRETAÇÃO CONCEITUAL	PROPÓSITO
ontologia como disciplina	Nomear a ontologia como disciplina filosófica e os princípios metafísicos subjacentes
ontologia como artefato	Se refere a qualquer artefato computacional construído com alguma linguagem declarativa de representação
modelo ontológico	Se refere a qualquer ontologia como artefato que faz uso dos princípios da ontologia como disciplina para obter boa fundamentação na modelagem

Fonte: adaptado de Almeida (2021).

O quadro distingue a ontologia compreendida como disciplina filosófica, voltada à reflexão sobre os princípios metafísicos fundamentais do ser; a ontologia como artefato, entendida como um produto computacional expresso em linguagens declarativas de representação do conhecimento; e o modelo ontológico, que articula fundamentos da ontologia filosófica à prática da

modelagem conceitual. Essa distinção evidencia a polissemia do conceito de ontologia e contribui para clarificar seu uso no âmbito da Ciência da Informação e das áreas afins, especialmente no desenvolvimento de sistemas de organização e representação do conhecimento.

Nesse contexto, Almeida (2021) sistematiza as vertentes conceituais de ontologia que orientam a compreensão contemporânea do termo, oferecendo um arcabouço organizado para distinguir ontologia enquanto uma disciplina filosófica, um artefato computacional e um modelo ontológico fundamentado.

## 2.2 TRADIÇÃO E LINHAS DE PESQUISA EM ONTOLOGIA

A ontologia, em sua concepção original, constitui um ramo da Filosofia dedicado ao estudo da existência e das formas mais gerais do ser. Nessa perspectiva, ela é concebida como um sistema de categorias que busca compreender a realidade, identificar o que existe e descrever as propriedades essenciais dos entes. Assim, a ontologia representa um conjunto de princípios de natureza metafísica destinados à fundamentação teórica da realidade, abrangendo noções trans categoriais que orientam processos de investigação e construção de teorias (Almeida, 2020).

Com o avanço das tecnologias da informação, o termo foi incorporado pela Ciência da Computação, consolidando-se como um artefato representacional utilizado para descrever formalmente um domínio, reduzir ambiguidades e permitir inferências automatizadas. Nesse campo, a ontologia é concebida como uma teoria lógica e um artefato representacional, construído a partir de linguagens formais de representação com o propósito de estabelecer vocabulários estruturados que possibilitam a modelagem e o processamento de conhecimento em sistemas de informação (Almeida, 2020).

Embora possuam forte caráter computacional, os estudos sobre ontologia transcendem a dimensão tecnológica (Almeida; Mendonça; Aganette, 2013). No âmbito da representação do conhecimento, a ontologia assume a forma de uma estrutura composta por conceitos articulados por meio de um vocabulário lógico, constituindo uma teoria que descreve fatos, relações e regras que organizam uma determinada realidade.

Dessa forma, as ontologias são aplicadas à modelagem conceitual e representação do conhecimento, uma vez que se fundamentam em princípios ontológicos voltados à compreensão e à modelagem de fenômenos, fornecendo descrições sistematizadas das entidades e de suas características em processos de construção de modelos conceituais. Em síntese, tratam-se de declarações formalizadas por meio de linguagens de representação elaboradas para execução por mecanismos de inferência automatizada (Almeida, 2014).

A incorporação do conceito pela Ciência da Informação reflete essa expansão uma vez que, nesse domínio, a ontologia é compreendida como um artefato representacional definido tanto como teoria informal quanto como sistema conceitual destinado à compreensão de um domínio, à classificação de termos e à construção de vocabulários controlados orientados à recuperação da informação (Almeida, 2020). Assim, o foco da área recai sobre a manipulação de dados que descrevem entidades em diversos contextos, nos quais os princípios ontológicos funcionam como base para estruturar sistemas de categorização e modelos de representação da informação (Almeida, 2014).

A abordagem interdisciplinar da ontologia evidencia sua relevância para o campo da Ciência da Informação. Os estudos não se limitam à representação da informação, mas, abrangem também sua dimensão social ao considerar a comunidade científica como produtora, mediadora e usuária de informação. Nesse sentido, a Ciência da Informação possui potencial para explicar o ambiente social que envolve a análise realizada pelo cientista da informação, articulando princípios ontológicos que se encontram presentes, de forma convergente, na Filosofia, na Ciência da Computação e na própria Ciência da Informação (Almeida, 2014).

Para que essa articulação interdisciplinar se fortaleça é necessário que os princípios ontológicos estabeleçam claramente quais entidades podem existir no mundo e como tais entidades coexistem de acordo com as evidências empíricas. Por um lado, a Filosofia contribui ao oferecer um teste sistemático desses princípios, possibilitando novas interpretações teóricas e epistemológicas. Por outro, as Ciências Aplicadas reforçam a necessidade de fundamentos sólidos que sustentem investigações empíricas robustas (Almeida, 2014).

O Quadro 6 evidencia a interdisciplinaridade das abordagens conceituais de ontologia, ao relacionar distintas interpretações do conceito aos campos do conhecimento nos quais ele é mobilizado, bem como aos seus respectivos propósitos.

Quadro 6 - Interdisciplinaridade das abordagens conceituais de ontologia

DISTINÇÃO	CAMPO	CONCEITO	PROPÓSITO
Ontologia como uma disciplina	Filosofia	Ontologia como um sistema de categorias	Entender a realidade, as coisas que existem e suas características
ontologia como um artefato	Ciência da Computação	ontologia como uma teoria baseada em lógica	Entender um domínio e reduzi-lo a modelos
		ontologia como um artefato de <i>software</i>	Criar um vocabulário para representação em sistemas e gerar inferências
	Ciência da Informação	ontologia como um sistema conceitual informal	Criar vocabulário controlado para recuperação da informação

Fonte: adaptado de Almeida (2014).

A ontologia, enquanto disciplina filosófica, está associada ao campo da Filosofia e orienta-se para a compreensão da realidade, das entidades existentes e de suas características fundamentais. Já no âmbito da Ciência da Computação, a ontologia é concebida tanto como uma teoria baseada em lógica quanto como um artefato de software, com a finalidade de modelar domínios, estruturar vocabulários formais e possibilitar representação computacional e inferência automática. Na Ciência da Informação, por sua vez, a ontologia assume a forma de um sistema conceitual informal, direcionado à criação de vocabulários controlados que apoiam a organização, a recuperação e o uso da informação.

Essa articulação entre campos reforça o caráter transversal da ontologia e sua relevância teórica e aplicada em diferentes contextos científicos. A interação entre essas perspectivas demonstra que há amplo espaço para o desenvolvimento de pesquisas interdisciplinares em ontologia (Almeida, 2014). Essa análise é reforçada por estudos que caracterizam o campo como um espaço de coexistência de múltiplas abordagens conceituais, resultando em interpretações diversificadas e complementares (Almeida, 2021).

### 2.3 ONTOLOGIA NA PERSPECTIVA DA CIÊNCIA DA INFORMAÇÃO

De acordo com Borko (1968) a Ciência da Informação é definida como a disciplina que investiga as propriedades e o comportamento da informação, bem como as forças que governam seu fluxo e os significados associados ao processamento informacional, visando sempre à acessibilidade e usabilidade. Desse modo, dedica-se ao estudo do corpo de conhecimentos relacionados à origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação.

Sob essa perspectiva, Borko (1968) caracteriza a Ciência da Informação como uma ciência essencialmente interdisciplinar. Essa natureza manifesta-se em dois eixos complementares: o da ciência pura e o da ciência aplicada. No âmbito da ciência pura, a área se ocupa da investigação de seu objeto sem considerar, inicialmente, sua aplicação prática. Assim, busca construir um *corpus* teórico consistente capaz de fundamentar e aprimorar instituições e processos voltados à acumulação e à transmissão do conhecimento.

Por outro lado, quando observada como ciência aplicada, a Ciência da Informação se volta ao desenvolvimento de serviços, produtos e soluções informacionais. Esse enfoque evidencia sua função social e prática, permitindo que a disciplina opere como um campo de pesquisa aplicada comprometido com a resolução de problemas reais relacionados ao uso, à circulação e comunicação da informação no contexto científico e social (Borko, 1968).

Assim, a articulação entre ciência pura e ciência aplicada reafirma o caráter híbrido e multifacetado da área, sustentando sua consolidação como um domínio científico voltado tanto à compreensão teórica da informação quanto à produção de intervenções que qualifiquem o ambiente informacional e os processos de comunicação do conhecimento científico (Borko, 1968).

Buckland (1991) identifica três ambiguidades associadas ao termo “informação”: informação como processo, entendida como o ato de informar; informação como conhecimento, correspondente ao ato de transmitir conhecimento sobre determinado objeto e informação como coisa, referente aos objetos informativos que carregam ou materializam o conteúdo informacional.

No presente trabalho, adotamos essa tripla concepção proposta por Buckland (1991), considerando cada dimensão de acordo com sua função analítica. A informação como processo é mobilizada para compreender o ato de informar e sua dinâmica comunicacional. A informação como coisa é empregada para abordar a representação do conhecimento, especialmente no contexto dos sistemas de informação, nos quais os objetos informativos desempenham o papel de evidências científicas. Esse entendimento torna-se ainda mais relevante diante dos avanços das tecnologias da informação que ampliam o escopo para criação, manipulação e uso da informação como coisa, reconhecendo que seu caráter informativo é sempre situacional. Por fim, a informação como conhecimento é considerada ao quantificar e interpretar os dados de pesquisa, o que permite avaliar a produção científica e compreender os fluxos de conhecimento do campo analisado. Assim, a perspectiva tripartida de Buckland (1991) fornece o suporte conceitual necessário para operacionalizar o termo “informação” nesta pesquisa.

De acordo com Saracevic (1995) a Ciência da Informação é uma disciplina indissociavelmente ligada às tecnologias da informação. Essa relação intrínseca a caracteriza, por natureza, como um campo interdisciplinar que participa ativa e deliberadamente da evolução da sociedade da informação, contexto em que as dimensões sociais e humanas assumem um papel estratégico.

A literatura evidencia que a Ciência da Informação se dedica tanto às questões científicas quanto à prática profissional voltada para os problemas da comunicação do conhecimento e de seus registros, considerando os contextos sociais, institucionais e individuais referentes ao uso da informação. Nesse sentido, o autor debate sobre as relações existentes entre a informação como fenômeno e a comunicação como processo, uma vez que ambos os conceitos interagem de maneira complexa onde mutuamente influenciam-se (Saracevic, 1996).

De acordo com o autor, quando se considera o imperativo tecnológico que impulsiona o desenvolvimento e a aplicação de produtos e serviços de informação, observa-se que os papéis econômicos e sociais conferem à área importância estratégica tanto para o desenvolvimento nacional quanto para o avanço científico. O autor menciona que tal importância decorre da capacidade

da informação sustentar processos de inovação, fortalecer a produção de conhecimento e gerar vantagens competitivas na sociedade da informação. Conseqüentemente, as relações interdisciplinares tornam-se cada vez mais visíveis em aplicações teóricas e experimentais no campo, reforçando a natureza multifacetada e socialmente relevante da Ciência da Informação (Saracevic, 1996).

Macias-Chapula (1998) considera a ciência como um sistema social. Nessa perspectiva, a ciência é compreendida como um processo social no qual ações, práticas e comportamentos científicos são determinados pelo contexto em que se inserem. Enquanto sistema social, a ciência cumpre funções essenciais, tais como a disseminação do conhecimento, a preservação de padrões científicos e a atribuição de créditos e reconhecimento à comunidade científica. Sendo a produção de conhecimento a essência da atividade científica, logo, esse conhecimento precisa ser transformado em informação acessível para circulação e uso pela comunidade científica. Assim, a informação enquanto conhecimento, representa e evidencia as atividades científicas, contribuindo diretamente para o desenvolvimento e a avaliação do campo científico.

Capurro e Hjørland (2003) discutem a informação como um conceito interdisciplinar no âmbito da Ciência da Informação, destacando que seu significado é sempre dependente de contexto e associado a fenômenos específicos. Por se tratar de um conceito subjetivo, os critérios que definem os objetos informativos são moldados por processos socioculturais e científicos que produzem uma compreensão situacional da informação. Assim, a informação passa a ser entendida como um reflexo direto da função social desempenhada pelos sistemas de informação, cuja dinâmica é condicionada por valores, práticas e estruturas culturais que constituem determinado ambiente informacional.

Nessa perspectiva, a Ciência da Informação é reconhecida como uma ciência social aplicada que emergiu em meados do século XX. Embora inicialmente vinculada à computação, a disciplina ampliou progressivamente seu escopo e a partir da década de 1970, consolidou sua inserção no campo das ciências sociais (Capurro; Hjørland, 2003).

Os primeiros estudos, orientados pelos modelos positivista e funcionalista, abordavam a realidade social de maneira predominantemente estatística e quantitativa. Contudo, ainda naquela década, emergiu a compreensão de que o estudo da informação deveria incorporar a historicidade dos sujeitos e dos objetos, reconhecendo a totalidade dos fenômenos sociais e as tensões inerentes às relações de poder. Essa mudança de perspectiva, ancorada em uma abordagem crítica, trouxe à tona a importância do conflito, do contexto cultural e da inserção social dos fatos informacionais como elementos fundamentais para explicar os fenômenos informacionais (Capurro; Hjørland, 2003).

Nesse movimento, o enfoque interpretativo passou a orientar novas abordagens, permitindo compreender a informação como um processo que se configura de acordo com o sujeito, suas experiências e seus contextos socioculturais. A adoção desse enfoque consolidou a aceitação da natureza interdisciplinar da Ciência da Informação e marcou sua transição para uma postura epistemológica alinhada ao pensamento pós-moderno (Araújo, 2003).

Sabe-se que as ontologias constituem um tipo de sistema de organização do conhecimento cujo objetivo é estruturar e representar formalmente o conhecimento em determinados domínios (Sörgel, 1997). No âmbito da *web semântica*, a ontologia ganha relevância à medida que sustenta sistemas computacionais baseados em linguagens como a *Web Ontology Language* (OWL), permitindo a realização de inferências automáticas na *web* (Berners-Lee *et al.*, 2001). Assim, desempenham papel relevante na recuperação da informação ao organizar conteúdos provenientes de diferentes fontes de dados e contribuir para a disposição e o uso do conhecimento em sistemas informacionais (Almeida; Bax, 2003).

A ontologia aplicada articula métodos filosóficos a teorias e práticas das ciências aplicadas, visando atender às necessidades de organização do conhecimento. Ela diferencia-se dos tesouros tradicionais ao possibilitar inferências automáticas e maior rigor relacional, representando uma evolução dos sistemas de linguagens documentárias (Almeida *et al.*, 2005). Embora tesouros e ontologias compartilhem elementos como termos e relações, a natureza dessas relações difere: nos tesouros, são menos específicas e não

permitem raciocínios automatizados, o que limita o potencial inferencial (Marcondes; Campos, 2008).

Ao observar a ontologia sob a perspectiva da Ciência da Informação, reconhece a sua estreita relação com as práticas de gestão do conhecimento. A natureza interdisciplinar da área favorece estudos ontológicos e promove o engajamento em projetos colaborativos uma vez que as ontologias estabelecem um entendimento comum sobre um domínio e oferecem uma estrutura semântica padronizada para uso compartilhado (Marcondes; Campos, 2008).

Nessa direção, a Ciência da Informação contribui significativamente para o desenvolvimento ontológico, especialmente por sua preocupação com a compreensão semântica, a representação rigorosa e a implementação de inferências automatizadas em sistemas de informação. Porém, a literatura também destaca a insuficiência de controle e rigor nos processos de construção de algumas ontologias, diferenciando-as dos vocabulários controlados que são tradicionalmente normalizados (Silva; Souza; Almeida, 2008).

O conceito de ontologia no contexto da organização e recuperação da informação está fortemente vinculado à *web semântica* uma vez que ontologias definem conceitos e relações de forma consensual dentro de uma área do conhecimento com o propósito de facilitar a interoperabilidade entre sistemas (Brascher; Carlan, 2010). Nessa abordagem, as ontologias constituem estruturas representacionais baseadas em vocabulários formais que possibilitam inferências automáticas e enriquecem os processos de busca e recuperação da informação (Almeida; Souza; Coelho, 2010).

Considerando que a modelagem conceitual integra o processo de desenvolvimento de sistemas de informação, descrevendo partes da realidade e representando processos em contextos sociais específicos, o estudo das ontologias caracteriza-se pela convergência de múltiplas abordagens disciplinares. Os profissionais da Ciência da Informação dispõem de uma formação que privilegia o conhecimento sobre informação, o que os torna necessários para atividades de modelagem ontológica. Assim, a interdisciplinaridade, a abrangência conceitual e o caráter multifacetado da área favorecem sua contribuição ao desenvolvimento e aplicação de ontologias (Almeida; Oliveira; Coelho, 2010).

O termo “ontologia” tem sido amplamente utilizado em distintos campos científicos e é frequentemente associado a artefatos de representação do conhecimento, especialmente no desenvolvimento de sistemas especialistas. A literatura evidencia dois ramos principais de pesquisa sob essa designação: a ontologia aplicada e a *web semântica* (Almeida; Mendonça; Aganette, 2013).

Embora ontologias sejam teorias formais sobre entidades do mundo real e não linguagens, as linguagens de representação podem ser adotadas para sua especificação e aplicação em sistemas de recuperação da informação, utilizando princípios oriundos das linguagens documentárias (Almeida; Mendonça; Aganette, 2013).

A literatura confirma que a Ciência da Informação reconhece a utilidade e relevância das ontologias, evidenciando aproximações com tesouros e taxonomias devido às preocupações comuns com o controle terminológico e a organização do conhecimento. Contudo, as características de padronização e rigor das ontologias não devem ser confundidas com os princípios dos vocabulários controlados (Almeida, 2013).

Na perspectiva da área, a ontologia é um instrumento de representação que atua sobretudo no controle terminológico (Santos, 2014). A partir da década de 1990 seu uso cresceu substancialmente, impulsionado pela *web semântica* e pela necessidade de interoperabilidade entre os sistemas de informação, além do compartilhamento de dados estruturados (Santos, 2014).

A literatura da Ciência da Informação fornece bases importantes para o desenvolvimento de ontologias voltadas à representação do conhecimento. Estudos evidenciam que o campo constitui ambiente fértil para o avanço das pesquisas ontológicas (Café; Santos; Barros, 2015), reforçando sua contribuição para o domínio da Ciência da Informação (Figueiredo; Almeida, 2017) em função de seu interesse pelas relações semânticas (Tomoyose *et al.*, 2018).

Dessa forma, a ontologia configura-se como instrumento para a representação da informação em ambientes digitais não apenas pela descrição de conteúdos, mas, também pela representação dos próprios recursos informacionais. Essa perspectiva amplia seu papel no delineamento e na construção ontológica mediante os processos de modelagem conceitual que

asseguram a consistência, a padronização das relações e a interoperabilidade dos sistemas informacionais (Castro; Simionato, 2020).

O Quadro 7 apresenta os princípios ontológicos que orientam a concepção, o desenvolvimento e a manutenção de ontologias científicas, evidenciando fundamentos teóricos e metodológicos essenciais para sua consistência e aplicabilidade.

Quadro 7 - Princípios ontológicos

PRINCÍPIOS ONTOLÓGICOS	DESCRIÇÃO
Realismo	O objetivo da ontologia é descrever uma realidade
Perspectivismo	É necessária uma multiplicidade de ontologias para abranger os variados aspectos da realidade
Falibilismo	Implica na necessidade de estabelecer estratégias de versionamento para garantir aderência às novas descobertas científicas
Adequatismo	Os artefatos ontológicos devem ser desenvolvidos pautados pela interoperabilidade com outros artefatos
Reutilização	As ontologias existentes são consideradas referência e devem ser prioritariamente reutilizadas
Utilidade	Remete à revisão da ontologia em função de avanços científicos
Atualização	Ontologias científicas sempre estão sujeitas à atualização a partir de avanços no conhecimento
Facilidade	Orienta a construção de um artefato a partir dos casos mais simples para entendimento e posterior avanço para aspectos mais complexos do domínio

Fonte: adaptado de Emygdio, Almeida e Teixeira (2021).

Entre esses princípios, destacam-se o realismo, que enfatiza a descrição de uma realidade objetiva; o perspectivismo, que reconhece a necessidade de múltiplas ontologias para contemplar diferentes aspectos da realidade; e o falibilismo, que admite a natureza provisória do conhecimento científico e a necessidade de versionamento contínuo. O quadro também ressalta princípios operacionais, como adequatismo, reutilização e interoperabilidade, que reforçam a integração entre ontologias e outros artefatos conceituais, além da utilidade, atualização e facilidade, voltadas à evolução contínua, aplicabilidade prática e compreensão progressiva do domínio modelado. Em conjunto, esses princípios configuram um referencial normativo que assegura a qualidade, a sustentabilidade e a relevância das ontologias no contexto científico e informacional.

Os princípios ontológicos apoiam a construção de estruturas categóricas para a representação do conhecimento. Entretanto, a atuação da ontologia na Ciência da Informação não se limita à representação temática pois inclui a representação do recurso informacional enquanto objeto pertencente a uma comunidade de usuários. Desse modo, a ontologia se constitui como objeto de pesquisa da área devido ao seu potencial para explicar o ambiente social e os contextos epistemológicos nos quais uma comunidade científica produz conhecimento (Almeida, 2013).

A incorporação da ontologia aplicada ao corpo de conhecimento da Ciência da Informação contribui para ampliar o reconhecimento da área no cenário científico uma vez que consolida sua atuação na representação e organização do conhecimento. A literatura demonstra que as ontologias constituem elementos de arquitetura semântica, funcionando como pressupostos para a recuperação da informação em ambientes digitais. Nesse sentido, desempenham um papel estratégico ao permitir que técnicas tradicionais de organização e representação do conhecimento sejam transportadas para sistemas automatizados por meio de princípios ontológicos que atuam na atenuação de inconsistências semânticas presentes no domínio do conhecimento que possibilitam o raciocínio automático do conteúdo representado, potencializando a interoperabilidade e a precisão informacional (Emygdio, Almeida; Teixeira, 2021).

### **3 BIBLIOMETRIA E ESTUDO MÉTRICO DA INFORMAÇÃO**

A seção dedicada à bibliometria e estudo métrico da informação apresenta o marco teórico e a definição conceitual que fundamenta a análise quantitativa da produção científica.

A bibliometria, entendida como o estudo dos aspectos quantitativos da informação registrada, constitui um método consolidado de mapeamento científico que permite identificar padrões, dinâmicas e tendências em campos do conhecimento científico.

Sustentada por princípios estatísticos aplicados às unidades de informação, a bibliometria utiliza indicadores bibliométricos para mensurar as atividades científicas.

Assim, a seção contextualiza os fundamentos teóricos e metodológicos que orientam a análise bibliométrica adotada na pesquisa, ressaltando sua relevância para a compreensão da estrutura e evolução do campo investigado.

#### **3.1 MARCO TEÓRICO E DEFINIÇÃO CONCEITUAL DE BIBLIOMETRIA**

Os estudos sobre bibliometria têm suas raízes nas primeiras investigações estatísticas aplicadas à produção científica. O trabalho pioneiro de Cole e Eales (1917) marcou esse movimento ao empregar métodos quantitativos na análise de uma bibliografia especializada, dando continuidade a uma tradição que já se esboçava desde 1890, quando surgiram as primeiras referências à chamada estatística bibliográfica. Em 1923, Hulme retomou e consolidou essa abordagem ao cunhar o termo bibliografia estatística, reconhecendo explicitamente a contribuição de Cole e Eales (1917) para o desenvolvimento do campo.

A evolução terminológica e conceitual da área ganhou novo impulso com Paul Otlet (1934) que introduziu formalmente o termo bibliometria em seu Tratado da Documentação, empregando-o para designar medições relativas aos livros e documentos, especialmente no que se refere à dimensão física dessas unidades.

Pouco depois, Ranganathan (1948) ampliou a perspectiva ao propor a biblioteconomia, incluindo o uso de métodos estatísticos à biblioteconomia e sinalizando a expansão das técnicas quantitativas para diferentes dimensões da organização do conhecimento.

A consolidação da bibliometria como campo para a Ciência da Informação avançou significativamente com as contribuições de Price (1964) que demonstrou o potencial do tratamento estatístico da informação para fortalecer as bases teóricas da área e estabelecer vínculos evidentes entre o conhecimento científico e sua manifestação documental. Nesse movimento de afirmação conceitual, a definição contemporânea da bibliometria foi estabelecida por Pritchard (1969) que a caracterizou como o estudo destinado a quantificar os processos de comunicação científica mediante métodos matemáticos e estatísticos aplicados a livros e demais meios de comunicação.

Nesse sentido, a bibliometria passou a assumir um papel estrutural no estudo da atividade científica ao introduzir procedimentos sistemáticos, transparentes e reproduzíveis de análise. Autores como Crane (1972) reforçam essa tendência ao destacar que os métodos bibliométricos possibilitam a identificação de padrões e tendências no comportamento da comunidade científica, o que a torna relevante diante do crescente volume de informação produzida globalmente.

Braga (1973) e Price (1976) destacam que a análise quantitativa da produção científica possibilita uma visão abrangente e objetiva da atividade científica, permitindo identificar fluxos de produção, circulação e uso da informação em diferentes contextos. Assim, a bibliometria afirma-se como uma ferramenta necessária para mapear a evolução dos domínios do conhecimento, revelar dinâmicas de colaboração científica e compreender os processos de consolidação de áreas emergentes.

No Brasil, o campo ganhou impulso a partir da década de 1970, com a oferta da disciplina de processamento de dados na documentação no curso de mestrado do então Instituto Brasileiro de Bibliografia e Documentação (IBBD), atual Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Essa iniciativa fomentou as primeiras aplicações sistemáticas da bibliometria no país,

contribuindo para a formação de pesquisadores e para a incorporação definitiva das técnicas métricas ao campo da Ciência da Informação no Brasil.

No panorama brasileiro, observa-se a consolidação das pesquisas métricas e a atuação consistente de grupos especializados que vêm ampliando as metodologias e aprofundando a capacidade avaliativa da área (Noronha; Maricato, 2008). Em escala internacional, a bibliometria se fortalece como abordagem na avaliação quantitativa da produção científica, sendo reconhecida como um conjunto de métodos voltados ao estudo de textos e informações em grandes bases de dados (Cobo et al., 2010). A literatura nacional e internacional evidencia a amplitude do uso e das aplicações da bibliometria.

Segundo Pinto et al. (2012), a função basilar da bibliometria é medir e analisar a atividade científica e social, produzindo indicadores que refletem padrões de comportamento e impacto das pesquisas científicas. Vanti (2002) destaca o potencial da bibliometria para mensurar fluxos de informação, padrões de comunicação científica, núcleos de periódicos, obsolescência da literatura, produtividade de autores e instituições, padrões de colaboração, co-citação, desempenho de sistemas de informação e emergência de novos temas científicos.

Nesse movimento, Guedes e Borschiver (2005) reafirmam que a bibliometria atua como campo de estudo e método de análise, produzindo indicadores que subsidiam a gestão da informação e do conhecimento. Bufrem e Prates (2005) complementam essa perspectiva ao enfatizar sua relevância na compreensão dos processos quantitativos de produção, disseminação e uso da informação em um contexto em que o conhecimento se torna elemento estratégico para o desenvolvimento econômico e social da comunidade científica (Araújo, 2006).

Portanto, os estudos métricos da informação, enquanto subárea consolidada da Ciência da Informação, integram esse conjunto de abordagens ao se dedicarem à identificação, mensuração e avaliação da informação e de seus impactos. Curty e Delbianco (2020) ressaltam que ao empregar técnicas quantitativas para avaliar a produção científica, os estudos aprofundam a compreensão dos processos de geração, circulação e uso do conhecimento nos mais diversos canais informacionais. Por fim, observa-se na literatura um

fortalecimento crescente da bibliometria no Brasil, evidenciado pela ampliação de pesquisas que utilizam indicadores métricos para revelar padrões de produtividade, colaboração e impacto científico. Nesse cenário, as métricas de informação e comunicação assumem papel estruturante tanto metodológico quanto disciplinar no âmbito da Ciência da Informação (Pinheiro, 2021).

### 3.2 BIBLIOMETRIA: MÉTODO DE MAPEAMENTO CIENTÍFICO

Hjørland (2013) define a bibliometria como um campo interdisciplinar, uma vez que os estudos bibliométricos favorecem o mapeamento e a visualização de domínios do conhecimento. Nesse sentido, os mapas bibliométricos constituem uma contribuição relevante aos sistemas de organização do conhecimento, pois permitem representar a estrutura cognitiva de um campo científico. Por outro lado, a análise de citações é compreendida como uma abordagem social e dinâmica para a organização do conhecimento, considerando que uma disciplina acadêmica é simultaneamente um corpo de conhecimento intelectual e uma unidade social baseada na cooperação e na consolidação de padrões científicos. Dessa forma, o mapeamento científico destaca a importância dos métodos bibliométricos para evidenciar o desenvolvimento e a evolução de um campo científico.

Nessa perspectiva, Zupic e Cater (2015) destacam que os métodos bibliométricos revelam grande potencial para a confirmação quantitativa de categorias derivadas da produção científica, além de possibilitarem a exploração do cenário de pesquisa e a identificação do panorama do campo científico. Os autores ressaltam ainda que, ao promover conexões entre pesquisadores, tais métodos favorecem a colaboração científica, reforçando o caráter social do conhecimento.

Observa-se que os estudos bibliométricos têm sido aplicados em diversas áreas do conhecimento, o que evidencia sua interdisciplinaridade e sua aplicação prática como método de pesquisa. Essa característica contribui diretamente para a geração de novos conhecimentos e para o avanço científico (Medeiros; Vitoriano, 2015). A literatura aponta a bibliometria como uma ferramenta necessária para determinar as características estruturais e dinâmicas

de um campo científico ao estudar os aspectos quantitativos dos processos de produção, disseminação e uso da informação (Figueiredo; Almeida, 2017).

Nesse contexto, Curty e Delbianco (2020) conceituam a bibliometria como um método que aplica análises estatísticas à quantificação e ao monitoramento da produção, armazenamento, circulação e uso das informações registradas cujo objeto de estudo são as produções bibliográficas. As autoras ressaltam que as leis bibliométricas são fundamentais para a consolidação da área e para o aprofundamento das discussões científicas, sendo elas:

- Lei da Produtividade Científica de Autores, proposta por Lotka (1926)
- Lei da Dispersão Bibliográfica, de Bradford (1934)
- Lei do Menor Esforço, formulada por Zipf (1949)

Sob essa ótica, os estudos métricos da informação são compreendidos como um conjunto de pesquisas voltadas à avaliação do que é produzido por comunidades científicas. Assim, estes estudos englobam técnicas e métodos de mensuração e avaliação da produção científica que contribuem para a compreensão dos processos de construção e evolução do campo científico. Ao envolver questões interdisciplinares, tais estudos se relacionam com diferentes áreas do conhecimento (Curty; Delbianco, 2020).

O mapeamento científico, por sua vez, é utilizado para representar a estrutura cognitiva de um campo de pesquisa (Cobo *et al.*, 2010), constituindo uma representação espacial das relações entre os elementos que compõem a comunidade científica (Small, 1999). Dessa maneira, possibilita monitorar e delimitar áreas de pesquisa a fim de determinar a evolução e a configuração do campo científico (Noyons; Moed; Van Raan, 1999). No entanto, esse processo não deve ser aplicado diretamente aos dados obtidos das bases bibliográficas. É necessário realizar a tabulação dos dados na etapa de pré-processamento, garantindo sua consistência e confiabilidade para fins de análise bibliométrica (Cobo *et al.*, 2011).

A pesquisa bibliométrica configura-se, portanto, como uma abordagem relevante para resgatar a retrospectiva histórica e evolutiva de um campo científico (Öztürk *et al.*, 2024). Nesse contexto, o mapeamento científico, fundamentado na abordagem quantitativa, utiliza análises estatísticas para

investigar a estrutura e o desenvolvimento do campo científico (Zupic; Cater, 2015).

A metodologia empregada no mapeamento científico baseia-se na técnica de análise de palavras para visualizar o desenvolvimento de um domínio científico. O mapa de palavras representa, de maneira bidimensional, os termos fortemente relacionados (posicionados próximos) e os menos relacionados (colocados mais distantes), oferecendo uma visão geral da estrutura cognitiva do campo. Assim, o primeiro passo metodológico é a definição do campo científico, associada a uma perspectiva multidimensional (Van Raan, 2014).

A etapa de mapeamento científico consiste em aplicar um algoritmo sobre a rede formada pelos relacionamentos entre as unidades de análise selecionadas. Posteriormente, a análise de rede permite a realização de análises estatísticas sobre o mapa gerado. A análise temporal, por sua vez, possibilita identificar padrões, tendências e a evolução do campo científico enquanto a análise geoespacial evidencia os locais de produção científica e seus impactos (Cobo *et al.*, 2011).

O processo de normalização é fundamental para definir a relevância e as medidas de similaridade entre as unidades de informação. As unidades de análise mais utilizadas no mapeamento científico são: periódicos, documentos, referências citadas, autores e termos ou palavras descritivas (Börner *et al.*, 2003). Geralmente, esses elementos são utilizados como dados de co-ocorrência para a construção das redes bibliométricas (Cobo *et al.*, 2011). A relação entre as unidades pode ser representada como um grafo ou rede em que os nós representam as unidades de análise e as arestas indicam os vínculos entre elas.

Por fim, diferentes aspectos de um campo científico podem ser investigados a depender das unidades de análise escolhidas (Cobo *et al.*, 2011). A análise de co-autoria permite examinar a estrutura social do campo, considerando as interações entre os autores (Gänzel, 2001; Peter; Van Raan, 1991). A dimensão internacional pode ser explorada a partir das afiliações institucionais e geográficas dos pesquisadores (Gänzel, 2001; Peter; Van Raan, 1991). A estrutura conceitual do campo pode ser investigada por meio da análise de co-palavras (Callon *et al.*, 1983), enquanto a estrutura intelectual pode ser

analisada por meio da co-citação (Small, 1973) e do acoplamento bibliográfico (Kessler, 1963).

### 3.3 CONCEITOS ESTATÍSTICOS APLICADOS ÀS UNIDADES DE INFORMAÇÃO

A estatística descritiva constitui a base inicial do processo estatístico, reunindo procedimentos destinados à coleta, organização, descrição e apresentação dos dados. No campo das unidades de informação, sua aplicação permite estruturar e interpretar fenômenos documentais e informacionais por meio de conceitos basilares como população, amostra, unidade estatística e variável (Pinto *et al.*, 2012).

A população corresponde ao conjunto de elementos que compartilham ao menos uma característica relevante para o contexto de análise enquanto a amostra representa um subconjunto dessa população, selecionado com a finalidade de representá-la adequadamente. Cada elemento que compõe a população é denominado unidade estatística e a variável corresponde ao atributo mensurável que caracteriza o fenômeno estudado e cujos valores podem variar entre os elementos observados (Pinto *et al.*, 2012).

No processo de tratamento dos dados, as séries estatísticas desempenham o papel de organizar e evidenciar a distribuição dos valores referentes ao fenômeno analisado. Segundo Pinto *et al.* (2012), essas séries podem ser classificadas em três categorias, de acordo com o elemento variável: as séries históricas, cujo eixo é o tempo; as séries geográficas, que enfatizam variações espaciais; e as séries específicas, focadas nas características próprias do fenômeno. Além disso, a distribuição de frequência destaca-se como uma técnica central, pois agrupa os dados em classes ou intervalos, mantendo constantes os elementos fato, local e época, de modo a revelar a concentração dos valores dentro de cada classe (Pinto *et al.*, 2012).

A representação gráfica e tabular dos dados constitui uma etapa indispensável à análise estatística uma vez que facilita a interpretação dos resultados. As tabelas devem apresentar informações de forma clara, identificando o conteúdo e as relações estabelecidas enquanto os gráficos

devem ser autoexplicativos e facilitar a compreensão visual do fenômeno (Pinto *et al.*, 2012). Entre as ferramentas visuais utilizadas, o diagrama de dispersão se destaca por ilustrar a relação entre duas variáveis em um plano cartesiano, permitindo analisar associações, comparar tratamentos aplicados a um mesmo indivíduo ou avaliar diferenças entre medições realizadas em momentos distintos (Pinto *et al.*, 2012).

No âmbito da distribuição de frequência, diversas medidas estatísticas contribuem para caracterizar e interpretar o comportamento dos dados. As medidas de posição, como as medidas de tendência central e as separatrizes, indicam a localização relativa da distribuição no eixo horizontal. A média representa o ponto de equilíbrio da distribuição, a mediana identifica o valor central e a moda corresponde ao valor mais frequente. As separatrizes, por sua vez, dividem a distribuição em partes iguais, englobando mediana, decis, quartis e percentis (Pinto *et al.*, 2012).

As medidas de dispersão complementam essa análise ao indicar o grau de afastamento dos valores em relação ao ponto de referência. O ranking corresponde à diferença entre o maior e o menor valor observado, o desvio-padrão quantifica o espalhamento dos dados em torno da média, sendo maior quanto mais afastados estiverem os valores extremos, a variância expressa, em média, a distância entre os valores observados e o valor esperado e o coeficiente de variação permite comparar distribuições que diferem em escala ou magnitude, oferecendo uma medida relativa da variação (Pinto *et al.*, 2012).

Além dessas, outras métricas contribuem para a caracterização completa da distribuição. As medidas de assimetria permitem analisar a relação entre média, mediana e moda, evidenciando a direção e intensidade do desbalanceamento da distribuição. Já as medidas de curtose avaliam o grau de achatamento ou "alongamento" da curva de frequência em comparação a uma distribuição normal de referência, revelando o comportamento dos valores extremos (Pinto *et al.*, 2012).

Por fim, os números-índices constituem instrumentos que sintetizam e comparam grupos de variáveis relacionadas ou variações de uma mesma variável ao longo do tempo. Seu uso permite captar mudanças significativas em áreas previamente definidas e, no caso dos índices ponderados, a análise

incorpora o peso proporcional que cada componente exerce no valor total das transações estudadas, tornando a interpretação mais robusta e representativa (Pinto *et al.*, 2012).

### 3.4 INDICADORES BIBLIOMÉTRICOS DE ATIVIDADES CIENTÍFICAS

Os indicadores bibliométricos desempenham um papel estratégico na mensuração da atividade científica pois permitem quantificar dados, descrever o panorama de um campo e analisar sua dinâmica de desenvolvimento. Como destaca Pinto *et al.* (2012), os estudos métricos da informação ampliam o potencial de geração destes indicadores, ao fornecer instrumentos capazes de avaliar a eficácia, o impacto e a efetividade social das atividades científicas. Dessa forma, os indicadores constituem ferramentas necessárias para a análise de políticas científicas, para o monitoramento da produtividade e para a identificação de padrões de visibilidade e circulação do conhecimento.

Nesse contexto, embora sejam frequentemente associados apenas ao levantamento estatístico, os estudos bibliométricos vão muito além dessa dimensão quantitativa, pois possibilitam análises complexas e multifacetadas sobre os comportamentos informacionais, científicos e sociais. Ferreira (2010) reforça que se trata de um método robusto que integra abordagens quantitativas à interpretação das dinâmicas cognitivas, sociais e organizacionais do fazer científico, permitindo compreender não apenas o volume, mas também a qualidade e a influência da produção científica. Assim, os indicadores bibliométricos e científicos se consolidam como instrumentos estratégicos para subsidiar processos de avaliação, planejamento e formulação de políticas de fomento.

Pinto e Matias (2011) enfatizam que tais indicadores devem ser flexíveis e sensíveis às especificidades de cada área do conhecimento, visto que sua função não se limita à mensuração, mas também ao reconhecimento das necessidades e características do campo científico. Por essa razão, tornam-se necessários para a compreensão de cenários e para a orientação de políticas voltadas ao desenvolvimento científico.

A confiabilidade dos indicadores depende diretamente de padrões adequados de organização e representação da informação. Pinto e Matias (2011) alertam que a ausência de padronização ou a presença de erros nos metadados podem comprometer a qualidade dos indicadores, gerando resultados imprecisos e pouco úteis para decisões estratégicas. Nesse sentido, o aperfeiçoamento contínuo dos indicadores é necessário, pois eles devem acompanhar a evolução das práticas científicas e das transformações metodológicas da própria ciência. Esse processo depende de registros informacionais consistentes, de metadados padronizados e de modelos de organização que garantam precisão, recuperação eficiente e contextualização adequada dos dados.

Nessa perspectiva, os indicadores bibliométricos se afirmam como produtos de um ecossistema informacional complexo, cuja qualidade afeta diretamente a relevância das análises realizadas. Segundo García-Villar e García Santos (2021), os indicadores bibliométricos são dados numéricos que descrevem diversas dimensões da atividade científica relacionada à produção, ao uso e ao impacto da informação. Eles podem ser classificados em diferentes categorias, entre as quais se destacam: indicadores de atividade científica; indicadores de produção; indicadores de colaboração; indicadores de disseminação; indicadores de produtividade e indicadores de impacto. Cada categoria permite observar aspectos específicos da dinâmica científica, desde a produtividade individual e institucional até padrões de coautoria, circulação do conhecimento e impacto social das pesquisas científicas.

Além disso, a bibliometria, consolidada como campo teórico-metodológico, oferece um conjunto de leis estatísticas que favorecem a produção e a interpretação dos indicadores. A Lei de Bradford permite estimar a relevância e a centralidade dos periódicos; a Lei de Lotka mede a produtividade e o prestígio de autores e a Lei de Zipf analisa a frequência de ocorrência de termos na comunicação científica. O Ponto T de Goffman contribui, por sua vez, para identificar a transição entre termos de baixa e alta frequência, revelando padrões discursivos característicos de um domínio. Complementarmente, a análise de citações se destaca como um dos indicadores mais reconhecidos na

mensuração da influência científica, ao possibilitar o mapeamento de redes de colaboração, impacto intelectual e circulação do conhecimento (Guedes, 2012).

Como disciplina fundamentada em princípios estatísticos e empíricos, a bibliometria oferece uma base teórica sólida para a Ciência da Informação, ao mesmo tempo em que se constitui como ferramenta estratégica para a gestão da ciência e tecnologia. Suas abordagens teóricas, práticas e descritivas contribuem para a mensuração da informação, para a geração de indicadores relevantes e para o aprimoramento dos processos de avaliação, planejamento e desenvolvimento científico. Dessa forma, os indicadores bibliométricos se afirmam não apenas como instrumentos analíticos, mas como elementos estruturantes de um sistema científico orientado pela transparência, eficiência e busca pela excelência (Guedes, 2012).

O Quadro 8 sistematiza as principais leis e princípios bibliométricos, evidenciando seus focos de estudo e respectivas aplicações no âmbito da análise da produção científica e da comunicação do conhecimento.

Quadro 8 - Leis e princípios bibliométricos

LEIS E PRINCÍPIOS	FOCO DE ESTUDO	APLICAÇÕES
Lei de Bradford	Título de Periódicos	Estimativa dos graus relativos de relevância de títulos de periódicos em áreas específicas do conhecimento
Lei de Lotka	Autores	Estimativa dos graus relativos de relevância de autores em áreas específicas do conhecimento
Lei de Zipf	Palavras	Análise conceitual da escrita científica e indexação automática ou semiautomática de artigos científicos
Ponto de Transição de Goffman		Análise conceitual da escrita científica e indexação automática ou semiautomática de artigos científicos e tecnológicos
Colégios Invisíveis	Citações	Identificação da elite de pesquisadores em áreas específicas do conhecimento
Fator de Imediatismo ou Fator de Impacto		Estimativa dos graus de relevância de artigos, cientistas e títulos de

		periódicos científicos em áreas específicas do conhecimento
Acoplamento Bibliográfico		Estimativa dos graus relativos de ligação de dois ou mais artigos através de uma análise retrospectiva
Co-Citação		Estimativa dos graus relativos de ligação de dois ou mais artigos através de uma análise prospectiva
Obsolescência da Literatura		Estimativa do grau de declínio da literatura em áreas específicas do conhecimento
Vida Média		Estimativa da vida média de uma unidade da literatura em áreas específicas do conhecimento
Teoria Epidêmica de Goffman		Estimativa do grau de crescimento e de declínio de uma área de assunto e a importância de linhas de pesquisa em áreas específicas do conhecimento
Lei do Elitismo		Identificação e descrição da elite formada por autores que participaram intensivamente da produção científica em áreas específicas do conhecimento
Frente de Pesquisa		Identificação de um padrão de relações múltiplas entre autores que se citam na literatura e o reconhecimento dos colégios invisíveis em áreas específicas do conhecimento
Lei dos 80/20	Demanda de Informação	Processos de tomada de decisão ligados à composição, ampliação e redução de acervos em sistemas de recuperação da informação

Fonte: adaptado de Guedes (2012).

As abordagens apresentadas contemplam diferentes unidades de análise — periódicos, autores, palavras, citações e áreas temáticas — e oferecem subsídios teóricos e metodológicos para a mensuração da relevância, da produtividade, do impacto e da dinâmica da literatura científica. Leis clássicas, como as de Bradford, Lotka e Zipf, são articuladas a princípios voltados à

estrutura social da ciência, como colégios invisíveis, co-citação, acoplamento bibliográfico e frentes de pesquisa, além de indicadores temporais, como obsolescência e vida média da literatura. Em conjunto, o quadro evidencia o papel da bibliometria como instrumento analítico para compreender padrões de produção, disseminação e uso da informação científica, bem como para apoiar processos de avaliação, gestão e tomada de decisão em ciência e tecnologia.

## 4 BOAS PRÁTICAS CIENTÍFICAS EM CIÊNCIA ABERTA E INCLUSIVAS

A presente seção dedica-se a discutir a comunicação científica em sua correlação com a internacionalização da ciência e o movimento de acesso aberto em publicações científicas, destacando seu papel na democratização do acesso à informação e promoção da equidade na disseminação do conhecimento científico.

A comunicação científica, entendida como o alicerce que sustenta a circulação dos resultados de pesquisa, torna-se ainda mais relevante em um cenário global marcado pela necessidade de transparência, colaboração e inclusão. Nesse contexto, as práticas de ciência aberta e inclusiva, emergem como uma política científica orientada para a gestão da informação, ampliando a visibilidade e o impacto da produção científica ao assegurar que dados, métodos e publicações estejam acessíveis a diferentes comunidades científicas, instituições e países.

Assim, as boas práticas científicas não apenas favorecem a redução das desigualdades no acesso ao conhecimento científico, mas, também fortalecem a integridade da ciência, estimulando a cooperação internacional e promovendo um ecossistema informacional mais justo, participativo e sustentável.

### 4.1 A COMUNICAÇÃO CIENTÍFICA E A DEMOCRATIZAÇÃO DO ACESSO À INFORMAÇÃO

Borko (1968) identifica a comunicação científica como um fenômeno da Ciência da Informação ao afirmar que essa área do conhecimento emerge justamente para investigar os processos envolvidos na produção, transmissão e uso da informação científica. Desse modo, compreender a definição conceitual de “comunicação científica” torna-se fundamental para analisar seu papel na democratização do acesso à informação e promoção da equidade na disseminação do conhecimento científico.

Nessa perspectiva, Garvey (1979) conceitua a comunicação científica como um conjunto de atividades que abrangem a produção, disseminação e uso da informação, destacando sua natureza processual e dinâmica. Já Meadows

(1999) define a comunicação científica como o processo de disseminação da informação especializada entre membros da comunidade científica, enfatizando que sua efetividade depende da escolha adequada dos canais de comunicação, da natureza das informações transmitidas e do público-alvo. Complementarmente, Price (1976) aponta o artigo científico como o principal canal formal de comunicação destinado à divulgação e validação do saber científico, por meio do registro, da avaliação e da circulação do conhecimento produzido.

Avançando para uma perspectiva contemporânea, Maricato e Mendes (2015) destacam que a comunicação científica assume também uma dimensão social ao ampliar a participação da sociedade no processo de construção da ciência. Sob essa ótica, a comunicação científica deixa de ser apenas um mecanismo de circulação entre pares e passa a desempenhar um papel na democratização do acesso à informação, promovendo a visibilidade do campo científico, o compartilhamento de resultados e a difusão do conhecimento.

Maricato e Mendes (2015) citam que ao fortalecer o acesso, a transparência e a circulação da informação científica, a comunicação científica contribui de forma significativa para o avanço do conhecimento, para a redução de desigualdades informacionais e para a promoção de uma ciência socialmente relevante.

Os autores apresentam a definição conceitual de termos e conceitos relacionados à comunicação científica, destacando três elementos: a difusão científica, entendida como o processo e os recursos utilizados para comunicar uma informação científica e/ou tecnológica; a divulgação científica, voltada ao público-geral e não especialista; e a disseminação científica, direcionada aos cientistas e especialistas de uma comunidade científica (Maricato; Mendes, 2015).

A distinção desses conceitos é relevante, sobretudo no contexto dos países em desenvolvimento, que enfrentam desafios para garantir a visibilidade da produção científica e assegurar o acesso equitativo ao conhecimento científico global. Nesse cenário, torna-se pertinente refletir sobre o papel da comunicação científica na democratização do acesso à informação como

instrumento de promoção da equidade na disseminação do conhecimento científico (Maricato; Mendes, 2015).

O Quadro 9 apresenta uma distinção conceitual entre os termos difusão científica, divulgação científica e disseminação científica, evidenciando suas especificidades no âmbito da comunicação científica.

Quadro 9 - Termos e conceitos relacionados à comunicação científica

TERMO	DIFUSÃO CIENTÍFICA	DIVULGAÇÃO CIENTÍFICA	DISSEMINAÇÃO CIENTÍFICA
CONCEITO	Refere-se ao processo e recurso utilizado para comunicar uma informação científica e/ou tecnológica	Refere-se ao processo de comunicação científica dirigido ao público-geral, ou seja, não especialista	Refere-se ao processo de comunicação científica direcionado aos cientistas e especialistas de uma comunidade científica

Fonte: adaptado de Maricato e Mendes (2015).

A difusão científica é compreendida como um processo amplo de comunicação do conhecimento científico e tecnológico, envolvendo diferentes meios e públicos. A divulgação científica, por sua vez, refere-se à comunicação da ciência dirigida ao público leigo, caracterizando-se pela tradução e adaptação da linguagem especializada para torná-la acessível a não especialistas. Já a disseminação científica está associada à comunicação formal entre cientistas e especialistas, no interior de uma comunidade científica, com foco na circulação do conhecimento validado. Ao explicitar essas distinções, o quadro contribui para a compreensão dos diferentes fluxos e estratégias de comunicação do conhecimento científico, os quais são relevantes para análises sobre produção, circulação e apropriação da informação científica.

A literatura indica que a troca de informações, a validação de descobertas científicas e a construção do conhecimento científico configuram os fundamentos da ciência da informação como uma ciência moderna (Saracevic, 1996; Araújo, 2003; Gäal; Martins, 2022). Esses processos sustentam a comunicação científica como uma prática científica que contribui com o avanço do campo científico uma vez que possibilita a circulação, o compartilhamento e a consolidação do conhecimento. Assim, a comunicação científica não fortalece

apenas a dinâmica entre produção e uso da informação, mas, viabiliza a participação social no campo científico, ampliando sua legitimidade e função social.

Com essa perspectiva, observa-se que o desenvolvimento da cultura científica e o estímulo à prática de ciência, tecnologia e inovação voltados à promoção da inclusão social e à redução de desigualdades motivaram a criação de um órgão consultivo junto ao Ministério da Ciência, Tecnologia e Inovação (MCTI). Esse órgão tem como objetivo apoiar as ações do Programa Pop Ciência (Brasil, 2023). O Decreto nº 11.754/2023, ao instituir o Programa Nacional de Popularização da Ciência (Pop Ciência e Comitê Pop) formaliza o compromisso com a democratização do conhecimento científico e a promoção da cultura científica, reforçando o papel da comunicação científica como mecanismo de inclusão (Brasil, 2023).

Os estudos reforçam que a promoção, o incentivo e o apoio às atividades, espaços e iniciativas científicas estão alinhados ao escopo da ciência aberta e inclusiva. Entende-se que a popularização da ciência busca democratizar o acesso ao conhecimento, tornando-o compreensível e acessível à sociedade. Nesse sentido, a comunicação científica aproxima ciência e público, desmistifica conceitos complexos e promove o engajamento social com os avanços científicos. Desse modo, a ciência aberta configura-se como um termo abrangente que envolve práticas, iniciativas, perspectivas e implicações vinculadas ao acesso aberto e à ciência cidadã (Shintaku et al., 2024).

Nesse contexto, Falangola, Santos e Rodrigues (2025) investigam o papel da comunicação científica na construção e disseminação do conhecimento com o objetivo de compreender como a divulgação científica influencia a produção do conhecimento ao considerar a difusão global das publicações científicas. Os autores argumentam que a adoção de práticas científicas que contemplem a qualidade dos dados de pesquisa, a contribuição para o desenvolvimento científico e o impacto na sociedade favorece uma avaliação mais equitativa e representativa da diversidade do conhecimento científico.

O estudo demonstra que tais práticas promovem uma comunicação científica mais inclusiva, colaborativa e abrangente. Contudo, apesar dos avanços tecnológicos proporcionarem novas formas e canais de disseminação

do conhecimento, ainda persistem barreiras relacionadas à acessibilidade, ao financiamento e à avaliação da qualidade das publicações científicas. Portanto, a comunicação científica deve buscar o equilíbrio entre inovação e credibilidade, a fim de atender ao papel social da ciência e fortalecer sua relevância no desenvolvimento humano e social (Falangola; Santos; Rodrigues, 2025).

#### 4.2 O MOVIMENTO DE ACESSO ABERTO EM PUBLICAÇÕES CIENTÍFICAS

A Declaração de Budapeste (2002) conceitua o acesso aberto como um movimento que visa remover barreiras econômicas e legais ao conhecimento científico, tendo como mérito a defesa de duas estratégias:

1. o autoarquivamento da produção científica;
2. o desenvolvimento de uma nova geração de revistas científicas em acesso aberto.

Os estudos indicam que esse movimento se consolidou com a *Budapest Open Access Initiative*, em 2002, como resposta ao fenômeno conhecido como “crise dos periódicos”, ocorrido em meados da década de 1980, quando houve uma redução gradual do número de assinaturas de periódicos científicos, comprometendo a disseminação do conhecimento.

Nesse contexto, Shintaku, Gottschalg Duque e Suaiden (2016) destacam que o acesso à informação torna o processo de publicação científica mais democrático, ampliando o impacto da comunicação científica em escala global. Segundo os autores, o movimento de acesso aberto exerce forte influência sobre a comunicação científica, pois é amplamente aceito e utilizado como estratégia para disseminar a produção científica, aumentando sua visibilidade e impacto. Eles apontam que as práticas científicas alinhadas aos princípios da ciência aberta e inclusiva contribuem para a democratização da ciência uma vez que expandem o acesso à informação científica e promovem a participação equitativa de diferentes atores no ecossistema científico.

Nesse mesmo escopo, Costa e Leite (2017) propõem princípios e recomendações para orientar a comunicação dos dados de pesquisa em nível internacional, destacando a necessidade de formalizar a responsabilidade sobre a disponibilização pública dos dados, respeitar a conformidade legal, garantir a interoperabilidade, assegurar a qualidade e preservação dos dados, além de

estabelecer padrões, licenças abertas e transparência científica. Tais diretrizes fortalecem os processos de produção, compartilhamento e reutilização de dados de pesquisa, considerados necessários para a aceleração das descobertas científicas e para o avanço da ciência em ambientes colaborativos.

O Quadro 10 sistematiza os princípios e recomendações para a comunicação de dados científicos, destacando diretrizes normativas, éticas, técnicas e jurídicas que orientam o compartilhamento, o uso e a preservação dos dados de pesquisa.

Quadro 10 - Princípios e recomendações para a comunicação de dados científicos

RECOMENDAÇÕES	PRINCÍPIOS
Responsabilidade formal	É a formalização do compromisso em tornar o dado de pesquisa público
Conformidade legal	Destaca que a obrigatoriedade do compartilhamento e o reuso dos dados de pesquisa devem ser flexíveis aos sistemas jurídicos do país ou região
Interoperabilidade e qualidade dos dados	Evidencia a necessidade de padrões internacionais para a descrição, o armazenamento e a recuperação dos dados
Equilíbrio de interesses	Considera exceções na obrigatoriedade do compartilhamento dos dados de pesquisa em função dos interesses de pesquisadores e financiadores
Proteção da propriedade intelectual	Considera a possibilidade de restrição ao compartilhamento e o reuso dos dados de pesquisa em função dos direitos de propriedade intelectual e segredos comerciais
Custo mínimo para reprodução dos dados	Determina que o acesso ou uso do dado de pesquisa gerem o mínimo de custo possível para o usuário
Transparência para a ciência	Destaca a importância do acesso aos dados de pesquisa para tornar os resultados apresentados incorruptíveis e refutáveis
Prestação de contas	É a disponibilização dos dados de pesquisa como forma de comprovar a realização da pesquisa e o cumprimento de seus objetivos
Sem restrição de uso	Recomenda-se que não seja atribuída nenhuma restrição de uso, nem mesmo de uso comercial
Conjunto de dados	É a requisição do conjunto de dados produzidos e/ou utilizados pela pesquisa para a comprovação dos resultados
Preservação	É a necessidade de evidenciar formas e financiamentos para a preservação dos dados em longo prazo
Profissionalismo	É o uso de valores e técnicas profissionais para o tratamento, armazenamento e a disponibilização dos dados de pesquisa
Compartilhamento da licença	É a implicação da necessidade de obras derivadas ou que utilizem os dados de pesquisa compartilharem da mesma

RECOMENDAÇÕES	PRINCÍPIOS
	licença a fim de garantir a abertura dos dados
Dados em formato modificável	Indica que sempre que possível os dados de pesquisa devem ser compartilhados em formato modificável

Fonte: adaptado de Costa e Leite (2017).

O quadro evidencia aspectos como responsabilidade formal, conformidade legal, interoperabilidade, qualidade dos dados, transparência, preservação e prestação de contas, além de considerar questões relacionadas à propriedade intelectual, custos de reprodução e licenciamento. Ao reunir esses princípios, o quadro oferece uma visão integrada das boas práticas necessárias para garantir o acesso aberto, a reutilização responsável e a sustentabilidade dos dados científicos, reforçando o papel dos dados como insumos importantes para a validação, a reprodutibilidade e o avanço da ciência.

Em continuidade, Costa e Leite (2018) investigam os fatores de influência na comunicação de dados de pesquisa, evidenciando a complexidade da prática de compartilhamento científico. Entre esses fatores estão os imperativos da ciência moderna, tais como a transparência, o reuso dos dados, a visibilidade científica, os aspectos éticos e legais, a infraestrutura tecnológica, os sistemas de informação, os padrões de processos e os contextos sociais e institucionais. Esses elementos moldam a comunicação científica e evidenciam o compartilhamento de dados como componente necessário para a reprodutibilidade, a confiabilidade e a validação da produção científica.

O Quadro 11 apresenta os principais fatores de influência na comunicação dos dados de pesquisa, organizados em quatro eixos analíticos: imperativos da ciência moderna, infraestrutura, contexto social e financiamento. O Quadro 11 apresenta os fatores de influência na comunicação dos dados de pesquisa em Ciência da Informação.

Quadro 11 - Fatores de influência na comunicação dos dados de pesquisa

IMPERATIVOS DA CIÊNCIA MODERNA	<ul style="list-style-type: none"> <li>● Aceleração da ciência</li> <li>● Reuso dos dados de pesquisa</li> <li>● Transparência da ciência</li> <li>● Assédio aos pesquisadores</li> <li>● Visibilidade da pesquisa</li> <li>● Recompensa aos pesquisadores</li> <li>● Cultura acadêmica</li> <li>● Aspectos legais</li> <li>● Aspectos éticos</li> </ul>
INFRAESTRUTURA	<ul style="list-style-type: none"> <li>● Sistemas de informação</li> <li>● Padrões de processos, procedimentos e sistemas</li> <li>● Políticas como instrumento de comunicação</li> </ul>
CONTEXTO SOCIAL	<ul style="list-style-type: none"> <li>● Cenário de aplicação</li> </ul>
FINANCIAMENTO	<ul style="list-style-type: none"> <li>● Instituições financiadoras</li> <li>● Pesquisa científica</li> <li>● Dados científicos</li> <li>● Bem público</li> <li>● Custo associado</li> </ul>

Fonte: adaptado de Costa e Leite (2018).

O quadro evidencia que a comunicação dos dados científicos é condicionada tanto por demandas estruturais da ciência contemporânea, como o reuso de dados, a transparência, a visibilidade da pesquisa e os aspectos legais e éticos, quanto por elementos institucionais e tecnológicos, a exemplo dos sistemas de informação, padrões e políticas de comunicação. Além disso, destaca a influência do contexto social e das fontes de financiamento, que moldam o cenário de aplicação, os custos envolvidos e a compreensão dos dados como bem público. Em conjunto, esses fatores revelam a complexidade multidimensional que envolve a comunicação dos dados de pesquisa, ressaltando a necessidade de abordagens integradas que considerem dimensões científicas, técnicas, sociais e econômicas.

Com base nessa perspectiva, Cedro e Gottschalg Duque (2020) reforçam que a disseminação da informação científica é um critério fundamental para a transparência e validação das práticas científicas, destacando que a divulgação científica atua na democratização do conhecimento ao tornar a informação acessível por meio de linguagem compreensível. Os autores argumentam que garantir a confiabilidade e a qualidade dos dados de pesquisa é indispensável para legitimá-los como fonte de evidência científica, o que fortalece o papel do compartilhamento de dados como prática científica.

A literatura mostra que o acesso aberto busca remover barreiras de acesso às publicações científicas e promover a colaboração, contribuindo diretamente para o progresso científico. Ao eliminar barreiras financeiras, permite a participação de autores, instituições e países com recurso reduzido, ampliando a representatividade e a diversidade na comunicação científica (Gaal; Martins, 2022). Pereira (2022) acrescenta a necessidade de alfabetização científica como uma condição para que a divulgação científica cumpra seu papel social e democrático uma vez que apenas a exposição dos resultados científicos não garante sua compreensão ou apropriação pela sociedade.

Sob a perspectiva das políticas públicas, Santos (2022) apresenta um panorama das iniciativas governamentais brasileiras relativas ao direito de acesso à informação, ressaltando a importância do acesso aberto para a transformação das práticas científicas. A autora destaca que a ciência aberta promove transparência, produtividade e reprodutibilidade da pesquisa científica, criando novas oportunidades de colaboração, reconhecimento e inovação. No entanto, aponta que a transição para a ciência aberta é um processo multidimensional que exige a transformação da cultura científica, a implementação de políticas de dados abertos, o aprimoramento da infraestrutura tecnológica e a promoção da interoperabilidade entre sistemas de informação.

Nesse sentido, Santos (2022) enfatiza que a ciência aberta tem potencial para aumentar a qualidade, o alcance e os benefícios da ciência, tornando-a mais confiável e responsiva aos desafios da sociedade, além de favorecer a otimização da produção científica por meio do compartilhamento e da reutilização dos dados de pesquisa. O compartilhamento de dados, nesse contexto, emerge como uma estratégia metodológica que potencializa o alcance da comunicação científica, estimulando a participação colaborativa e o protagonismo dos pesquisadores.

Todavia, há desafios e contradições no escopo do acesso aberto, como observam Araújo, Barcellos e Sacramento (2024), ao discutirem a transferência dos custos de publicação para os autores por meio das taxas de processamento de artigos (APCs) e os impactos desse modelo na equidade científica. Tais condições limitam a participação de pesquisadores e instituições com orçamento

financeiro limitado, além de suscitar debates sobre a qualidade editorial, a sustentabilidade financeira e as práticas predatórias no campo científico.

Por fim, Marques (2025) reforça que o movimento de acesso aberto em publicações científicas representa um elemento fundamental para a democratização do conhecimento científico uma vez que promove a equidade no acesso à informação, estimula a inovação e gera impacto social. Ao acelerar o fluxo da divulgação científica, o acesso aberto fortalece a interconexão entre ciência, sociedade e tecnologia, consolidando-se como ferramenta relevante para a construção de uma ciência mais participativa, transparente e colaborativa.

### 4.3 A INTERNACIONALIZAÇÃO DA CIÊNCIA

O Brasil destaca-se por seu pioneirismo no movimento de acesso aberto embora ainda enfrente limitações significativas para manter e expandir a infraestrutura necessária ao suporte da ciência aberta. A literatura indica que a consolidação dessa cultura é um processo de longo prazo, mas, que deve ser impulsionado por ações gradativas, de baixo custo e orientadas à gestão e ao compartilhamento responsável de dados científicos. Nesse sentido, recomenda-se que as agências de fomento colaborem de forma sistemática com a implementação do Plano de Ação Nacional de Ciência Aberta, contribuindo para estruturar e fortalecer a política nacional voltada à abertura da ciência (IBICT, 2025).

A adoção das práticas de ciência aberta tem potencial para ampliar a inovação e elevar o impacto da pesquisa científica brasileira. Para tanto, torna-se necessário articular ações concretas que envolvam o Ministério da Ciência, Tecnologia e Inovação (MCTI), as agências de fomento, os pesquisadores, as instituições científicas e tecnológicas e as sociedades científicas como um todo. Trata-se então de promover mudanças graduais, porém consistentes, rumo a um modelo de ciência mais inclusiva, acessível e alinhada às tendências e esforços internacionais. O objetivo é assegurar que o conhecimento produzido com recursos públicos seja amplamente aproveitado e colocado a serviço da sociedade (IBICT, 2025).

Heinz e Miranda (2024) reforçam que a ciência aberta potencializa o desenvolvimento científico e social embora enfatizem desafios estruturais, sobretudo a ausência de padronização normativa e de políticas consolidadas que orientem a sua adoção. Para os autores, a ciência aberta deve ser compreendida como um ecossistema capaz de promover avanços sociais ao articular evidências científicas, ampliar o alcance das publicações e estabelecer regras mais justas e equitativas para a comunicação científica.

Nessa mesma direção, Araújo, Barcellos e Sacramento (2024) discutem a necessidade de consolidar uma política brasileira de ciência aberta que esteja em sintonia com o sistema nacional de avaliação científica. Os autores defendem que a incorporação de boas práticas de avaliação, sensíveis à diversidade do fazer científico e às dinâmicas contemporâneas de circulação do conhecimento, constitui um passo importante para a construção de uma ciência inclusiva e global.

O Quadro 12 sistematiza as recomendações da política brasileira de ciência aberta associadas à adoção de boas práticas de avaliação científica, evidenciando diretrizes voltadas à promoção da abertura, da equidade e da qualidade no sistema científico nacional.

Quadro 12 - Recomendações da política brasileira de ciência aberta e adoção de boas práticas de avaliação científica

RECOMENDAÇÕES DA POLÍTICA BRASILEIRA DE CIÊNCIA ABERTA E ADOÇÃO DE BOAS PRÁTICAS DE AVALIAÇÃO CIENTÍFICA
<ul style="list-style-type: none"> <li>• Investir em infra estruturas abertas e orientadas para a colaboração e abertura do conhecimento científico em benefício público</li> </ul>
<ul style="list-style-type: none"> <li>• Fomentar a sustentabilidade e a qualificação dos periódicos nacionais de acesso aberto diamante mediante editais de financiamento</li> </ul>
<ul style="list-style-type: none"> <li>• Valorizar o trabalho das equipes editoriais dos periódicos de acesso aberto diamante</li> </ul>
<ul style="list-style-type: none"> <li>• Valorizar as publicações em periódicos de acesso aberto diamante como critério de avaliação científica e apoiar os modelos editoriais não comerciais e sem taxas de processamento de artigos para contribuir com a equidade na ciência e o enfraquecimento dos periódicos e das práticas predatórias</li> </ul>
<ul style="list-style-type: none"> <li>• Incentivar a utilização de bases bibliográficas de dados não comerciais de acesso aberto e de maior cobertura global, regional e local para a avaliação da ciência brasileira</li> </ul>
<ul style="list-style-type: none"> <li>• Reformular as políticas de avaliação científica com vista à adequação aos princípios de acesso aberto</li> </ul>

<ul style="list-style-type: none"> <li>• Incentivar uma internacionalização globalizada e menos colonial que valorize as interações e as trocas entre os países do sul global</li> </ul>
<ul style="list-style-type: none"> <li>• Incentivar o multilinguismo na prática científica diminuindo o imperativo do inglês como a língua franca da ciência com vista a garantir a equidade</li> </ul>
<ul style="list-style-type: none"> <li>• Incentivar a formação, a qualificação e a valorização dos pareceristas de mérito científico a partir do treinamento dos pesquisadores com o objetivo de contribuir para a melhoria do processo de avaliação por pares</li> </ul>
<ul style="list-style-type: none"> <li>• Incentivar outras abordagens de avaliação da ciência além das centradas no periódico científico</li> </ul>
<ul style="list-style-type: none"> <li>• Promover a conscientização e o debate com pesquisadores de outras áreas não especialistas do tema sobre a importância da abertura do conhecimento científico e das recomendações de boas práticas de avaliação científica</li> </ul>
<ul style="list-style-type: none"> <li>• Promover uma mudança cultural de ciência aberta a partir da inclusão do tema como parte da formação de pesquisadores em início de carreira</li> </ul>
<ul style="list-style-type: none"> <li>• Incentivar o desenvolvimento de pesquisas para propor novas formas de avaliação da ciência do ponto de vista quantitativo e qualitativo</li> </ul>

Fonte: adaptado de Araújo, Barcellos e Sacramento (2024).

As recomendações abrangem ações estruturantes, como o investimento em infraestruturas abertas e colaborativas, o fortalecimento dos periódicos de acesso aberto diamante e a valorização do trabalho editorial, bem como propostas de reformulação dos modelos tradicionais de avaliação científica, historicamente centrados no periódico. O quadro também enfatiza a importância da internacionalização menos colonial, do multilinguismo, da formação de avaliadores e da diversificação dos critérios e métodos de avaliação, incluindo abordagens quantitativas e qualitativas. Em conjunto, essas recomendações apontam para uma mudança cultural na ciência brasileira, alinhada aos princípios da ciência aberta, à democratização do conhecimento e à construção de um sistema de avaliação mais justo, transparente e socialmente responsável.

As recomendações apresentadas visam sensibilizar a comunidade científica e os usuários da informação acerca da importância de uma política nacional de acesso aberto e da consolidação de uma cultura voltada à democratização do conhecimento no país. O estudo enfatiza que, para avançar nesse propósito, torna-se imprescindível enfrentar as barreiras históricas que restringem a circulação e a apropriação do conhecimento científico (Araújo; Barcellos; Sacramento, 2024).

Nesse cenário, os autores propõem um manifesto que incorpora boas práticas de política científica alinhadas ao acesso aberto e sustentadas por critérios contemporâneos de avaliação da produção científica. O documento assume papel significativo para a democratização do conhecimento, pois incentiva uma mudança cultural em direção à adoção das práticas de ciência aberta, ao mesmo tempo em que reforça princípios de equidade, transparência e responsabilidade na comunicação científica (Araújo; Barcellos; Sacramento, 2024).

A literatura aponta que a democratização e a universalização do acesso ao conhecimento constituem condições relevantes para o desenvolvimento igualitário e sustentável do campo científico. Para que a ciência seja verdadeiramente global, ela precisa ser inclusiva, considerando e valorizando toda a diversidade de atores, contextos e sistemas de produção do conhecimento e não apenas aqueles oriundos de nações com maior poder econômico que historicamente concentram a visibilidade e a legitimidade no cenário científico internacional. Nessa perspectiva, pensar uma ciência global pressupõe o enfrentamento das desigualdades estruturais e a eliminação dos obstáculos que limitam o acesso ao conhecimento (Araújo; Barcellos; Sacramento, 2024).

No entanto, os autores alertam que a inclusão exige atenção rigorosa aos mecanismos de avaliação científica frequentemente baseados em indicadores que privilegiam a ciência produzida no Norte Global e perpetuam a sub-representação de revistas não comerciais ou oriundas de países periféricos. Esses instrumentos, ao reforçarem a dependência de sistemas hegemônicos de circulação do conhecimento acabam sustentando práticas coloniais de dominação e desigualdade científica que dificultam o reconhecimento pleno da diversidade epistemológica e institucional existente no Sul Global (Araújo; Barcellos; Sacramento, 2024).

Nesse contexto, é importante reconhecer que a internacionalização da ciência brasileira já integra as políticas da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) desde a década de 1990, constituindo-se em um eixo estratégico para o fortalecimento dos Programas de Pós-Graduação e para a ampliação da inserção do país no cenário científico internacional. A

articulação entre internacionalização, acesso aberto e práticas científicas de ciência aberta torna-se, portanto, um caminho necessário para promover maior visibilidade, justiça cognitiva e equidade no ecossistema global da comunicação científica.

O Quadro 13 apresenta os objetivos do Programa Institucional de Internacionalização (PrInt), evidenciando as diretrizes estratégicas estabelecidas para fortalecer a inserção internacional das instituições brasileiras de ensino superior e pesquisa.

Quadro 13 - Objetivos do Programa Institucional de Internacionalização (PrInt)

OBJETIVOS PrInt
<ul style="list-style-type: none"> <li>• Fomentar a construção, a implementação e a consolidação de planos estratégicos de internacionalização das instituições contempladas nas áreas priorizadas do conhecimento</li> </ul>
<ul style="list-style-type: none"> <li>• Estimular a formação de redes de pesquisas internacionais com vistas a aprimorar a qualidade da produção acadêmica vinculadas à pós-graduação</li> </ul>
<ul style="list-style-type: none"> <li>• Ampliar as ações de apoio à internacionalização na pós-graduação das instituições contempladas</li> </ul>
<ul style="list-style-type: none"> <li>• Promover a mobilidade de docentes e discentes com ênfase em doutorandos, pós-doutorandos e docentes para o exterior e do exterior para o Brasil vinculados a programas de pós-graduação stricto sensu com cooperação internacional</li> </ul>
<ul style="list-style-type: none"> <li>• Integrar outras ações de fomento da CAPES ao esforço de internacionalização</li> </ul>

Fonte: adaptado de CAPES (2019).

Os objetivos destacam a construção e a consolidação de planos institucionais de internacionalização alinhados às áreas e temas estratégicos do conhecimento, bem como o estímulo à formação de redes de pesquisa e à elevação da qualidade da produção acadêmica, especialmente no âmbito da pós-graduação. O quadro também enfatiza a ampliação das ações de apoio à internacionalização, a promoção da mobilidade acadêmica de docentes, discentes, doutorandos e pós-doutorandos, e a integração das iniciativas da CAPES aos esforços institucionais. Em conjunto, esses objetivos reforçam o papel do PrInt como instrumento estruturante para a internacionalização sustentável, qualificada e articulada da pós-graduação brasileira.

O Quadro 14 apresenta as ações de inserção internacional, reunindo um conjunto de iniciativas estratégicas voltadas ao fortalecimento da presença e do reconhecimento internacional de pesquisadores e instituições científicas.

Quadro 14 - Ações de inserção internacional

AÇÕES DE INSERÇÃO INTERNACIONAL
<ul style="list-style-type: none"> <li>• Atrair pesquisadores de destaque de fora do país para realizarem pesquisa em cooperação</li> </ul>
<ul style="list-style-type: none"> <li>• Atrair estudantes de fora do país para seus laboratórios</li> </ul>
<ul style="list-style-type: none"> <li>• Ser do quadro editorial de periódicos internacionais indexados</li> </ul>
<ul style="list-style-type: none"> <li>• Ter posições em instituições de ensino e pesquisa no exterior</li> </ul>
<ul style="list-style-type: none"> <li>• Ter cooperações institucionais com centros de pesquisa no exterior</li> </ul>
<ul style="list-style-type: none"> <li>• Pertencer a diretorias de sociedades internacionais</li> </ul>
<ul style="list-style-type: none"> <li>• Pertencer a academias de ciências de outros países e participar de atividades internacionais</li> </ul>
<ul style="list-style-type: none"> <li>• Atrair e organizar congressos internacionais, simpósios e mesas redondas em eventos internacionais ou proferir palestras</li> </ul>
<ul style="list-style-type: none"> <li>• Rever propostas de pesquisa competitivas e publicações em revistas indexadas</li> </ul>
<ul style="list-style-type: none"> <li>• Possuir financiamentos competitivos de fontes estrangeiras</li> </ul>
<ul style="list-style-type: none"> <li>• Ser convidado para escrever revisões em revistas indexadas</li> </ul>

Fonte: adaptado de CAPES (2016).

As ações elencadas evidenciam práticas relacionadas à atração de pesquisadores e estudantes estrangeiros, à participação em corpos editoriais e diretorias de sociedades científicas internacionais, bem como ao estabelecimento de cooperações institucionais com centros de pesquisa no exterior. O quadro também destaca a importância da organização e participação em eventos científicos internacionais, da produção e publicação em periódicos indexados, da captação de financiamentos externos e do convite para atuação como revisor em revistas de impacto. Em conjunto, essas ações configuram mecanismos para ampliar a visibilidade, a circulação internacional do conhecimento e a integração das instituições brasileiras aos fluxos globais da ciência.

A literatura demonstra que as recomendações propostas visam, sobretudo, sensibilizar a comunidade científica e os usuários da informação quanto à importância estratégica de uma política nacional de acesso aberto e de uma cultura voltada à democratização do conhecimento no país. O estudo destaca que é necessário enfrentar barreiras históricas que limitam a disseminação da produção científica, propondo um manifesto fundamentado nas boas práticas de avaliação que integra o acesso aberto como princípio estruturante. Tal manifesto se configura como um instrumento para promover a

mudança cultural necessária à adoção de práticas científicas de ciência aberta (Araújo; Barcellos; Sacramento, 2024).

Os autores defendem que a democratização e a universalização do acesso ao conhecimento são condições necessárias para o desenvolvimento científico sustentável e igualitário. Assim, argumentam que uma ciência verdadeiramente global deve ser inclusiva, reconhecendo e valorizando a diversidade de contextos e práticas científicas, especialmente aquelas provenientes de países com poder econômico reduzido. Nesse sentido, a construção de uma ciência global pressupõe a redução ou mesmo a eliminação de barreiras que dificultam o acesso ao conhecimento científico (Araújo; Barcellos; Sacramento, 2024).

Contudo, os estudos evidenciam que a inclusão científica ainda é limitada por mecanismos de avaliação que privilegiam indicadores baseados majoritariamente na produção do Norte Global, além de apresentarem sub-representação de revistas não comerciais. Esse cenário reforça práticas coloniais de dominação e aprofunda desigualdades estruturais uma vez que perpetua a dependência de sistemas hegemônicos de comunicação científica (Araújo, Barcellos e Sacramento, 2024).

No contexto da pós-graduação, observa-se que a internacionalização da ciência brasileira está presente na política da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) desde os anos 1990. Estudos como o apresentado por Leão, Nogueira e Castro (2024) analisam as políticas de internacionalização do Programa Institucional de Internacionalização (CAPES-PrInt), destacando que, embora o programa contribua para a inserção internacional, ele é altamente seletivo e beneficia majoritariamente programas de excelência consolidada. Essa seletividade, por sua vez, aprofunda assimetrias regionais e institucionais, o que compromete a construção de um modelo solidário e inclusivo de internacionalização.

A literatura aponta que a internacionalização polariza debates no meio científico. Para alguns pesquisadores, ela amplia a difusão global do conhecimento. Para outros, ela é uma expressão de estruturas de poder e dominação que reproduzem lógicas de mercantilização do conhecimento,

afetando a autonomia intelectual e a diversidade cultural das nações menos favorecidas (Fioravante; Robaina; Nabozny, 2023).

Somam-se a essas questões as críticas ao atual sistema de avaliação científica, marcado pelo produtivismo quantitativista, que tem gerado tensões e impulsionado a discussão sobre a necessidade de reformulação dos critérios avaliativos. A ciência aberta se apresenta, nesse movimento, como uma alternativa capaz de fomentar práticas mais colaborativas, transparentes e coerentes com os princípios de democratização do conhecimento (Trinca; Albagli, 2023). Estudos como o de Muck e Caregnato (2023) reforçam que compreender as práticas de ciência aberta é fundamental para a formulação de políticas científicas inclusivas que ampliem o impacto social da pesquisa onde a disseminação da informação científica é relevante e necessária. A Declaração de Budapeste (2002) emerge como marco desse movimento ao definir o conceito de acesso aberto e defender as estratégias de autoarquivamento da produção científica e de desenvolvimento de periódicos em acesso aberto.

A literatura reforça que a democratização da informação científica não depende apenas da abertura de acesso, mas, de ações que promovam o letramento científico e uma cultura de participação pública no fazer científico (Pereira, 2022). Além disso, autores como Velho (2011) relacionam a evolução das políticas científicas às mudanças no conceito dominante de ciência, destacando a transição para uma visão que considera a ciência como bem público e instrumento de desenvolvimento social.

Convergindo com essa perspectiva, a UNESCO (2021) recomenda que os Estados-membros implementem políticas que favoreçam a ciência aberta, incluindo a revisão dos sistemas de avaliação e a remoção de barreiras do compartilhamento científico. Tais recomendações revelam desafios para as políticas de Ciência, Tecnologia e Inovação (CT & I), especialmente no que concerne à necessidade de mudanças culturais e articulação entre atores nacionais e internacionais.

No Brasil, a internacionalização da ciência continua sendo interpretada como uma estratégia de desenvolvimento científico, tecnológico e econômico (Araújo; Fernandes, 2021). No entanto, estudos mostram que políticas de mobilidade e cooperação, como o Programa Ciência sem Fronteiras, embora

tenham ampliado oportunidades de formação e aprendizado de idiomas, também reforçaram assimetrias globais e favoreceram países do Norte (Silva; Silva, 2021).

Por fim, a literatura destaca que a cooperação internacional exerce influência positiva na visibilidade e qualidade da pesquisa científica. Todavia, alerta para a necessidade de estratégias institucionais e governamentais que reduzam desigualdades internas e externas, promovendo uma internacionalização equitativa, inclusiva e alinhada aos princípios da ciência aberta (Carvalho; Araújo, 2020; Bortoli; Oliveira; Pezarico, 2020; Gheno *et al.*, 2020).

## 5 DELINEAMENTO METODOLÓGICO

A pesquisa científica compreende um conjunto de processos sistemáticos, práticos e empíricos relacionados ao estudo aprofundado de um fenômeno. No âmbito metodológico, o enfoque quantitativo recorre a coleta estruturada de dados para testar hipóteses, fundamentando-se na medição numérica e na análise estatística como meios para identificar padrões e sustentar teorias (Sampieri; Collado; Lucio, 2013).

Dentro desse enfoque, os estudos de alcance descritivo têm como objetivo especificar propriedades, características e traços relevantes do fenômeno investigado. Assim, os desenhos transversais descritivos realizam a coleta de dados em um único momento, permitindo identificar a incidência de modalidades, categorias ou níveis de uma ou mais variáveis em uma população definida (Sampieri; Collado; Lucio, 2013).

O presente estudo caracteriza-se como uma pesquisa de natureza empírica com abordagem quantitativa e alcance descritivo. Adotou-se um desenho de pesquisa transversal, no qual os dados foram coletados em um único momento para descrever as características e a estrutura do campo científico (Sampieri, Collado; Lucio, 2013).

A estratégia metodológica adotada para atingir os objetivos propostos no estudo é a análise bibliométrica, um método consolidado na Ciência da Informação para realizar o mapeamento científico, permitindo a identificação de padrões, tendências, redes de colaboração e o panorama do campo (Zupic; Cater, 2015).

O percurso metodológico foi estruturado em quatro etapas principais, adaptadas do fluxo de trabalho de mapeamento científico proposto por Börner, Chen e Boyack (2003) e Aria e Cuccurullo (2017):

- Desenho do estudo e seleção da fonte de dados
- Coleta e tratamento dos dados
- Análise e visualização dos dados
- Interpretação dos resultados

As seções a seguir detalham os procedimentos adotados em cada uma dessas etapas, garantindo a transparência e a reprodutibilidade da pesquisa.

## 5.1 FLUXO DE MAPEAMENTO CIENTÍFICO

O desenho de pesquisa constitui o plano de ação que orienta a execução do estudo para o alcance de seus objetivos (Sampieri; Collado; Lucio, 2013). Neste trabalho, a estratégia metodológica adotada é o mapeamento científico, uma abordagem que permite não apenas identificar a base de conhecimento e a estrutura intelectual do campo, mas também examinar a sua frente de pesquisa e as redes de colaboração que compõem a sua estrutura social (Zupic; Cater, 2015).

Para a condução do mapeamento científico, o percurso metodológico adotado foi estruturado em um fluxo de trabalho consolidado na literatura por Börner, Chen e Boyack (2003) e Zupic e Cater (2015).

A primeira etapa consiste no desenho do estudo, fase em que se define a questão de pesquisa e se especificam os métodos bibliométricos apropriados. Em seguida procede-se à coleta de dados, que envolve a seleção da base de dados e a definição dos critérios de busca para filtrar o conjunto de documentos e a exportação dos registros bibliográficos. Esta etapa de pré-processamento de dados inclui a recuperação dos dados, sua conversão para um formato compatível com as ferramentas de análise e um rigoroso processo de limpeza dos dados para corrigir inconsistências e garantir a qualidade da análise bibliométrica.

As etapas subsequentes são a análise e a visualização dos dados, nas quais se empregam ferramentas estatísticas e *softwares* de mapeamento científico para processar os registros bibliográficos e gerar representações visuais como mapas de redes científicas. Por fim, a etapa de interpretação consiste na análise descritiva dos resultados e das visualizações geradas, permitindo extrair conclusões significativas sobre a estrutura e a dinâmica do campo científico (Börner; Chen; Boyack, 2003).

O Quadro 15 apresenta uma análise comparativa do fluxo de mapeamento científico a partir das propostas metodológicas de Börner, Chen e Boyack (2003) e de Zupic e Čater (2015), evidenciando convergências e distinções entre os dois referenciais.

Quadro 15 - Análise comparativa do fluxo de mapeamento científico

AUTORES PROPÓSITO	BÖRNER; CHEN; BOYACK (2003)	ZUPIC; CATER (2015)
DESENHO DO ESTUDO	Definir a questão de pesquisa e especificar os métodos bibliométricos apropriados para o objetivo do estudo, considerando o intervalo temporal para representar a imagem estática e representacional do campo científico	Definir a problematização da pesquisa e escolher os métodos bibliométricos apropriados
COLETA DE DADOS	Selecionar a base de dados bibliográfica que contém os dados bibliométricos, filtrar o conjunto de documentos que serão objeto de análise e exportar os dados bibliométricos	Selecionar o banco de dados e filtrar o conjunto de documentos para exportar os dados bibliométricos
ANÁLISE DE DADOS	Adotar as ferramentas bibliométricas de natureza estatística para atender as necessidades do estudo	Escolher as ferramentas que atendam as necessidades computacionais da pesquisa
VISUALIZAÇÃO DE DADOS	Decidir o método de visualização usado nos resultados e empregar o <i>software</i> de mapeamento científico apropriado	Decidir o método de visualização usado para representar os resultados da pesquisa e realizar o mapeamento científico
INTERPRETAÇÃO	Interpretar e descrever as descobertas científicas	Interpretar e descrever as descobertas do campo científico

Fonte: elaborado pela autora (2025).

O quadro organiza o processo de mapeamento científico em etapas fundamentais, desenho do estudo, coleta de dados, análise de dados, visualização e interpretação, permitindo compreender como cada abordagem estrutura o percurso metodológico da pesquisa bibliométrica. Enquanto ambas enfatizam a definição clara do problema, a seleção criteriosa das bases de dados e o uso de ferramentas adequadas para análise e visualização, observam-se diferenças no nível de detalhamento conceitual e computacional atribuído a cada etapa. Assim, o quadro contribui para a sistematização do mapeamento científico como um processo analítico estruturado, servindo de referência para a escolha metodológica em estudos de natureza bibliométrica.

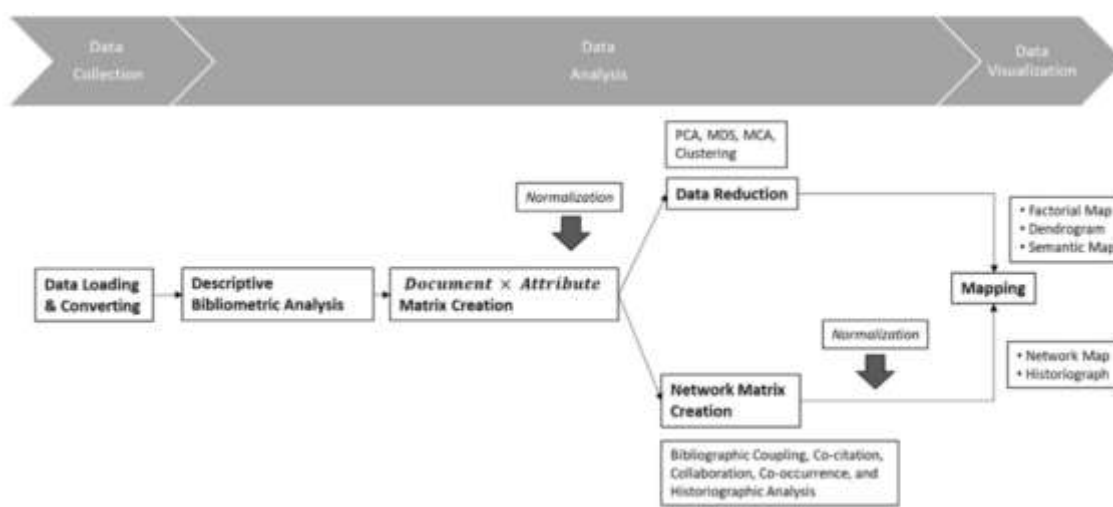
Ambos os modelos estruturam o processo em etapas: desenho do estudo, coleta, análise, visualização e interpretação dos dados. As semelhanças são evidentes em quase todas as fases, com os dois referenciais teóricos concordando sobre a necessidade de selecionar uma base de dados, filtrar documentos, escolher ferramentas de análise e, por fim, interpretar os resultados.

As distinções, embora sutis, residem no nível de detalhamento e na ênfase de cada etapa. Na fase de desenho do estudo, por exemplo, Börner, Chen e Boyack (2003) são mais prescritivos ao especificarem a necessidade de definir o "intervalo temporal" da análise, um detalhe não explicitado por Zupic e Cater (2015). Da mesma forma, na análise de dados, a primeira abordagem enfatiza a natureza "estatística" das ferramentas enquanto a segunda abordagem adota um viés mais pragmático, focando na adequação das ferramentas às "necessidades computacionais" da pesquisa.

No entanto, essas variações não alteram o fluxo de mapeamento científico, que permanece consistente entre os dois marcos teóricos, reforçando a existência de um protocolo metodológico amplamente aceito e consolidado para realizar estudos bibliométricos, conforme demonstrado na Figura 1.

A Figura 1 apresenta, de forma esquemática, o fluxo metodológico de uma análise bibliométrica, evidenciando as principais etapas que compõem o processo de coleta, análise e visualização dos dados científicos.

Figura 1 - Fluxo de mapeamento científico



Fonte: Aria e Cuccurullo (2017).

O procedimento inicia-se com o carregamento e a conversão dos dados bibliográficos, seguido pela realização de análises bibliométricas descritivas, que permitem caracterizar a produção científica sob diferentes aspectos. Em seguida, são construídas matrizes do tipo documento × atributo e matrizes de redes, as quais passam por processos de normalização para garantir a qualidade dos dados de pesquisa. A etapa de análise contempla tanto técnicas de redução de dados, como Análise de Componentes Principais (PCA), Escalonamento Multidimensional (MDS), Análise de Correspondência Múltipla (MCA) e métodos de clusterização, quanto a criação de redes baseadas em acoplamento bibliográfico, co-citação, coocorrência, colaboração e análise historiográfica. Por fim, os resultados são representados por meio de diferentes formas de mapeamento e visualização, incluindo mapas fatoriais, dendrogramas, mapas semânticos, mapas de rede e historiografias, permitindo a interpretação integrada da estrutura intelectual, conceitual e social do campo científico.

## 5.2 BASE DE DADOS BIBLIOGRÁFICA: *WEB OF SCIENCE*

A *Web of Science (WoS)*, mantida pela *Clarivate Analytics*, é uma base de dados bibliográficos multidisciplinar que se distingue por seu critério seletivo de indexação. Seu componente principal é a *Web of Science Core Collection (WoS CC)*, que agrupa diversos índices de citação especializados. Entre eles, destacam-se o *Science Citation Index Expanded (SCIE)*, o *Arts & Humanities Citation Index (A&HCI)* e o *Emerging Sources Citation Index (ESCI)*. Para os propósitos deste estudo, o escopo de análise foi delimitado ao *Social Sciences Citation Index (SSCI)*, alinhando a coleta de dados ao domínio da Ciência da Informação e das Ciências Sociais.

A seleção da fonte de dados é uma decisão metodológica em estudos bibliométricos, pois, como aponta a literatura, as bases de dados não possuem uma cobertura uniforme dos campos científicos, tornando a escolha um ato que deve ser rigorosamente justificado (Waltman, 2016; Zupic; Cater, 2015).

O Quadro 16 apresenta uma análise comparativa entre três bases de dados bibliográficas amplamente utilizadas em estudos científicos — *Library*,

*Information Science & Technology Abstracts (LISTA)*, Web of Science (WoS) e Base de Dados em Ciência da Informação (BRAPCI) —, considerando um conjunto de características relevantes para pesquisas em Ciência da Informação e para a aplicação de métodos bibliométricos.

Quadro 16 - Análise comparativa de bases de dados bibliográficas

CARACTERÍSTICA	LISTA	WoS	BRAPCI
ESCOPO E COBERTURA	Especializada e Internacional: focada exclusivamente em Biblioteconomia, Ciência da Informação e áreas correlatas  Cobertura internacional com predomínio de publicações em língua inglesa	Multidisciplinar e Internacional: abrange todas as áreas do conhecimento com categorias específicas para Ciência da Informação  Cobertura global Com viés conhecido para periódicos de alto impacto em inglês	Especializada e Nacional: focada exclusivamente na produção científica brasileira em Ciência da Informação  Cobertura nacional, indexa os principais periódicos e eventos da área no Brasil
TIPO DE CONTEÚDO	Indexa periódicos, livros, relatórios de pesquisa e anais de eventos	Indexa periódicos de alto impacto, anais de eventos e livros selecionados	Indexa artigos de periódicos e trabalhos apresentados em eventos nacionais
CRITÉRIOS DE SELEÇÃO	Curadoria focada na relevância para o campo da Ciência da Informação	Altamente Seletiva: processo de curadoria rigoroso baseado em critérios de impacto, influência e qualidade editorial, sendo considerada a "elite" das publicações	Curadoria focada na produção científica nacional em Ciência da Informação, incluindo os principais veículos de comunicação da área no Brasil
DADOS DE CITAÇÃO	Limitados ou Ausentes: por ser uma base de dados referencial e de resumo, geralmente não fornece os dados completos de referências citadas, o que a torna inadequada para análises de cocitação ou acoplamento bibliográfico	Completos e Padronizados: este é seu principal diferencial, pois fornece metadados de citação completos e estruturados para cada registro, sendo o padrão-ouro para análises de citação	Presentes, mas não Padronizados: fornece as referências dos artigos, mas a padronização pode não ser tão rigorosa quanto na WoS, o que pode exigir um trabalho maior de tratamento de dados para análises de citação
ADEQUAÇÃO À BIBLIOMETRIA	Baixa: excelente para revisões de literatura, mas inadequada para estudos bibliométricos	Excelente: considerada a base de dados ideal para estudos bibliométricos devido à qualidade,	Boa: indicada para mapear a produção científica brasileira permite análises de

CARACTERÍSTICA	<i>LISTA</i>	<i>WoS</i>	BRAPCI
	que dependem de análise de citações	profundidade e padronização dos seus metadados	produtividade e colaboração, mas análises de citação podem exigir tratamento de dados
PÚBLICO-ALVO E ACESSO	Pesquisadores e profissionais da área de Biblioteconomia e Ciência da Informação  Acesso geralmente por assinatura	Pesquisadores de todas as áreas  Acesso por assinatura	Pesquisadores, estudantes e profissionais da Ciência da Informação no Brasil  Acesso aberto e gratuito

Fonte: elaborado pela autora (2025).

A comparação abrange aspectos como escopo e cobertura, tipo de conteúdo indexado, critérios de seleção, disponibilidade e padronização de dados de citação, adequação à bibliometria e público-alvo com respectivas condições de acesso. A partir dessa sistematização, o quadro evidencia diferenças estruturais e funcionais entre bases internacionais e nacionais, permitindo compreender suas potencialidades, limitações e níveis de adequação conforme os objetivos da pesquisa, especialmente no que se refere à análise de impacto, colaboração científica e mapeamento do campo.

A análise comparativa evidencia que a escolha da base de dados depende necessariamente dos objetivos da pesquisa. Para um estudo focado exclusivamente na produção científica brasileira em Ciência da Informação, a BRAPCI é a base de dados recomendada. Entretanto, para realizar uma revisão de literatura abrangente sobre um tema na área da Ciência da Informação, a *LISTA* é uma excelente opção devido à sua cobertura especializada. Contudo, para um estudo bibliométrico de escopo internacional que visa analisar não apenas a produção, mas também a estrutura intelectual e as redes de colaboração por meio de análise de citações, a *Web of Science (WoS)* se destaca como a escolha mais adequada.

A Figura 2 apresenta a avaliação da completude dos metadados da coleção bibliométrica, composta por 406 documentos indexados no *Institute for Scientific Information (ISI)*.

Figura 2 - Validação dos metadados do corpus de referência

## Completude dos metadados -- 406 documentos do Isi

Metadados	Descrição	Contagens Ausentes	Ausente %	Status
AU	Autor	0	0,00	Excelente
DT	Tipo de documento	0	0,00	Excelente
ENTÃO	Jornal	0	0,00	Excelente
Los Angeles	Linguagem	0	0,00	Excelente
PY	Ano de publicação	0	0,00	Excelente
Banheiro	Categorias de Ciências	0	0,00	Excelente
TI	Título	0	0,00	Excelente
TC	Total de citações	0	0,00	Excelente
AB	Resumo	1	0,25	Bom
CR	Referências citadas	1	0,25	Bom
DI	DOI	8	1,97	Bom
C1	Afiliação	9	2,22	Bom
RP	Autor correspondente	9	2,22	Bom
DE	Palavras-chave	105	25,86	Pobre
EU IA	Palavras-chave Plus	110	27,09	Pobre

Fonte: dados da pesquisa (2025).

A figura sintetiza, para cada campo de metadado, a descrição correspondente, o número absoluto de registros ausentes, o percentual de ausência e o respectivo status de qualidade. Observa-se que a maioria dos metadados essenciais — tais como autor (AU), tipo de documento (DT), periódico, idioma, ano de publicação (PY), categorias de ciências, título (TI) e total de citações (TC) — apresenta completude total (0% de ausência), sendo classificada como *excelente*, o que indica elevada confiabilidade da base para análises descritivas e estruturais. Campos como resumo (AB), referências citadas (CR), DOI (DI), afiliação institucional (C1) e autor correspondente (RP) exibem percentuais reduzidos de ausência, mantendo status *bom* e não comprometendo significativamente as análises bibliométricas. Em contraste, os metadados relacionados às palavras-chave dos autores (DE) e às palavras-chave Plus (ID) apresentam percentuais elevados de ausência, sendo

classificados como *pobres*, o que sinaliza uma limitação específica para análises baseadas exclusivamente em co-ocorrência de termos. De modo geral, a figura evidencia a confiabilidade e a qualidade dos dados de pesquisa, ao mesmo tempo em que delimita com clareza os campos que requerem cautela interpretativa nas análises bibliométricas.

Portanto, a escolha da *WoS* em detrimento das outras bases é justificada pela busca de rigor metodológico, confiabilidade dos dados e pela necessidade de realizar análises que dependem intrinsecamente de metadados de citação de alta qualidade, que não são oferecidos com o mesmo nível de excelência pela *LISTA* ou pela BRAPCI.

Ao contrário de bases de dados abrangentes, a *WoS* adota um critério seletivo, incluindo apenas periódicos e publicações de alto impacto que passaram por um criterioso processo de avaliação por pares. Essa curadoria garante um padrão de qualidade e confiabilidade do *corpus* documental, atuando como um pré-requisito para realizar uma análise bibliométrica confiável. Além disso, a *WoS* se destaca pela padronização e completude dos metadados de citação, que são importantes para a execução de análises de cocitação e acoplamento bibliográfico, que buscam desvendar a estrutura intelectual do campo científico.

Embora sua cobertura multidisciplinar seja uma vantagem, para os fins deste estudo, o escopo foi delimitado ao Social Sciences Citation Index (SSCI), garantindo o alinhamento com o domínio da Ciência da Informação. Reconhece-se, contudo, a limitação da base de dados quanto a um possível viés para publicações em língua inglesa, um fator considerado na interpretação dos resultados.

O Quadro 17 sistematiza os principais critérios que fundamentam a escolha da base de dados Web of Science (*WoS*) para a realização de estudos bibliométricos e análises científicas de caráter internacional.

Quadro 17 - Critérios de escolha da Base de Dados *Web of Science*

CRITÉRIOS DE ESCOLHA DA BASE DE DADOS WEB OF SCIENCE	
QUALIDADE E CURADORIA	A <i>WoS</i> inclui apenas revistas e conferências que passam por um rigoroso processo de avaliação e curadoria que asseguram a qualidade e a

CRITÉRIOS DE ESCOLHA DA BASE DE DADOS WEB OF SCIENCE	
DOS DADOS	relevância das publicações científicas indexadas em sua base de dados bibliográfica. As fontes de informação bibliográficas são amplamente reconhecidas e respeitadas na comunidade acadêmica, garantindo que os trabalhos citados e analisados tenham passado por revisões rigorosas por pares
FERRAMENTAS DE ANÁLISE AVANÇADA	Oferece métricas detalhadas que são recomendadas para o escopo da análise bibliométrica. As métricas auxiliam na avaliação da influência e relevância das publicações e das redes de relacionamentos da comunidade científica. A WoS permite realizar análises avançadas de citações, redes de coautoria e tendências de pesquisa, proporcionando <i>insights</i> profundos sobre o impacto e a evolução de áreas do conhecimento científico
COBERTURA MULTIDISCIPLINAR	Apesar da ênfase em ciências naturais e exatas, a WoS também abrange as ciências sociais, artes e humanidades, oferecendo uma visão abrangente do panorama científico global
CONSISTÊNCIA E ATUALIZAÇÃO	As atualizações frequentes garantem o acesso às pesquisas mais recentes. A estrutura consistente dos dados facilita o processo de busca e análise, proporcionando uma experiência consolidada para os pesquisadores
INTEROPERABILIDADE E INTEGRAÇÃO	A compatibilidade com ferramentas de gestão de referências como <i>EndNote</i> e <i>RefWorks</i> facilita a organização e a citação das fontes de dados. A integração com outras bases de dados renomadas permite realizar comparações e análises cruzadas
RECONHECIMENTO INSTITUCIONAL	As instituições acadêmicas sugerem o uso da base de dados devido a sua reputação e a qualidade dos dados, influenciando o financiamento e a avaliação de pesquisas em âmbito internacional

Fonte: adaptado de Costa (2024).

O quadro destaca dimensões como a qualidade e a curadoria rigorosa dos dados, a disponibilidade de ferramentas avançadas de análise, a cobertura multidisciplinar, a consistência e atualização das informações, bem como a interoperabilidade com softwares de gestão de referências e outras bases de dados. Além disso, evidencia o reconhecimento institucional da WoS no contexto acadêmico, reforçando sua legitimidade como fonte confiável para avaliações científicas, análises de impacto, redes de colaboração e acompanhamento da evolução das áreas do conhecimento.

A *Web of Science* possui uma posição dominante no meio acadêmico devido a sua rigorosa curadoria, as ferramentas avançadas de análise, cobertura multidisciplinar, consistência dos dados e integração com outras plataformas e ferramentas de análise de dados. As características citadas apontam a WoS

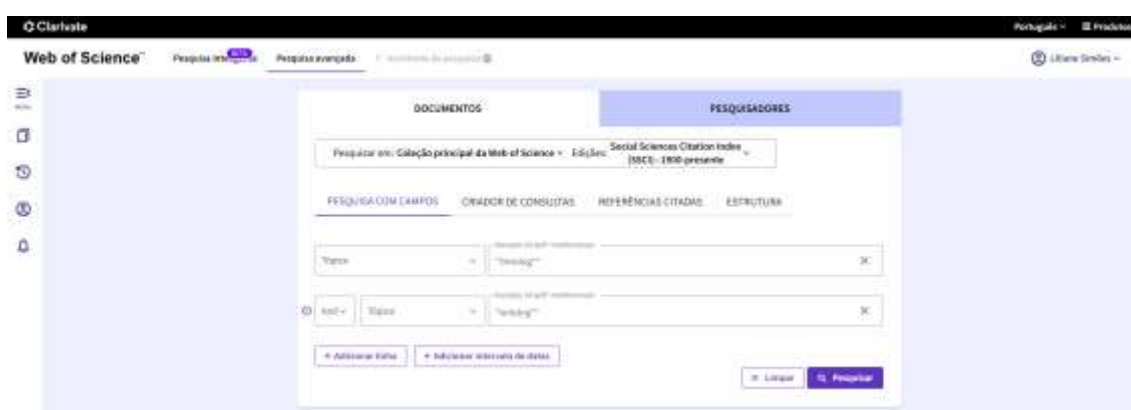
como uma escolha adequada aos critérios objetivos deste estudo, considerando a natureza interdisciplinar e de análise bibliométrica da pesquisa que exige qualidade, confiabilidade e profundidade das investigações.

### 5.3 COLETA, ANÁLISE, VISUALIZAÇÃO E INTERPRETAÇÃO DE DADOS

O mapeamento científico inicia-se com a coleta e o pré-processamento dos dados bibliométricos. Esta etapa envolve a recuperação dos registros em bases de dados bibliográficas e sua conversão para um formato adequado.

A Figura 3 ilustra a interface de pesquisa avançada da base de dados *Web of Science*, da Clarivate, evidenciando o procedimento adotado para a recuperação dos documentos que compõem o corpus da pesquisa.

Figura 3 - Coleta de dados na *Web of Science*



Fonte: dados da pesquisa (2025).

A Figura ilustra o processo de busca realizado na base de dados *Web of Science*. A consulta foi realizada no dia 21/11/2025.

Ao acessar a página da base de dados bibliográfica, navegou-se até o *menu* “Pesquisa avançada”, selecionando a aba “DOCUMENTOS”. A busca foi realizada na “Coleção principal da *Web of Science*”, considerando a “Edição - Índice de Citações em Ciências Sociais - SSCI” que indexa conteúdo desde 1900. O critério de busca adotado foi “PESQUISA COM CAMPOS”, elegendo-se o campo “Tópico” que pesquisa o título, o resumo, o *keyword plus* e as palavras-chave do autor. O descritor utilizado na busca foi “Ontolog\*” AND “ontolog\*”, a raiz do termo foi mantida e adotou-se o uso do asterisco “\*” para realizar o

truncamento da palavra, recuperando variações da palavra, abrangendo termos relacionados e ampliando o alcance da pesquisa.

Observa-se que a busca foi realizada na Coleção Principal da *Web of Science*, com recorte específico para a edição *Social Sciences Citation Index* (SSCI), abrangendo o período de 1900 até o momento presente. A estratégia de busca utiliza a opção Pesquisa com Campos, empregando o campo *Tópico* e o operador de truncamento (“\*”), aplicado ao termo *ontolog*”, de modo a recuperar variações semânticas relacionadas ao conceito de ontologia. A combinação de linhas por meio do operador lógico AND demonstra o cuidado metodológico na delimitação do escopo temático, assegurando maior precisão e abrangência dos resultados. Dessa forma, a figura explicita o processo de construção da estratégia de busca, reforçando a transparência, a reprodutibilidade e o rigor metodológico da coleta de dados bibliográficos utilizada na análise bibliométrica.

A Figura 4 apresenta a tela de resultados da pesquisa realizada na base de dados *Web of Science*, evidenciando

Figura 4 - Documentos na edição “*Social Sciences Citation Index - SSCI*”



Fonte: dados da pesquisa (2025).

Observa-se que a consulta, aplicada ao campo *Tópico* com o uso do operador de truncamento (*ontolog*), resultou em 21.359 documentos indexados na edição *Social Sciences Citation Index* (SSCI). A interface destaca, de forma clara, a expressão de busca utilizada, bem como as opções de refinamento e exploração dos resultados, como a sugestão de palavras-chave relacionadas (por exemplo, *ontology*, *social ontology*, *ontological politics* e *ontological turn*). A figura demonstra a abrangência inicial da estratégia de recuperação, evidenciando a necessidade de etapas posteriores de refinamento e filtragem para delimitação do corpus final da pesquisa. Dessa forma, a visualização

reforça a transparência do processo de coleta de dados e a consistência metodológica adotada na construção da base documental.

A Figura 5 apresenta a etapa de refinamento dos resultados da busca realizada na base de dados *Web of Science*, evidenciando a aplicação de critérios adicionais para a delimitação do corpus documental.

Figura 5 - Artigos indexados na edição “SSCI”



Fonte: dados de pesquisa (2025).

Observa-se que, após a estratégia inicial de recuperação dos registros no *Social Sciences Citation Index (SSCI)* com o termo *ontolog* no campo *Tópico*, foi aplicado o filtro referente ao tipo de documento, restringindo os resultados exclusivamente a artigos científicos. Em decorrência desse refinamento, o conjunto de documentos foi reduzido para 19.016 registros, indicando a exclusão de outros tipos documentais. A figura ilustra uma etapa metodológica do processo de coleta de dados voltada ao controle da qualidade e da homogeneidade do corpus analisado, garantindo maior consistência às análises bibliométricas.

A Figura 6 apresenta uma etapa adicional de refinamento dos resultados da busca realizada na base de dados *Web of Science*, evidenciando a aplicação de critérios para a consolidação do corpus de análise.

Figura 6 - Artigos em acesso aberto na edição SSCI



Fonte: dados da pesquisa (2025).

Após a recuperação inicial dos registros no *Social Sciences Citation Index* (SSCI) com o termo *ontolog* no campo *Tópico* e a restrição ao tipo de documento artigo, foi aplicado o filtro referente ao acesso aberto, selecionando apenas os documentos classificados como *open access*. Como resultado desse refinamento, o conjunto de publicações foi reduzido para 6.286 documentos, indicando um recorte mais específico e alinhado aos princípios de acessibilidade e transparência científica. A figura ilustra uma fase metodológica voltada à seleção de um corpus consistente, verificável e replicável, assegurando o rigor metodológico às análises bibliométricas.

A Figura 7 apresenta a etapa final de refinamento e delimitação do corpus documental utilizado na pesquisa, a partir da base de dados *Web of Science*, especificamente da edição *Social Sciences Citation Index* (SSCI).

Figura 7 - Artigos em acesso aberto na Ciência da Informação



Fonte: dados da pesquisa (2025).

Observa-se que, após a aplicação sucessiva dos filtros relacionados ao termo *ontolog* no campo *Tópico*, ao tipo de documento artigo, ao acesso aberto e à área de pesquisa *Information Science & Library Science*, o conjunto de resultados foi reduzido para 406 documentos. A figura evidencia, de forma clara, o processo metodológico de seleção rigorosa da amostra, demonstrando a convergência entre critérios temáticos, tipológicos e disciplinares. Esse procedimento assegura a coerência, a relevância temática e a consistência científica do corpus analisado, conferindo maior robustez e confiabilidade às análises bibliométricas desenvolvidas ao longo do estudo.

O Quadro 18 apresenta, de forma sistematizada, as etapas de coleta de dados realizadas na base de dados *Web of Science*, explicitando os critérios metodológicos adotados para a construção do corpus da pesquisa.

Quadro 18 - Etapas de coleta de dados na *Web of Science*

BASE DE DADOS <i>WEB OF SCIENCE</i>
COLEÇÃO PRINCIPAL
PESQUISA AVANÇADA
DOCUMENTOS
CAMPO PESQUISÁVEL: TÓPICO TÍTULO; RESUMO; PALAVRA-CHAVE PLUS E PALAVRAS-CHAVE AUTOR
DESCRITOR: "Ontolog*" AND "ontolog*"
EDIÇÃO: <i>SOCIAL SCIENCES CITATION INDEX (SSCI)</i> = 21.359 RESULTADOS
FILTRO: ARTIGO = 19.016 RESULTADOS
FILTRO: ACESSO ABERTO = 6.286 RESULTADOS
FILTRO: ÁREA DE PESQUISA "INFORMATION SCIENCE LIBRARY SCIENCE" = 406 RESULTADOS
<b>TOTAL: 406 RESULTADOS</b>

Fonte: elaborado pela autora (2025).

O quadro detalha o percurso de busca desde a seleção da coleção principal e o uso da pesquisa avançada, passando pela definição dos campos pesquisáveis (título, resumo, *keywords plus* e palavras-chave do autor) e do descritor temático utilizado, até a aplicação sucessiva de filtros relacionados à

edição da base, tipo de documento, acesso aberto e área de pesquisa. Ao explicitar cada etapa e os respectivos quantitativos de resultados, o quadro assegura transparência, reprodutibilidade e rigor metodológico ao processo de delimitação do conjunto final de documentos analisados, totalizando 406 registros que compõem o corpus de referência.

Em seguida, foi realizado o tratamento dos dados uma vez que a qualidade da análise bibliométrica depende da normalização dos dados de pesquisa que compõem o *corpus* de referência do estudo. De acordo com as boas práticas em gestão de dados bibliométricos, a normalização dos dados de pesquisa inclui a detecção e a correção de erro ortográfico e de inconsistências, a eliminação de duplicatas e a aplicação de filtros para refinar o escopo de atuação do *corpus* de referência (Aria; Cuccurullo, 2017; Zupic; Cater, 2015).

O Quadro 19 apresenta os critérios de normalização aplicados aos dados de pesquisa que compõem o corpus de referência, constituído por 406 registros.

Quadro 19 - Critérios de normalização dos dados de pesquisa

CORPUS DE REFERÊNCIA: 406 REGISTROS
DETECÇÃO E CORREÇÃO DE ERRO ORTOGRÁFICO
DETECÇÃO E CORREÇÃO DE INCONSISTÊNCIAS
ELIMINAÇÃO DE DUPLICATAS
APLICAÇÃO DE FILTROS PERSONALIZADOS PARA REFINAR O ESCOPO DA ANÁLISE

Fonte: elaborado pela autora (2025).

O quadro explicita as etapas adotadas para garantir a qualidade, a consistência e a confiabilidade dos dados, contemplando a detecção e correção de erros ortográficos, a identificação e o tratamento de inconsistências informacionais, bem como a eliminação de registros duplicados. Além disso, destaca a aplicação de filtros personalizados com o objetivo de refinar o escopo da análise, assegurando maior precisão metodológica e robustez aos resultados das análises bibliométricas.

A normalização dos dados de pesquisa foi realizada a partir da exportação do registro completo, incluindo as referências citadas. Foram exportados 406 registros no formato arquivo de texto sem formatação “.txt”. A normalização dos

dados de pesquisa buscou detectar e corrigir erros de ortografia e de inconsistência, bem como eliminar duplicatas. O arquivo que compõe o *corpus* de referência do estudo foi importado no ambiente de desenvolvimento integrado RStudio, onde foi realizado a normalização dos dados de pesquisa. Não foram encontrados arquivos duplicados visto que adotou-se uma única fonte bibliográfica para extrair os registros bibliométricos. Portanto, não houve cruzamento de dados e/ou duplicatas. O *corpus* de referência foi utilizado para refinar o escopo da pesquisa bibliométrica com a finalidade de alcançar o objetivo proposto na análise bibliométrica.

Uma vez que os dados estão normalizados, as etapas de análise, visualização e interpretação concentram-se na construção e no estudo do mapeamento científico. Existem diferentes abordagens para extrair as redes de conhecimento científico, que são compostas por nós (vértices), representando as unidades de análise (autores, palavras-chave), e por arestas (relações), que indicam a força da conexão existente entre eles.

A construção das redes científicas exige um processo de normalização que utiliza medidas de similaridade para quantificar tais relações. A partir disso, é possível aplicar diversas técnicas bibliométricas para mapear o conhecimento e a estrutura social do campo investigado.

As principais técnicas de análise bibliométrica empregadas no mapeamento científico incluem a análise da estrutura social, conceitual e intelectual do campo.

Para investigar a estrutura social, utiliza-se a análise de coautoria, que examina as redes de colaboração entre pesquisadores e suas afiliações (Glänzel, 2020). A estrutura conceitual é desvendada pela análise de co-palavras, que mapeia as relações entre os termos para revelar os principais temas e tendências de pesquisa (Callon; Courtial; Turner; Bauin, 1983). Por fim, a estrutura intelectual é explorada por meio da análise de citações, que pode ser decomposta em duas técnicas distintas: a análise de acoplamento bibliográfico, que conecta documentos com base nas referências que eles citam em comum (Kessler, 1963) e a análise de cocitação, que agrupa trabalhos com base na frequência com que são citados juntos por outros autores (Small, 1973).

## 6 FERRAMENTAS PARA ANÁLISE BIBLIOMÉTRICA

A condução de uma análise bibliométrica demanda um ambiente computacional que ofereça suporte às etapas de coleta, tratamento, análise e visualização de dados. A escolha das ferramentas adequadas é um passo metodológico relevante, pois impacta diretamente na qualidade e reprodutibilidade dos resultados de pesquisa.

Nesse contexto, este trabalho empregou um conjunto de *softwares* que de forma integrada viabilizaram o fluxo de mapeamento científico desde o pré-processamento dos registros bibliográficos até a geração dos indicadores bibliométricos das atividades científicas. As ferramentas foram selecionadas com base em sua ampla aceitação na comunidade científica de estudos métricos da informação, considerando a sua flexibilidade e a capacidade de se complementarem, garantindo um processo metodológico com rigor científico.

O fluxo de trabalho foi estruturado em duas etapas principais onde cada etapa é apoiada por *softwares* específicos.

### 6.1 BIBEXCEL

A primeira etapa, focada no tratamento e na normalização dos dados, foi realizada com o auxílio do *Bibexcel*. Este software foi desenvolvido para auxiliar no processo de análise de dados bibliográficos, permitindo a importação de registros em diversas bases de dados e a geração de arquivos de dados tabulados. Sua função nesta pesquisa foi de preparar o *corpus* bruto, extraído da *Web of Science*, para realizar a etapa de análise subsequente. O *BibExcel* permitiu a identificação e correção de inconsistências e a padronização de campos, logo, o seu uso garantiu a confiabilidade da análise bibliométrica. Embora não possua módulos próprios de visualização, a sua flexibilidade na manipulação e exportação de dados o torna uma ferramenta útil na fase de pré-processamento (Persson; Dannel; Schneider, 2009; Cobo *et al.*, 2011; Moral-Muñoz *et al.*, 2020).

## 6.2 RSTUDIO

A segunda etapa, dedicada à análise e visualização dos dados, foi desenvolvida no ambiente de desenvolvimento integrado RStudio, uma interface para a linguagem de programação R, que é um recurso da computação estatística e gráfica recomendado para estudos bibliométricos (Aria; Cuccurullo, 2017).

## 6.3 BIBLIOMETRIX

Para a execução da análise bibliométrica foi utilizado o pacote *Bibliometrix*, uma ferramenta de código aberto programada em R e projetada para realizar análises abrangentes de mapeamento científico. Este pacote oferece suporte desde o carregamento dos dados até a aplicação de diversas técnicas de análise bibliométrica (Aria; Cuccurullo, 2017).

## 6.4 BIBLIOSHINY

Complementarmente, a interface *Biblioshiny*, que é alimentada pelo *Bibliometrix*, foi empregada para a geração interativa de gráficos e visualizações de desempenho. Desenvolvida por Massimo Aria e Corrado Cuccurullo (2017), esta interface baseada na *web* facilita a exploração dos dados por meio de uma plataforma intuitiva e bem-organizada, permitindo a exportação dos resultados visuais em diversos formatos (Moral-Muñoz *et al.*, 2020).

A integração entre *RStudio*, *Bibliometrix* e *Biblioshiny* constituiu o núcleo tecnológico desta pesquisa, permitindo a transformação dos dados bibliográficos em indicadores quantitativos e representações visuais que revelam a estrutura e a dinâmica do campo científico investigado. A integração do ambiente de desenvolvimento integrado foi indispensável para alcançar os objetivos propostos no estudo e apresentar os resultados da análise bibliométrica.

## 7 ANÁLISE BIBLIOMÉTRICA

A análise bibliométrica constitui um método robusto para avaliar a produção científica de forma objetiva, sistemática e transparente, sendo amplamente reconhecida como um instrumento confiável para compreender a dinâmica, a evolução e o panorama do campo científico (Raan, 2014). Sua aplicação permite identificar padrões de comportamento, mapear redes de colaboração científica, revelar tendências emergentes e ressaltar os pilares teóricos que sustentam a área do conhecimento. Diante desse potencial metodológico, esta seção apresenta os resultados da análise bibliométrica da produção científica de ontologia em Ciência da Informação no cenário internacional.

A organização dos resultados de pesquisa foi estruturada de modo a oferecer uma visão abrangente do *corpus* analisado.

Inicialmente, o panorama do campo científico descreve a evolução temporal das publicações, destacando o ritmo de crescimento, as transições entre períodos de maior e menor produtividade e a consolidação do tema ao longo dos anos.

Na sequência, a subseção dedicada às fontes de publicações científicas identifica os periódicos mais destacados, evidenciando os veículos que concentram a maior parte da produção científica e exercem influência na difusão do conhecimento científico.

Logo, os autores são analisados em termos de produtividade, impacto e influência científica, revelando a contribuição para o desenvolvimento do campo científico. Complementarmente, a subseção de afiliações examina as instituições e centros de pesquisa, permitindo compreender a concentração institucional da produção científica e o papel das organizações na formação de redes.

A subseção de países amplia essa perspectiva ao oferecer uma visão geográfica da produção científica, analisando tanto a distribuição internacional das publicações quanto os padrões de colaboração entre os países. Esse olhar evidencia a presença dos países produtivos, sua influência global e a configuração das redes de colaboração científicas.

Na subseção de documentos, são destacados os estudos mais citados e relevantes, contribuindo para a consolidação teórica e metodológica do tema.

Em seguida, a análise de agrupamentos revela os principais *clusters* temáticos, permitindo observar como os tópicos se organizam em comunidades científicas e como autores, instituições e países interagem dentro desses agrupamentos.

A estrutura conceitual é apresentada por meio do mapeamento dos termos-chave, que compõem o núcleo temático do campo. A partir dela, avança-se para a estrutura intelectual, que identifica os referenciais teóricos, as correntes epistemológicas e os estudos que formam a base do conhecimento.

Por fim, a estrutura social evidencia as redes de coautoria, parcerias institucionais e fluxos internacionais de colaboração científica, revelando a formação de comunidades científicas e a dinâmica social da produção do conhecimento.

Portanto, essa organização articulada permite não apenas compreender a amplitude do campo, mas também visualizar a interdependência entre seus elementos, compondo, assim, um panorama científico integrado, coerente e fundamentado.

## 7.1 PANORAMA DO CAMPO CIENTÍFICO

A etapa inicial da análise bibliométrica consiste na apresentação da visão geral da coleção de documentos analisada, permitindo compreender o escopo, a composição e as características da produção científica de ontologia no âmbito da Ciência da Informação.

Nesse sentido, esta seção reúne os indicadores quantitativos e descritivos que caracterizam o *corpus* de referência, constituindo uma etapa importante para contextualizar a robustez da amostra.

A Figura 8 apresenta o painel de informações principais da coleção bibliométrica, sintetizando os indicadores que caracterizam o escopo, a dinâmica e o impacto da produção científica.

Figura 8 - Visão geral do *corpus* de referência

Fonte: dados da pesquisa (2025).

A coleção analisada abrange um período temporal de 30 anos, compreendendo publicações entre 1996 e 2026, o que permite observar a evolução do campo tanto em perspectiva histórica quanto contemporânea. Ao longo desse intervalo, foram identificados 406 documentos, distribuídos em 59 fontes de informação científica. Esse quantitativo indica uma produção moderada, mas distribuída entre diversos canais de publicação, sugerindo relativa dispersão e possível multidisciplinaridade temática.

A taxa de crescimento anual negativa (-1,34%) revela uma leve tendência de estabilidade ou declínio na produção, indicando possivelmente que o tema se consolidou e atingiu certo grau de maturidade enquanto área de pesquisa.

Essa interpretação é reforçada pela idade média dos documentos (8,75 anos), demonstrando que a literatura inclui contribuições clássicas e seminais que são importantes para a compreensão conceitual.

Em termos de impacto, os documentos da coleção apresentam média de 19,78 citações por publicação, valor considerado elevado para os padrões das Ciências Sociais Aplicadas, evidenciando a relevância e a influência dos estudos. Além disso, o conjunto inclui 16.792 referências bibliográficas, fator que potencializa análises baseadas em citações, como o acoplamento bibliográfico e a co-citação.

A coleção apresenta 1.179 palavras-chave de autor (DE) e 637 palavras-chave Plus (ID), demonstrando ampla diversidade terminológica e possibilitando análises temáticas aprofundadas. A expressiva variedade de termos indica um campo conceitualmente dinâmico, com múltiplas perspectivas de abordagem e potencial para a identificação de frentes de pesquisa emergentes.

No que diz respeito à participação dos pesquisadores, a coleção reúne 1.371 autores, dos quais apenas 67 publicaram trabalhos de autoria individual.

Esse dado, associado à média de 3,8 co-autores por documento, evidencia predominância de práticas colaborativas, característica de áreas interdisciplinares e em consolidação. A co-autoria internacional representa 23,4 % das publicações, revelando um nível moderado de cooperação entre pesquisadores de diferentes países e sugerindo oportunidades para o fortalecimento da internacionalização da área.

Por fim, a produção científica é composta exclusivamente por artigos científicos (406 documentos), sendo 8 artigos publicados em anais de conferências e 3 artigos em acesso antecipado.

Os indicadores revelam que os estudos de ontologia no campo da Ciência da Informação em escala internacional, considerando apenas os artigos publicados em acesso aberto, apresentam consolidação temática, impacto científico e diversidade conceitual, demonstrando maturidade e relevância científica.

Esses elementos fornecem uma base sólida para as análises bibliométricas, permitindo explorar com profundidade a estrutura conceitual, intelectual e social do campo científico da ontologia na Ciência da Informação.

A tabela 1 apresenta o panorama geral sobre os dados de pesquisa que foram obtidos a partir de um corpus de documentos científicos selecionados de acordo com critérios previamente definidos, abrangendo um período temporal contínuo e representativo da evolução do campo científico. O conjunto de dados contempla informações bibliográficas e bibliométricas relativas ao conteúdo dos documentos, autoria e padrões de colaboração científica, permitindo realizar uma análise sistemática da dinâmica de produção e disseminação do conhecimento científico. A partir desses dados, torna-se possível identificar tendências conceituais, características da estrutura social da pesquisa científica e o grau de interação entre os atores científicos, oferecendo subsídios para a compreensão do estágio de desenvolvimento e consolidação do campo no cenário científico internacional.

A Tabela 1 apresenta uma visão integrada dos principais indicadores bibliométricos do corpus analisado, permitindo caracterizar de forma geral a produção científica sobre ontologia no âmbito da Ciência da Informação.

Tabela 1 - Informações dos dados de pesquisa

Descrição	Resultados
<b>INFORMAÇÕES PRINCIPAIS SOBRE OS DADOS</b>	
Período de tempo	1996:2026
Fontes (Jornais, Livros, etc.)	59
Documentos	406
Taxa de crescimento anual %	-1,34
Idade média do documento	8,75
Citações médias por documento	19,78
Referências	16792
<b>CONTEÚDO DO DOCUMENTO</b>	
Palavras-chave Plus (ID)	637
Palavras-chave do autor (DE)	1179
<b>AUTORES</b>	
Autores	1371
Autores de documentos de autoria única	67
<b>COLABORAÇÃO ENTRE AUTORES</b>	
Documentos de autoria única	70
Coautores por documento	3,8
% de coautorias internacionais	23,4
<b>TIPOS DE DOCUMENTOS</b>	
artigo	395
artigo; acesso antecipado	3
artigo; anais de conferência	8

Fonte: dados da pesquisa (2025).

As informações apresentadas sobre os dados de pesquisa revelam que os aspectos relacionados ao conteúdo dos documentos, à autoria e à colaboração científica permite compreender a dinâmica conceitual e social do campo científico.

No que se refere ao conteúdo do documento, observa-se um expressivo volume de descritores, com 637 palavras-chave Plus (ID) e 1.179 palavras-chave atribuídas pelos autores (DE). Esse resultado evidencia uma elevada diversidade terminológica, indicando que o campo científico apresenta uma estrutura conceitual ampla e heterogênea marcada pela coexistência de múltiplas abordagens teóricas e metodológicas.

A predominância das palavras-chave fornecidas pelos próprios autores sugere um domínio científico no qual ainda não há completa padronização terminológica, característica recorrente em áreas interdisciplinares ou em processo de consolidação epistemológica. Essa configuração revela a presença de diferentes tradições conceituais que coexistem e se sobrepõem, contribuindo para a complexidade semântica do campo científico.

No que concerne à autoria, o corpus analisado reúne 1.371 autores, o que demonstra uma base de autores numerosa e relativamente dispersa. Apenas 67 autores apresentam produção científica de autoria única, indicando que a participação individual ou isolada não é frequente. Esse padrão sugere que a produção científica do campo não se encontra concentrada em um número restrito de pesquisadores, mas está distribuída entre diversos autores, muitos dos quais contribuem de forma pontual ou vinculada a projetos coletivos.

A colaboração entre os autores reforça essa interpretação. Do total de documentos, apenas 70 são de autoria única, enquanto a média de 3,8 co autores por documento revela uma forte tendência à produção científica colaborativa. Esse indicador aponta para a predominância de pesquisas desenvolvidas em equipes, frequentemente associadas a grupos de pesquisa institucionais ou interinstitucionais.

Observa-se que 23,4% das publicações resultam de coautorias internacionais, evidenciando um grau relevante de internacionalização da produção científica. Embora a colaboração internacional não seja majoritária, a sua presença indica a inserção do campo científico em redes científicas internacionais.

Esses resultados revelam que o campo científico apresenta uma estrutura social caracterizada por redes científicas colaborativas e relativamente densas, nas quais a produção do conhecimento científico ocorre predominantemente de forma coletiva, com baixa incidência de autoria individual. A dispersão de autores associada à elevada taxa de colaboração científica sugere um domínio científico aberto e com grupos de pesquisa atuando de maneira articulada.

Sob a perspectiva da estrutura conceitual, a grande diversidade de palavras-chave indica um campo científico em constante expansão temática, no qual diferentes conceitos, abordagens e objetos de estudo são mobilizados.

Essa multiplicidade conceitual reflete diretamente na configuração da estrutura social do campo científico, uma vez que as redes científicas com tendências colaborativas tendem a incorporar distintas tradições teóricas e perspectivas analíticas.

Assim, os indicadores bibliométricos de conteúdo, autoria e colaboração apontam para um campo científico dinâmico, interdisciplinar e em processo contínuo de consolidação, no qual a produção do conhecimento se organiza por meio de redes colaborativas e de uma estrutura conceitual ampla e em transformação.

O Quadro 20 apresenta de forma sistematizada os indicadores adotados na análise bibliométrica da pesquisa, explicitando para cada indicador sua definição operacional e a interpretação analítica no contexto dos estudos métricos da informação.

Quadro 20 - Indicadores adotados na análise bibliométrica

INDICADOR	DEFINIÇÃO	INTERPRETAÇÃO
PERÍODO DE TEMPO	O período temporal abrangido pela coleção, desde o ano de publicação mais antigo até o mais recente	Um período de tempo mais amplo permite a análise de tendências longitudinais e perspectivas históricas  Coleções que abrangem décadas são adequadas para o estudo da evolução da pesquisa e das mudanças de paradigma
FONTES	O número total de diferentes veículos de publicação representados na coleção	Um maior número de fontes sugere um campo de pesquisa multidisciplinar ou disperso  Enquanto um número menor indica concentração em alguns periódicos principais  Essa métrica é útil para identificar os principais veículos de publicação por meio da análise da Lei de Bradford
DOCUMENTOS	O número total de registros bibliográficos na coleção	Este é o tamanho de amostra fundamental para todas as análises

INDICADOR	DEFINIÇÃO	INTERPRETAÇÃO
		<p>Coleções maiores fornecem informações mais robustas, especialmente para análises de rede e de agrupamento</p> <p>(&gt;1.000 documentos)</p>
<p>TAXA DE CRESCIMENTO ANUAL</p>	<p>O aumento percentual médio no número de publicações por ano ao longo do período de vigência da coleção</p> <p>Fórmula: Taxa de Crescimento Anual Composta (CAGR) calculada como: <math>[(N_{\text{final}} / N_{\text{initial}})^{(1/\text{years})} - 1] \times 100</math></p>	<p>Uma taxa de crescimento positiva indica um campo de pesquisa em expansão</p> <p>Enquanto valores negativos ou próximos de zero sugerem maturidade ou declínio</p> <p>Altas taxas de crescimento (&gt;10%) geralmente sinalizam tópicos emergentes que atraem crescente visibilidade</p>
<p>AUTORES</p>	<p>O número total de autores únicos que contribuíram para os documentos da coleção</p>	<p>Essa métrica reflete o tamanho da comunidade de pesquisa</p> <p>Uma alta proporção de autores por documento sugere pesquisa colaborativa</p> <p>Enquanto uma baixa proporção pode indicar uma área dominada por poucos pesquisadores</p>
<p>AUTORES DE DOCUMENTOS DE AUTORIA ÚNICA</p>	<p>O número de autores que publicaram pelo menos um documento de autoria individual na coleção</p>	<p>Artigos de autoria individual são mais comuns nas ciências humanas e em disciplinas teóricas</p> <p>Uma baixa proporção sugere alta intensidade de colaboração, típica das ciências experimentais e de campos interdisciplinares</p>
<p>COAUTORIA INTERNACIONAL</p>	<p>A porcentagem de documentos elaborados por pesquisadores de vários países</p>	<p>Alta colaboração internacional (&gt;30%) indica redes de pesquisa globais e está frequentemente associada a um maior impacto de citações</p> <p>Essa métrica é um indicador de globalização da pesquisa e intercâmbio de conhecimento transfronteiriço</p>

INDICADOR	DEFINIÇÃO	INTERPRETAÇÃO
CO-AUTORES POR DOCUMENTO	O número médio de autores por documento na coleção	<p>Os valores normalmente variam de 2 (ciências sociais, humanidades) a 5 ou mais (ciências biomédicas, física)</p> <p>O aumento dos valores ao longo do tempo reflete a tendência em direção à redes científicas e a projetos colaborativos de grande escala</p>
PALAVRAS-CHAVE DO AUTOR	O número total de palavras-chave únicas fornecidas pelos autores (DE=Descritores) em todos os documentos	<p>Um conjunto rico de palavras-chave (mais de 1.000 termos únicos) permite uma análise temática robusta e a modelagem de tópicos</p> <p>A diversidade de palavras-chave reflete a amplitude conceitual do campo de pesquisa</p>
REFERÊNCIAS	O número total de referências citadas listadas nas bibliografias de todos os documentos da coleção	<p>Esta métrica é importante para análises baseadas em citações (co-citação, acoplamento bibliográfico, espectroscopia do ano de publicação da referência)</p> <p>Conjuntos de referências maiores permitem um mapeamento mais abrangente da estrutura intelectual</p>
IDADE MÉDIA DO DOCUMENTO	<p>O número médio de anos decorridos desde a publicação, calculado em relação ao ano atual</p> <p>Fórmula: Current Year - Mean (Publication Years)</p>	<p>Valores mais baixos (&lt;5 anos) indicam um foco em pesquisas recentes</p> <p>Enquanto valores mais altos sugerem a inclusão de literatura clássica ou histórica</p> <p>Essa métrica ajuda a avaliar se a coleção é contemporânea ou retrospectiva</p>

INDICADOR	DEFINIÇÃO	INTERPRETAÇÃO
NÚMERO MÉDIO DE CITAÇÕES POR DOCUMENTO	O número médio de citações recebidas pelos documentos da coleção  (com base na contagem de citações do banco de dados)	Valores mais altos indicam pesquisas de alto impacto  As taxas médias de citação variam bastante conforme a disciplina  (por exemplo, ciências biomédicas >20, ciências sociais ~10)  Essa métrica é influenciada pela idade do documento, pelas normas da área e pela abrangência da base de dados

Fonte: adaptado de Aria e Cuccurullo (2017).

O quadro reúne métricas temporais, estruturais, colaborativas e de impacto como período de tempo, fontes, documentos, taxa de crescimento anual, autores, coautoria internacional, palavras-chave, referências, idade média dos documentos e número médio de citações por documento. Ao articular esses indicadores às respectivas interpretações, o quadro evidencia como cada métrica contribui para a compreensão da dinâmica do campo científico analisado, permitindo avaliar sua evolução histórica, grau de maturidade, padrões de colaboração, diversidade temática e impacto científico, constituindo a base metodológica para as análises bibliométricas desenvolvidas ao longo do estudo.

O Quadro 21 apresenta os modos de visualização dos indicadores bibliométricos, sistematizando as principais formas de representação utilizadas na análise e comunicação de dados científicos.

Quadro 21 - Modo de visualização dos indicadores bibliométricos

QUADROS	TABELAS	GRÁFICOS	REDES	MAPAS
Contêm, predominantemente, texto ou dados qualitativos, e, diferentemente das tabelas, possuem todas as suas	São formas não discursivas de apresentar informações, onde o dado numérico se destaca como	São recursos visuais utilizados para expressar estatísticas, dados ou valores de forma visual e simplificada,	O conceito está ligado à análise da conectividade e dinâmica comportamental dos dados	É um esquema representativo de uma superfície ou parte dela, sobre a qual são apresentadas informações

QUADROS	TABELAS	GRÁFICOS	REDES	MAPAS
bordas fechadas	informação principal	facilitando a leitura, compreensão e interpretação de tendências e padrões		quantitativas e qualitativas de eventos

Fonte: adaptado de Normas de Apresentação Tabular do IBGE (1993).

O quadro distingue cinco categorias, quadros, tabelas, gráficos, redes e mapas, e descreve suas características conceituais e funcionais, evidenciando as diferenças entre representações predominantemente textuais e qualitativas, estruturas numéricas não discursivas e recursos visuais para análise estatística, representações relacionais baseadas em redes e esquemas espaciais de natureza cartográfica. Ao se fundamentar nas Normas de Apresentação Tabular do IBGE (1993), o quadro fornece um referencial metodológico para a escolha adequada do tipo de visualização, contribuindo para a clareza, a interpretação e a análise dos padrões, tendências e dinâmicas do campo científico investigado.

O Quadro 22 sintetiza as estratégias de refinamento da análise bibliométrica, explicitando os principais eixos metodológicos empregados para aprofundar e qualificar a interpretação dos dados.

Quadro 22 - Estratégias de refinamento da análise bibliométrica

ESTRATÉGIA DE REFINAMENTO	FINALIDADE
FILTROS	Refinar a coleção, aplicando filtros de metadados
FONTES	Identificar os periódicos produtivos e analisar os padrões de publicação
AUTORES	Analisar a produtividade dos autores, as redes de colaboração e as métricas de impacto
ESTRUTURA CONCEITUAL	Explorar a evolução temática e o agrupamento de tópicos por meio da co-ocorrência de palavras-chave e mapas temáticos
ESTRUTURA INTELECTUAL	Investigar redes de citação por meio de análise de co-citação

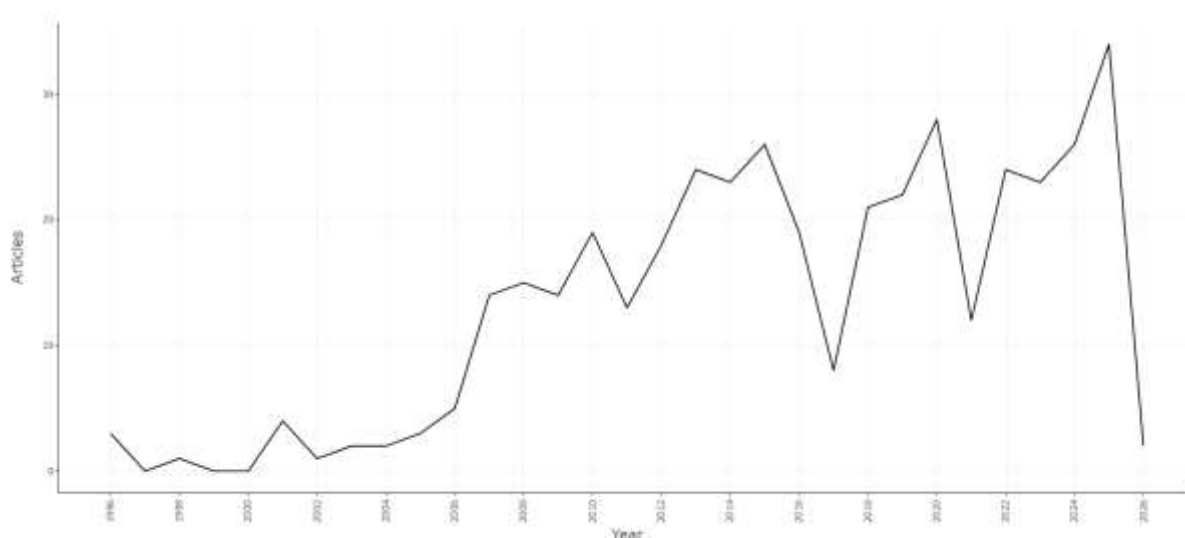
ESTRATÉGIA DE REFINAMENTO	FINALIDADE
ESTRUTURA SOCIAL	Compreender o comportamento das redes de colaboração científica

Fonte: adaptado de Aria e Cuccurullo (2017).

As estratégias apresentadas, filtros, fontes, autores, estrutura conceitual, estrutura intelectual e estrutura social, articulam-se a finalidades específicas que vão desde o ajuste do corpus por meio de metadados até a investigação das dinâmicas temáticas, intelectuais e colaborativas do campo científico. Ao destacar técnicas como a análise de produtividade, coautoria, co-ocorrência de palavras-chave e co-citação, o quadro evidencia um percurso analítico integrado, capaz de revelar padrões de publicação, relações entre pesquisadores, evolução temática e fundamentos intelectuais da área estudada.

O Gráfico 1 apresenta a evolução temporal da produção científica, evidenciando o número de artigos publicados por ano no período de 1996 a 2026, evidenciando a trajetória de amadurecimento do campo científico, marcada por expansão ao longo do tempo, ainda que com variações que refletem mudanças no interesse científico e no ritmo de publicação.

Gráfico 1 - Indicador da produtividade científica anual



Fonte: dados da pesquisa (2025).

O gráfico referente a produção científica anual revela um padrão de crescimento ao longo das últimas décadas, indicando a consolidação e a expansão do campo científico.

Nos anos iniciais (1996–2005), observa-se um volume reduzido e oscilante de publicações, característica típica de fases embrionárias de um domínio científico, quando a área ainda está se estruturando conceitualmente e por isso atrai um número limitado de pesquisadores.

A partir de 2006, contudo, evidencia-se um salto expressivo na produtividade, seguido por um movimento de estabilização, o que sugere a institucionalização do tema de pesquisa com a ampliação das redes de colaboração e a consolidação das agendas de pesquisa.

Entre 2010 e 2015, a curva apresenta um crescimento acentuado, alcançando picos superiores a 20 artigos anuais, acompanhando a tendência global de expansão da pesquisa em informação, ontologias e ciência de dados.

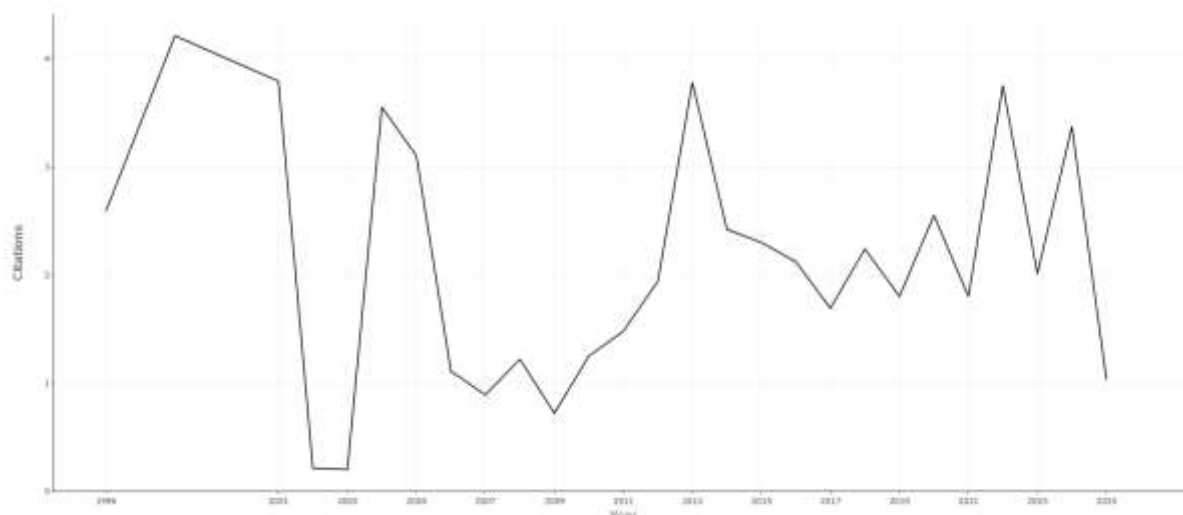
O período subsequente, embora marcado por oscilações, mantém níveis elevados de produção, com destaque para o pico observado em 2025, que supera a marca de 30 publicações, o maior valor registrado na série histórica. Tal comportamento indica não apenas a maturidade do campo, mas também a intensificação das práticas colaborativas e do volume de pesquisas, fatores que sustentam picos de produtividade.

A queda abrupta no último ponto (2026) é desconsiderada, pois representa dados incompletos para o ano corrente e não um declínio da atividade científica.

Portanto, o gráfico confirma a trajetória de ascensão contínua, típica de áreas em expansão e contribui para compreender a dinâmica temporal de desenvolvimento da produção científica.

O Gráfico 2 apresenta a evolução temporal da média de citações por artigo ao longo do período, evidenciando o comportamento do impacto da produção científica ao longo dos anos que revela um impacto relativamente consistente ao longo do tempo, refletindo a consolidação e a relevância do campo científico.

Gráfico 2 - Indicador da média de citações por ano



Fonte: dados da pesquisa (2025).

O gráfico revela um padrão oscilatório, caracterizado por períodos de maior visibilidade seguidos por quedas acentuadas, sem tendência linear clara. Nos primeiros anos (1996–2000), observa-se um crescimento inicial, indicando que os trabalhos publicados nesse período passaram a receber atenção da comunidade científica. Entretanto, entre 2001 e 2003, ocorreu uma redução brusca na média de citações, possivelmente associada ao intervalo entre novas publicações relevantes ou ao declínio do impacto dos documentos anteriores.

A partir de 2004, a média de citações volta a subir, embora de forma instável, oscilando entre picos moderados e quedas sucessivas até aproximadamente 2010. Após esse período, observa-se relativa recuperação, com aumento gradual até 2013, ano que apresenta um dos valores mais elevados da série, sugerindo reconhecimento das obras publicadas na década anterior. Nos anos subsequentes (2014–2025), o impacto anual mantém-se flutuante, com pequenos picos, como em 2020 e 2022, seguidos por reduções que reforçam a natureza irregular da série temporal.

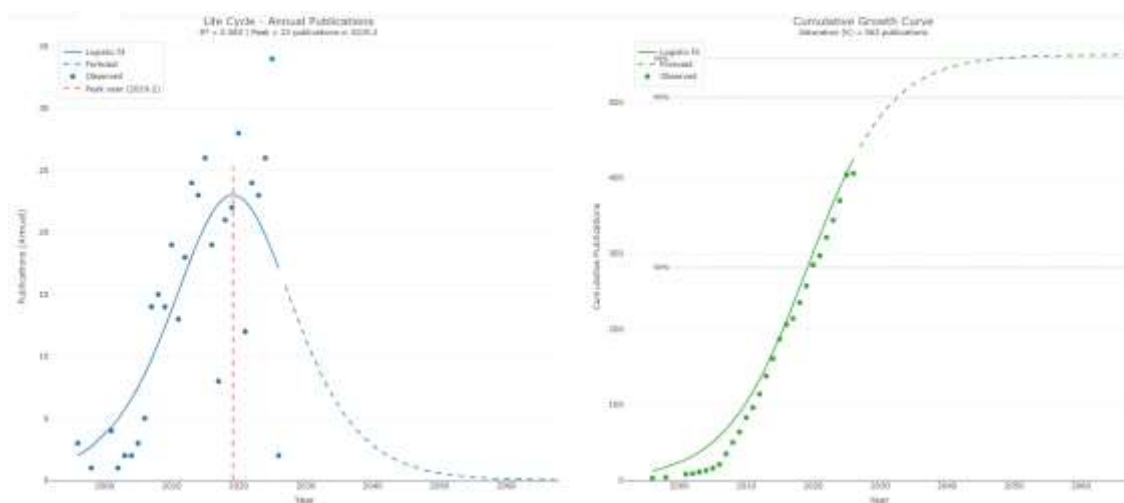
O comportamento indica que o impacto das publicações não cresce de forma contínua, mas está condicionado à dinâmica do campo científico, ao tempo de maturação das citações e à presença de trabalhos pontuais que concentram maior visibilidade. Essas oscilações refletem tanto na renovação temática quanto na dispersão de interesses dentro do campo científico, revelando um impacto temporal marcado por ciclos de atenção e períodos de menor repercussão.

O ciclo de vida da produção científica constitui uma etapa fundamental dos estudos bibliométricos, pois permite compreender a dinâmica temporal de desenvolvimento do campo científico, identificando fases de emergência, crescimento, maturidade e possível declínio. Ao examinar a evolução das publicações ao longo dos anos, torna-se possível avaliar o ritmo de expansão do domínio, reconhecer momentos de inflexão e estimar tendências futuras, contribuindo para a interpretação do estágio atual da área.

Nessa perspectiva, a análise do ciclo de vida da produção científica oferece subsídios para compreender não apenas o volume e a distribuição temporal das publicações, mas também a trajetória histórica do tema, sua consolidação e o grau de renovação temática que caracteriza seu desenvolvimento.

O Gráfico 3 apresenta a análise do ciclo de vida da produção científica por meio do ajuste de um modelo de crescimento logístico, contemplando duas representações complementares: a evolução das publicações anuais (à esquerda) e a curva de crescimento cumulativo (à direita), indicando que o campo analisado atravessou uma fase de crescimento acelerado e encontra-se próximo do estágio de maturidade, por isso, tende a apresentar estabilização no volume de publicações, caracterizando um processo típico de consolidação científica.

Gráfico 3 - Ciclo de vida da produção científica



Fonte: dados da pesquisa (2025).

Fundamentada na teoria da difusão científica e da inovação, essa abordagem se baseia no modelo de crescimento logístico, que descreve a evolução das publicações científicas ao longo do tempo, considerando a trajetória natural de um tópico desde sua emergência até sua possível saturação.

O modelo logístico ajusta uma curva ao número cumulativo de publicações do tema, indicando o comportamento evolutivo do campo conforme quatro estágios: emergência, quando o volume de publicações ainda é baixo e irregular; crescimento rápido, caracterizado pela aceleração da produção, atraindo maior visibilidade; maturidade, quando o crescimento desacelera e o campo se consolida; e declínio, momento em que o interesse científico tende à estabilização ou redução.

A curva logística, é representada pela função:

$$P(t) = \frac{K}{1 + e^{-b(t - t_0)}} \quad P(t) = \frac{K}{1 + e^{-b(t - t_0)}}$$

A curva logística permite estimar três parâmetros do ciclo evolutivo: K (saturação): estimativa do número máximo de publicações que o tema deverá produzir ao longo de sua existência; b (taxa de crescimento): indica a velocidade com que a produção científica se expande; e  $t_0$  (ano de inflexão): representa o momento em que o crescimento anual atinge o seu pico (máxima produtividade).

A partir desses parâmetros, o modelo gera quatro indicadores-chave que oferecem uma visão estratégica sobre o desenvolvimento da área, onde a Saturação (K): estima o potencial máximo de publicações, indicando que se trata de um tópico emergente, consolidado ou restrito a um nicho especializado; Ano de Pico ( $T_m$ ): revela o momento previsto de maior concentração de publicações anuais, permitindo identificar se o tema está em ascensão, estabilização ou retração; Pico Anual: quantifica o número máximo esperado de publicações por ano no momento de maior produtividade, evidenciando a intensidade do interesse no tema; Duração do Crescimento ( $\Delta t$ ): estima o intervalo de tempo entre os estágios iniciais (10% de K) e a fase de quase saturação (90% de K), sinalizando a velocidade de maturação do campo científico.

Além da previsão do comportamento futuro, o modelo permite classificar o estágio atual com base no percentual já atingido de K. Quando a produção

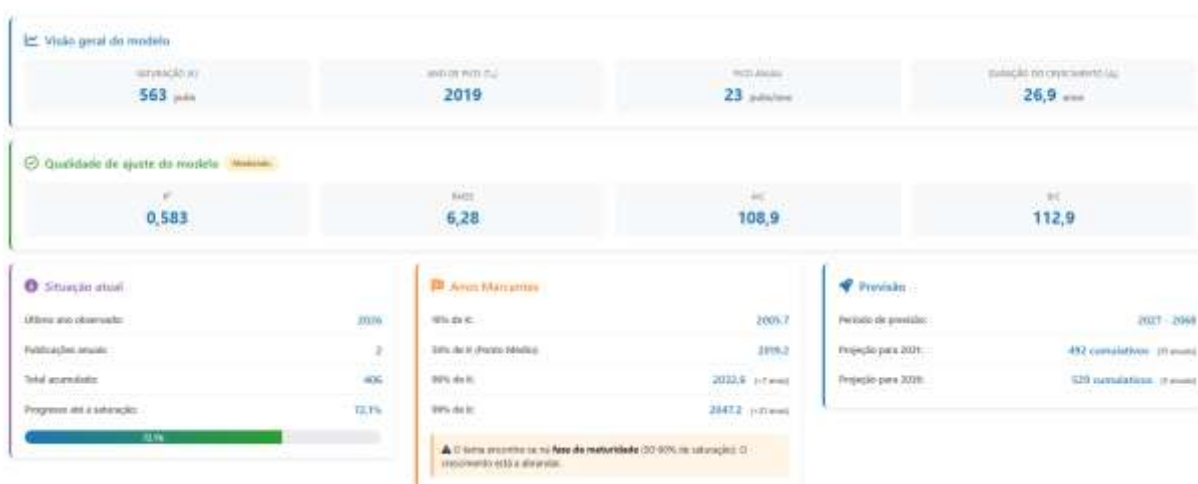
científica encontra-se abaixo de 30% de K, o tema é considerado emergente; entre 30% e 70%, está em fase de crescimento rápido; entre 70% e 90%, aproxima-se da maturidade; e acima de 90%, tende à estabilidade ou possível declínio.

Essa classificação auxilia na compreensão do posicionamento da área no ciclo evolutivo da ciência onde a avaliação da qualidade e ajuste do modelo é realizada por meio de métricas como  $R^2$ , RMSE, AIC e BIC, que indicam a aderência do modelo logístico aos dados. Valores elevados de  $R^2$  (>0,90) sugerem excelente ajuste e reforçam a validade das estimativas geradas, permitindo projeções confiáveis sobre a evolução do campo científico.

Dessa forma, a análise do ciclo de vida da produção científica é uma ferramenta importante para compreender o desenvolvimento da pesquisa, revelando seu nível de consolidação científica, o potencial de expansão futura e a demonstrando a fase atual de evolução enquanto área científica.

A Figura 9 apresenta a síntese dos resultados do modelo de ciclo de vida da produção científica, fundamentado no ajuste de uma curva de crescimento logístico, oferecendo uma visão integrada sobre a dinâmica temporal do campo científico.

Figura 9 - Estimativa da vida útil da produção científica



Fonte: dados da pesquisa (2025).

O ciclo de vida da produção científica, modelada por curvas logísticas de crescimento, permite avaliar o estágio de desenvolvimento da área e projetar sua

trajetória futura. A partir da série temporal de publicações anuais, observa-se que o campo científico passou por uma fase inicial de crescimento lento entre meados da década de 1990 e 2005, seguida de um período de expansão acelerada entre 2006 e 2018, quando ocorreu o aumento expressivo do número de publicações.

A modelagem logística indica que o pico da produtividade anual foi atingido em 2019, com aproximadamente 23 publicações, evidenciando o momento de maior intensidade. Após esse ponto, a curva ajustada sugere um movimento de declínio gradual nas décadas subsequentes, típico de campos que alcançam maturidade teórica e saturação temática. Essa dinâmica é reforçada pela curva cumulativa, que revela uma tendência em direção ao limite de aproximadamente 563 publicações, valor estimado como capacidade de saturação do domínio (K).

O ponto de máxima velocidade de crescimento, situado entre 2017 e 2018, indica que o campo já ultrapassou sua fase de difusão acelerada e ingressa em um estágio de estabilização, com previsões de que 90% da saturação sejam alcançados por volta da década de 2030 e 99% cerca de duas décadas depois.

O comportamento logístico das curvas aponta para um domínio em processo de consolidação, no qual o volume de novas publicações tende a estabilizar-se, refletindo uma estrutura temática madura, com menor emergência de frentes de pesquisa inéditas e maior prevalência de estudos incrementais ou de caráter aplicado. As projeções reforçam essa tendência, estimando crescimento cumulativo mais lento nas próximas décadas, o que evidencia a consolidação do campo científico e a transição para um estágio de estabilização da produção científica.

Sob outra perspectiva, o gráfico de três campos, também denominado *Three-Field Plot* ou diagrama de *Sankey*, constitui uma técnica de visualização amplamente empregada em estudos bibliométricos devido à sua capacidade de representar, de maneira simultânea, as relações entre três dimensões distintas de metadados. Essa abordagem permite compreender como autores, fontes de publicação, palavras-chave e referências citadas se articulam no domínio

científico, revelando padrões que contribuem para a interpretação da dinâmica intelectual, social e conceitual da área científica.

Esse tipo de visualização é composto por três colunas verticais, onde cada uma corresponde a um campo bibliográfico específico selecionado pelo pesquisador. As conexões entre os elementos desses campos são representadas por fluxos cuja espessura é proporcional à frequência de co-ocorrência entre os itens, de modo que vínculos mais robustos indicam associações recorrentes, enquanto fluxos mais estreitos sugerem relações pontuais ou especializadas.

Portanto, o diagrama permite identificar não apenas os principais atores, fontes e tópicos de pesquisa, mas também a forma como esses elementos interagem entre si.

Contudo, a interpretação do gráfico requer atenção a aspectos fundamentais, tais como a espessura dos fluxos, que indica a intensidade das associações; a centralidade dos nós, que evidencia elementos altamente conectados e, portanto, relevantes; a presença de fluxos isolados, que pode sinalizar nichos temáticos ou linhas emergentes de pesquisa; e a convergência de múltiplos fluxos para um mesmo nó, indicando potenciais interseções interdisciplinares.

Apesar de sua utilidade analítica, o gráfico apresenta limitações. Por representar, em geral, apenas os itens mais frequentes (*top-N*), essa técnica pode omitir contribuições menos recorrentes, porém significativas. Além disso, o diagrama expressa associações, mas não hierarquias ou relações causais, demandando interpretação cuidadosa por parte do pesquisador. Em casos com grande densidade de informações, a visualização pode tornar-se complexa, o que exige a seleção criteriosa do número de itens incluídos.

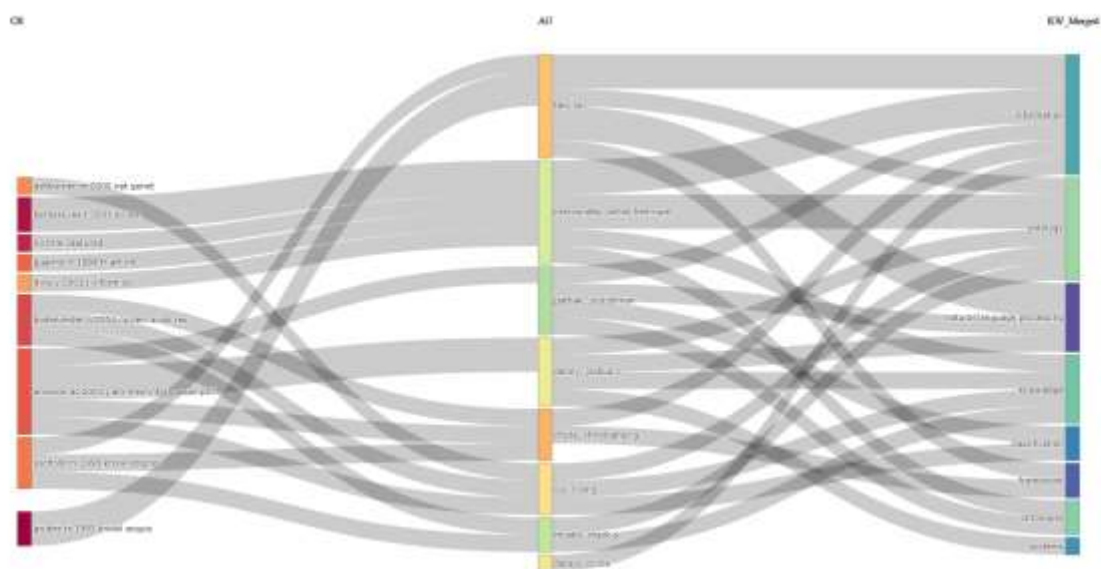
A construção do gráfico requer a seleção prévia dos metadados de interesse e, posteriormente, o processamento dos dados em *softwares* bibliométricos. Esses sistemas identificam as co-ocorrências entre os campos e geram automaticamente o diagrama de *Sankey*. Portanto, o gráfico configura-se como uma ferramenta útil para sintetizar e interpretar a estrutura intelectual e temática da literatura, contribuindo para a compreensão dos padrões que caracterizam o campo científico.

Os gráficos 4, 5, 6 e 7 demonstram as relações existentes na produção científica do *corpus* de referência, considerando os filtros de análise:

- Referências → Autores → Palavras-chave;
- Fontes → Autores → Países;
- Palavras-chave → Autores → Referências citadas;
- Países → Autores → Palavras-chave.

O Gráfico 4 evidencia as relações entre referências citadas (CR), autores (AU) e palavras-chave consolidadas (KW\_Merged) no conjunto documental.

Gráfico 4 - Análise Referências → Autores → Palavras-chave



Fonte: dados da pesquisa (2025).

O gráfico articula referências citadas (CR), autores da coleção (AU) e palavras-chave (KW\_Merged), permitindo compreender simultaneamente as bases teóricas, os atores e os eixos conceituais que estruturam o campo de pesquisa.

No primeiro campo (CR), destacam-se referências clássicas e amplamente reconhecidas, como Gruber (1993), Uschold (1996), Berners-Lee (2001) e Guarino (1998). Essas obras constituem marcos fundacionais nos estudos de ontologias, representação do conhecimento, *web semântica* e interoperabilidade, evidenciando que a área se apoia em um corpo teórico

consolidado e transdisciplinar. À esquerda, encontram-se as principais referências teóricas, que constituem a base intelectual do campo científico, destacando trabalhos clássicos e amplamente citados.

No campo intermediário (AU), observa-se que autores como *Cui Tao*, Carlos Henrique Marcondes, *Jyotishman Pathak*, *Joshua Denny*, *Christopher Chute*, *Licong Cui*, *Mark Musen* e *Cinzia Daraio* estabelecem conexões intensas com esse conjunto de referências. Isso indica que tais pesquisadores desempenham um papel na atualização, expansão e aplicação desses fundamentos teóricos, configurando-se como atores-chave na comunidade científica. No campo central, são apresentados os autores mais representativos, que atuam como mediadores entre essas referências e os temas investigados.

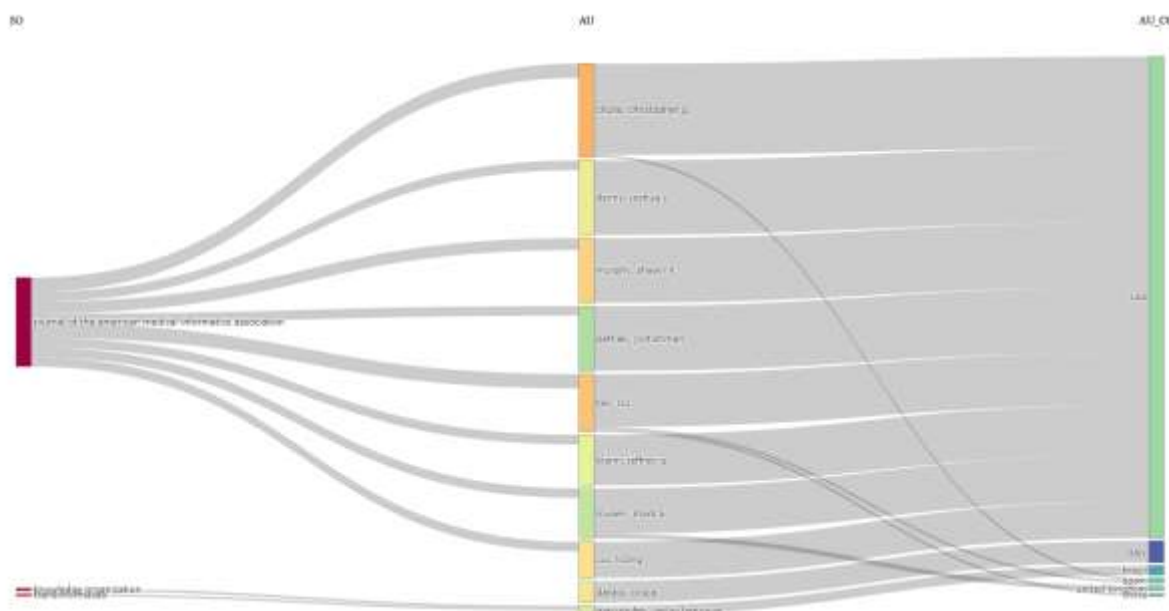
Por fim, no terceiro campo (KW\_Merged), é possível identificar os principais eixos temáticos que emergem da literatura, como *information*, *ontology*, *natural language processing*, *knowledge*, *classification*, *framework*, *ontologies* e *systems*. As ligações entre autores e palavras-chave mostram a especialização temática de cada pesquisador e revelam a convergência de suas contribuições à área. À direita, figuram as palavras-chave predominantes, que expressam os eixos conceituais da pesquisa, como *ontology*, *information*, *knowledge*, *classification* and *natural language processing*.

A largura dos fluxos indica a intensidade das associações entre os campos, revelando quais autores se apoiam em determinadas referências e como essas bases teóricas se articulam aos principais temas do domínio.

Dessa forma, a figura permite compreender, de maneira integrada, o fluxo do conhecimento, a articulação entre a estrutura intelectual e a estrutura conceitual do campo, bem como a centralidade de determinados autores e conceitos na produção científica, evidenciando um ecossistema científico fortemente integrado, no qual referências clássicas sustentam o trabalho dos autores mais influentes, que, por sua vez, convergem em torno de um conjunto bem definido de tópicos consolidados, confirmando a existência de uma coerência intelectual, ao mesmo tempo em que permite observar como teoria, pesquisadores e temas se articulam para formar o núcleo conceitual do campo científico.

O Gráfico 5 evidencia as relações entre fontes de publicação (SO), autores (AU) e países de afiliação dos autores (AU\_CO) no corpus.

Gráfico 5 - Análise Fontes → Autores → Países



Fonte: dados da pesquisa (2025).

O gráfico estabelece uma relação entre fontes de publicação (SO), autores (AU) e países de afiliação (AU\_CO), permitindo visualizar como a produção científica se distribui entre periódicos, pesquisadores e contextos geográficos.

No primeiro campo, observa-se que o *Journal of the American Medical Informatics Association (JAMIA)* constitui a principal fonte de publicações, atuando como o principal pólo de difusão do conhecimento científico. Esse periódico conecta-se a um conjunto de autores produtivos, como *Chute, Christopher G., Denny, Joshua C., Murphy, Shawn N., Pathak, Jyotishman, Tao, Cui e Musen, Mark A.*, evidenciando que parte da literatura concentra-se em um núcleo editorial específico.

No segundo campo, a rede de autores revela uma forte concentração geográfica onde a maioria dos pesquisadores está vinculada aos *Estados Unidos*, indicando a dominância desse país na produção científica. Ainda que apareçam conexões pontuais com Espanha, Brasil, China, Reino Unido e Itália,

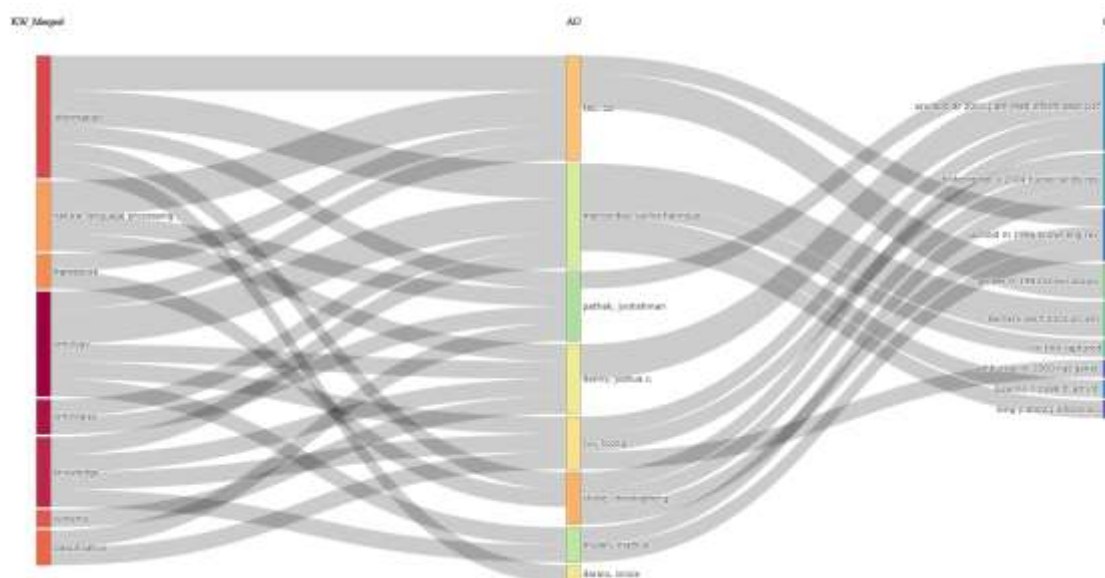
estas são menos expressivas, reforçando o domínio norte-americano no desenvolvimento das pesquisas vinculadas às fontes mais relevantes.

Por fim, a visualização demonstra a existência de um ecossistema científico, no qual um periódico de referência (*JAMIA*), um conjunto de autores de alto impacto e um país dominante (*EUA*) representam o núcleo principal da produção científica. Esse padrão sugere não apenas a concentração de capital científico, mas também a formação de comunidades de pesquisa consolidadas, que atuam de maneira interdependente, contribuindo para a continuidade e o aprofundamento das discussões no campo científico.

A largura dos fluxos indica a intensidade das relações entre os campos, permitindo compreender a estrutura social da produção científica, a centralização institucional e a distribuição geográfica dos autores envolvidos no campo.

O Gráfico 6 integra palavras-chave consolidadas (KW\_Merged), autores (AU) e referências citadas (CR), permitindo visualizar de forma articulada a estrutura conceitual e intelectual do campo.

Gráfico 6 - Análise Palavras-chave → Autores → Referências citadas



Fonte: dados da pesquisa (2025).

O gráfico revela a articulação entre os principais eixos temáticos, os autores produtivos e as referências que sustentam o domínio do conhecimento científico.

Observa-se que termos como *information*, *ontology*, *knowledge* e *natural language processing* constituem os pólos conceituais da área, estabelecendo múltiplas conexões com os autores representativos da coleção, o que evidencia a concentração temática e a consistência das linhas de pesquisa.

No campo intermediário, autores como Carlos Henrique Marcondes, *Cui Tao*, *Jyotishman Pathak*, *Joshua Denny*, *Christopher Chute* e *Mark Musen* destacam-se por sua vinculação a esses tópicos, revelando especializações e trajetórias científicas alinhadas aos núcleos conceituais do domínio.

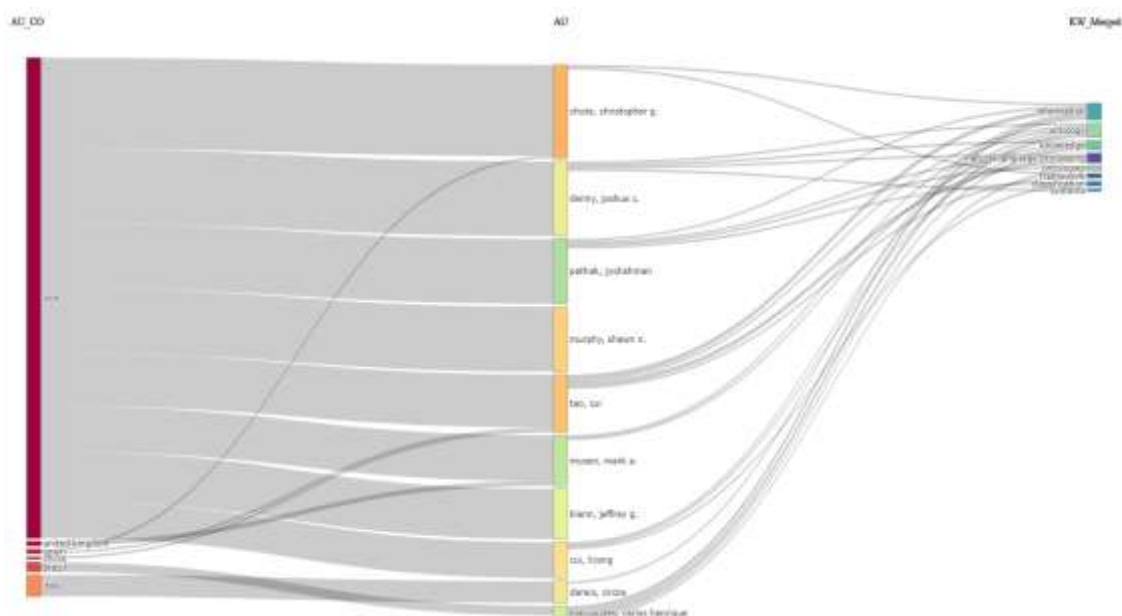
Por fim, o conjunto de referências citadas demonstra a dependência teórica de obras clássicas da área de ontologia, representação do conhecimento e *web semântica*, tais como os trabalhos de *Gruber* (1993), *Uschold* (1996), *Berners-Lee* (2001) e *Guarino* (1998).

A sobreposição dos fluxos entre tópicos, autores e referências mostra um campo consolidado, estruturado em torno de bases teóricas compartilhadas e agendas de pesquisa convergentes, permitindo compreender tanto a coesão intelectual quanto a circulação do conhecimento científico.

A largura dos fluxos indica a intensidade das relações entre conceitos, autores e referências, evidenciando como determinados temas se apóiam em bases teóricas específicas e como os autores articulam esses fundamentos em suas pesquisas. Dessa forma, a figura permite compreender o fluxo do conhecimento científico, revelando a articulação entre conceitos, autores intelectuais e obras de referência que estruturam o domínio.

O Gráfico 7 evidencia as relações entre países de afiliação dos autores (AU\_CO), autores (AU) e palavras-chave consolidadas (KW\_Merged), permitindo analisar a estrutura social articulada à dimensão conceitual da produção científica.

Gráfico 7 - Análise Países → Autores → Palavras-chave



Fonte: dados da pesquisa (2025).

O gráfico estabelece uma visualização integrada das relações entre países de afiliação dos autores (AU\_CO), autores (AU) e palavras-chave (KW\_Merged), permitindo identificar padrões na comunidade científica.

No primeiro campo, observa-se um predomínio absoluto dos Estados Unidos, que concentra a maior parte das ligações com os autores representados, evidenciando seu papel de dominância na produção científica. Outros países, como Reino Unido, Espanha, China, Brasil e Itália, apresentam conexões pontuais, mas relevantes, ao contribuírem com autores específicos.

O campo intermediário reúne os autores produtivos ou influentes. Entre eles, destacam-se *Christopher G. Chute*, *Joshua C. Denny*, *Jyotishman Pathak*, *Shawn N. Murphy*, *Cui Tao*, *Mark A. Musen*, *Jeffrey G. Klann*, *Licong Cui* e *Cinza Dario*, além de Carlos Henrique Marcondes, representando o Brasil no cenário internacional.

No terceiro campo encontram-se as principais palavras-chave, que representam os tópicos abordados por esses autores. Termos como *information*, *ontology*, *knowledge*, *natural language processing*, *framework*, *ontologies*, *classification* e *systems* revelam os eixos conceituais predominantes da área. A distribuição das linhas demonstra como os autores convergem para

determinados temas, destacando a formação de núcleos temáticos compartilhados.

Portanto, o gráfico evidencia não apenas a concentração geográfica da produção científica, mas também a especialização temática e a interconexão entre autores e conceitos-chave, oferecendo um panorama das dinâmicas intelectuais e colaborativas do campo científico.

A espessura dos fluxos mostra o volume relativo de produção associada a cada autor e sua origem geográfica. A largura dos fluxos indica a intensidade das associações entre países, autores e temas, evidenciando a centralidade dos Estados Unidos na liderança científica e a articulação temática dos autores mais produtivos. Dessa forma, a figura permite compreender de maneira integrada a distribuição geográfica da pesquisa, o protagonismo autoral e os focos conceituais predominantes do campo estudado.

## 7.2 FONTES DE PUBLICAÇÕES CIENTÍFICAS

As fontes de publicações científicas representam o principal canal de comunicação do conhecimento produzido pela comunidade científica e, por isso, desempenham um papel estratégico nos estudos bibliométricos.

A análise dessas fontes de publicações científicas permite identificar quais periódicos, conferências ou repositórios concentram a produção mais relevante, revelando tanto o núcleo editorial da área quanto a dispersão temática das pesquisas científicas.

Além disso, também possibilita compreender padrões de impacto, especialização e visibilidade científica, uma vez que diferentes periódicos exercem níveis distintos de influência na consolidação do campo científico.

Portanto, examinar as fontes de publicações científicas não apenas orienta o mapeamento das tendências científicas e da evolução da literatura, mas também evidencia os veículos que funcionam como espaços de debate e de contribuição à comunidade científica.

A Tabela 2 apresenta a distribuição das publicações por fonte, evidenciando os periódicos mais produtivos no corpus analisado.

Tabela 2 - Top 10: fontes de publicações científicas mais relevantes

Sources	Articles
JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	115
KNOWLEDGE ORGANIZATION	33
PROFESIONAL DE LA INFORMACION	20
INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE	17
TRANSINFORMACAO	14
JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	13
ETHICS AND INFORMATION TECHNOLOGY	11
JOURNAL OF DOCUMENTATION	11
JOURNAL OF INFORMATION SCIENCE	11
JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY	11

Fonte: dados da pesquisa (2025).

Observa-se forte concentração da produção científica em um conjunto restrito de periódicos, com destaque para o *Journal of the American Medical Informatics Association*, que lidera de forma expressiva com 115 artigos, indicando o protagonismo da área de Informática Médica e Saúde na produção sobre ontologias.

Em seguida, o periódico *Knowledge Organization* aparece como a principal fonte da Ciência da Informação, com 33 artigos, reforçando seu papel nas discussões teóricas e metodológicas sobre organização do conhecimento. Outros periódicos relevantes incluem *Profesional de la Información*, *International Journal of Geographical Information Science* e *Transinformação*, que, embora com menor volume, demonstram a dispersão temática e interdisciplinar do campo, abrangendo áreas como geoinformação, ética da informação e documentação.

Esse padrão de concentração e dispersão corrobora a dinâmica descrita pela Lei de Bradford, na qual poucos periódicos concentram a maior parte da produção científica, enquanto um conjunto mais amplo de fontes contribui de forma complementar para a consolidação e difusão do conhecimento no domínio estudado.

A Tabela 3 apresenta as principais fontes de publicação considerando o volume total de artigos indexados, evidenciando os periódicos com maior concentração da produção científica no conjunto de dados.

Tabela 3 - Top 10: fontes de publicações científicas locais mais citadas

Sources	Articles
J AM MED INFORM ASSN	438
LECT NOTES COMPUT SC	243
MIS QUART	211
J BIOMED INFORM	182
AMIA ANNU SYMP PROC	157
J AM SOC INF SCI TEC	147
ARXIV	134
J DOC	128
NUCLEIC ACIDS RES	123
KNOWL ORGAN	117

Fonte: dados da pesquisa (2025).

Observa-se o predomínio do *Journal of the American Medical Informatics Association (J Am Med Inform Assn)*, com 438 artigos, o que confirma o impacto da área de Informática em Saúde na produção relacionada a ontologias.

Em seguida, destacam-se *Lecture Notes in Computer Science* (243) e *MIS Quarterly* (211), indicando a forte contribuição da Ciência da Computação e dos Sistemas de Informação para o desenvolvimento teórico e aplicado do tema.

Periódicos como *Journal of Biomedical Informatics*, *AMIA Annual Symposium Proceedings* e *Journal of the Association for Information Science and Technology* reforçam o caráter interdisciplinar do campo, articulando Ciência da Informação, Computação e Ciências Biomédicas.

A presença de repositórios como *arXiv* e de periódicos clássicos da organização do conhecimento, como *Journal of Documentation* and *Knowledge Organization*, evidencia a coexistência de canais abertos de comunicação científica, revelando um ecossistema editorial diversificado.

A Tabela 4 apresenta a aplicação da Lei de Bradford às fontes de publicação do corpus analisado, organizando os periódicos segundo o ranking de produtividade (Rank), frequência absoluta de artigos (Freq), frequência acumulada (cumFreq) e sua respectiva zona de concentração.

Tabela 4 - Top 10: fontes de publicações científicas - Lei de Bradford

SO	Rank	Freq	cumFreq	Zone
JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	1	115	115	Zone 1
KNOWLEDGE ORGANIZATION	2	33	148	Zone 1
PROFESIONAL DE LA INFORMACION	3	20	168	Zone 2
INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE	4	17	185	Zone 2
TRANSINFORMACAO	5	14	199	Zone 2
JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	6	13	212	Zone 2
ETHICS AND INFORMATION TECHNOLOGY	7	11	223	Zone 2
JOURNAL OF DOCUMENTATION	8	11	234	Zone 2
JOURNAL OF INFORMATION SCIENCE	9	11	245	Zone 2
JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY	10	11	256	Zone 2

Fonte: dados da pesquisa (2025).

A tabela estabelece a distribuição da literatura científica em zonas de produtividade, permitindo identificar os periódicos dominantes da produção científica. Os dados revelam uma divisão entre Zona 1 (núcleo), composta pelos periódicos mais produtivos, e Zona 2, formada por fontes que apresentam menor, porém ainda relevante, participação.

Na Zona 1, encontram-se dois periódicos que concentram a maior frequência de publicações: *Journal of the American Medical Informatics Association (JAMIA)*, ocupando o 1º lugar com 115 publicações, consolidando-se como a principal fonte do *corpus* de referência; *Knowledge Organization*, no 2º lugar, com 33 publicações, evidenciando a relevância da organização e representação do conhecimento para o tema das ontologias. Esses dois periódicos, juntos, formam o núcleo da produção científica do *corpus* (frequência acumulada de 148 documentos), indicando que são as fontes principais da área.

A Zona 2 reúne os demais periódicos, que apresentam frequências menores, porém ainda significativas para o campo. Destacam-se: *Profesional de la Información* (20 publicações, 3º lugar); *International Journal of Geographical Information Science* (17 publicações, 4º lugar); *Transinformação* (14 publicações, 5º lugar); *Journal of the Association for Information Science and Technology* (13 publicações, 6º lugar); *Ethics and Information Technology* (11 publicações, 7º lugar); *Journal of Documentation*, *Journal of Information Science* e *Journal of the American Society for Information Science and Technology* (todos com 11 publicações, ocupando 8º, 9º e 10º lugares). Esses periódicos formam a zona de dispersão imediata do núcleo, contribuindo para o alargamento temático da produção científica. A frequência acumulada da Zona 2 alcança 256

publicações, evidenciando a importância dessas fontes na sustentação da produção científica.

A Tabela 5 apresenta os principais periódicos do corpus analisado segundo indicadores bibliométricos de impacto e produtividade, incluindo *h-index*, *g-index*, *m-index*, total de citações (TC), número de publicações (NP) e o ano inicial de publicação (PY\_start), evidenciando a coexistência de periódicos altamente produtivos e de outros com menor volume, porém com impacto significativo, revelando a diversidade estrutural e interdisciplinar do campo de estudos sobre ontologia.

Tabela 5 - Top 10: métricas de impacto local das fontes de publicações científicas

Source	h_index	g_index	m_index	TC	NP	PY_start
JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	34	53	1.133	3424	115	1996
INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE	12	17	0.480	847	17	2001
JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY	8	11	0.381	196	11	2005
JOURNAL OF THE ASSOCIATION FOR INFORMATION SYSTEMS	8	9	0.400	162	9	2006
JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	7	11	0.583	143	13	2014
JOURNAL OF DOCUMENTATION	6	10	0.300	101	11	2006
KNOWLEDGE ORGANIZATION	6	11	0.375	140	33	2010
PROFESIONAL DE LA INFORMACION	6	9	0.316	116	20	2007
PERSPECTIVAS EM CIENCIA DA INFORMACAO	5	6	0.278	51	11	2008
ETHICS AND INFORMATION TECHNOLOGY	4	11	0.250	258	11	2010

Fonte: dados da pesquisa (2025).

Observa-se que o *Journal of the American Medical Informatics Association* (JAMIA) ocupa posição de destaque, apresentando o maior conjunto de indicadores: *h-index* = 34, *g-index* = 53, *m-index* = 1,133, além de 3424 citações locais distribuídas em 115 publicações. Esses valores evidenciam sua dominância no campo científico, reforçando a liderança do periódico nas discussões sobre informática médica, ontologias e interoperabilidade semântica desde 1996, ano que marca seu início na série temporal da coleção.

Em seguida, destaca-se o *International Journal of Geographical Information Science*, com *h-index* = 12, *g-index* = 17 e 847 citações, distribuídas em 17 artigos. A presença expressiva desse periódico sublinha a relevância crescente das ontologias aplicadas ao domínio geográfico e à modelagem de informações espaciais.

Fontes da Ciência da Informação também se mostram relevantes. O *Journal of the American Society for Information Science and Technology* e o *Journal of the Association for Information Science and Technology* apresentam h-index entre 7 e 8, com *g-index* variando de 9 a 11 e impacto acumulado consistente. Essa estabilidade reforça sua relevância para as discussões teóricas e metodológicas que sustentam a representação e organização do conhecimento, temas importantes para a construção de ontologias.

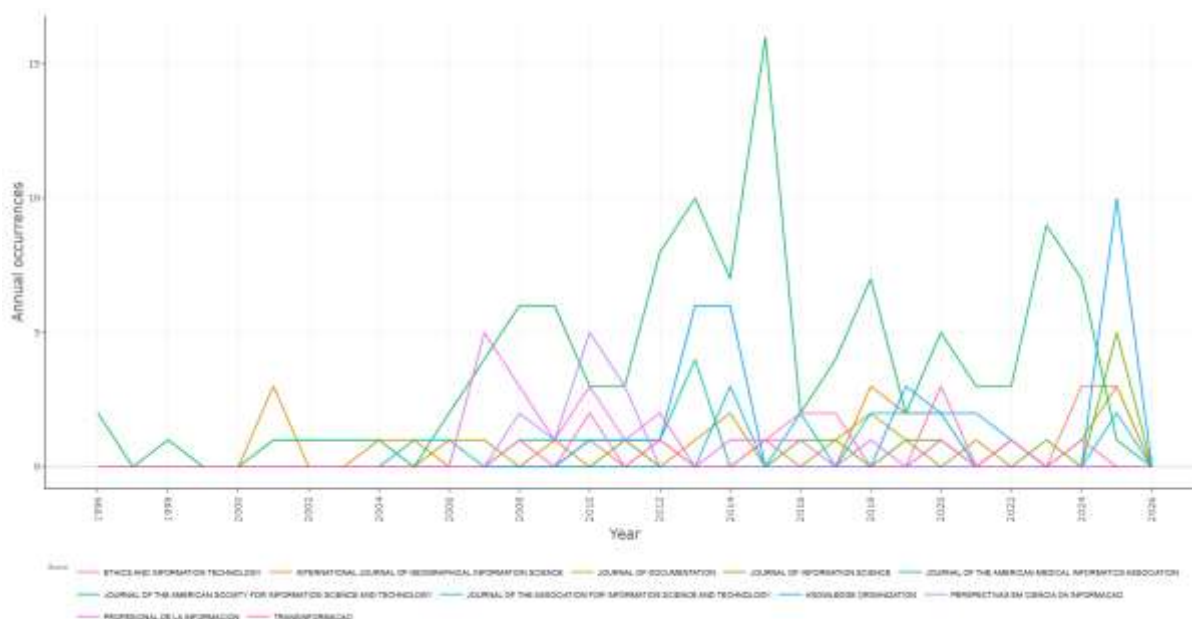
O *Journal of Documentation*, *Knowledge Organization* e *Profesional de la Información* exibem h-index igual a 6, apresentando volume significativo de citações e número expressivo de publicações (entre 11 e 33 artigos). Esses periódicos ocupam papel importante na sustentação dos debates epistemológicos e aplicados da área, contribuindo para o amadurecimento teórico do campo.

Por sua vez, o periódico brasileiro *Perspectivas em Ciência da Informação* aparece com h-index = 5 e 51 citações, mantendo-se como importante veículo regional na área de organização e gestão da informação. Embora apresente impacto inferior aos periódicos internacionais, destaca-se pela constância e pelo alinhamento temático.

Por fim, *Ethics and Information Technology*, com h-index = 4 e 258 citações, contribui introduzindo a dimensão ética e sociotécnica às discussões de ontologias e sistemas informacionais, ampliando o escopo interdisciplinar da coleção.

O Gráfico 8 apresenta a evolução temporal da frequência anual de publicações por periódico no período de 1996 a 2026, evidenciando a dinâmica das principais fontes de divulgação científica do campo.

Gráfico 8 - Evolução temporal das fontes de publicações científicas



Fonte: dados da pesquisa (2025).

Observa-se que a produção encontra-se distribuída entre um conjunto diversificado de periódicos nacionais e internacionais, com comportamentos distintos ao longo do tempo, observando-se um crescimento da produção científica ao longo da última década, especialmente em periódicos de maior relevância internacional na área de ontologias, Ciência da Informação e Ciência da Computação.

O *Journal of the American Medical Informatics Association (JAMIA)* destaca-se como a fonte com maior volume de publicações. Seus valores aumentam de forma contínua, passando de 74 artigos em 2016 para 115 artigos em 2025, indicando sua posição consolidada como principal veículo de disseminação científica em estudos que integram ontologias, interoperabilidade e sistemas de informação aplicados à área médica. Essa trajetória ascendente evidencia não apenas a crescente maturidade da área, mas também a dominância do periódico no cenário internacional.

Outras fontes de grande abrangência na Ciência da Informação, como *Knowledge Organization* e *Profissional de la Informacion*, também revelam tendência de crescimento ao longo do período. *Knowledge Organization* passou de 15 publicações em 2016 para 33 em 2025, refletindo o fortalecimento do interesse em ontologias e processos de organização do conhecimento. Já a

*Profissional de la Información* mantém produção relativamente estável, com oscilações entre 18 e 20 artigos por ano, o que indica uma continuidade temática sem grandes variações.

O *International Journal of Geographical Information Science*, embora com menor volume representativo, apresenta um incremento gradual (de 9 publicações em 2016 para 17 em 2025), o que sugere a ampliação das aplicações de ontologias no contexto da informação geográfica e dos sistemas espaciais.

Periódicos como *Transinformação*, *Ethics and Information Technology* e *Journal of Documentation* mostram movimentos de crescimento moderado. *Transinformação* evolui de 7 para 14 publicações, enquanto *Ethics and Information Technology* passa de 3 para 11, demonstrando que as discussões ética-informacionais e de representação do conhecimento científico ganham espaço no cenário internacional.

Os periódicos voltados à Ciência da Informação aplicada, como o *Journal of Information Science*, mantém crescimento contínuo, de 2 publicações em 2016 para 11 em 2025, refletindo a participação da área em pesquisas que dão suporte à construção de sistemas baseados em conhecimento.

Além disso, tanto o *Journal of the Association for Information Science and Technology (JASIST)* quanto o *Journal of the American Society for Information Science and Technology* mantêm produção estável ao longo do período, com valores uniformes (aproximadamente 11 artigos por ano), reforçando sua importância e constância como veículos consolidados da área.

Por fim, observa-se que a fonte *Perspectivas em Ciência da Informação* mantém presença consistente com 11 publicações anuais, indicando estabilidade editorial e constância temática do cenário brasileiro em âmbito internacional.

A visualização apresentada permite compreender a distribuição temporal da produção científica por fontes de publicações, destacando os periódicos mais produtivos e a evolução do protagonismo editorial no domínio estudado.

### 7.3 AUTORES

A análise de *perfis* de autores constitui uma etapa importante nos estudos bibliométricos, pois possibilita compreender a relevância científica dos pesquisadores e sua contribuição para o campo científico.

Nesse contexto, a análise do perfil do autor oferece uma abordagem analítica estruturada em dois níveis complementares: o Perfil Global e o Perfil Local. Essa dupla perspectiva permite contrastar a atuação científica geral do pesquisador com seu papel particular no recorte temático definido pela coleção bibliométrica.

O Perfil Global apresenta uma visão abrangente da produção científica do autor, englobando todas as suas publicações registradas em bases indexadas, independentemente de integrarem ou não o conjunto documental analisado. Os metadados são obtidos diretamente da *API OpenAlex*, via pacote *openalexR*, o que assegura atualização e amplitude da cobertura.

O Perfil Local, por sua vez, restringe-se às produções do autor que compõem a coleção definida para o estudo bibliométrico. Trata-se de um recorte que revela sua participação exclusivamente no domínio temático investigado, evidenciando o grau de centralidade, especialização ou complementaridade de sua contribuição.

A análise integrada dos *Perfis* Global e Local permite avaliar o alinhamento temático dos autores com a área investigada, bem como comparar a sua influência geral com a sua relevância em relação ao campo delimitado. Autores com alto impacto global, mas baixa representatividade na coleção, tendem a apresentar atuação periférica no domínio estudado; ao passo que aqueles com impacto local, mesmo que menos expressivos globalmente, podem constituir referências de dominância, demonstrando a especialidade no tema de pesquisa.

A Figura 10 apresenta o perfil global do pesquisador Maurício Barcellos Almeida, elaborado a partir de dados provenientes da base OpenAlex, oferecendo uma visão abrangente de sua produção científica e impacto acadêmico.

Figura 10 - Perfil global Almeida, M. B



Fonte: dados da pesquisa (2025).

O perfil global do pesquisador Maurício Barcellos Almeida, vinculado à Universidade Federal de Minas Gerais (UFMG), evidencia uma trajetória acadêmica consolidada no cenário internacional. O autor contabiliza 207 publicações, com um total de 534 citações, refletindo uma produção científica extensa e diversificada. Seu índice h igual a 12 indica que ao menos doze de seus trabalhos receberam doze ou mais citações, o que confirma consistência e impacto nas áreas em que atua. O índice i10 = 14 reforça essa tendência, demonstrando a presença de um núcleo significativo de publicações com visibilidade consolidada.

A citação média de dois anos (0,6) sugere um ritmo moderado, mas contínuo, de absorção da produção recente, condizente com um pesquisador que transita por múltiplas áreas temáticas e que possui um histórico extenso, com picos de impacto concentrados em diferentes momentos.

A análise das tendências de publicação dos últimos dez anos revela uma produção intensa e estável, com picos relevantes, especialmente nos anos de

2018 (21 publicações), 2016 (17 publicações) e 2021 (13 publicações). Embora haja variação natural entre os anos, o volume publicado permanece elevado e consistente, refletindo o envolvimento contínuo do autor em projetos e temáticas de pesquisa consolidadas.

Atualmente, observa-se a manutenção da produtividade, com publicações em 2023, 2024 e 2025, o que indica a estabilidade das atividades científicas. Os principais tópicos de pesquisa associados ao autor demonstram um perfil interdisciplinar, com destaque para: web semântica e ontologias (76); Ciência da Informação e bibliotecas (63); mineração de texto biomédico e ontologias (52), estudos de linguística e terminologia (24) e estudos de negócios e gestão (23).

Esse conjunto revela que o autor atua em eixos estratégicos da Ciência da Informação, articulando fundamentos teóricos da representação e organização do conhecimento com aplicações tecnológicas, como a mineração de texto biomédico, a interoperabilidade semântica e os sistemas baseados em ontologias. A amplitude temática indica não apenas diversidade, mas também profundidade em domínios que dialogam com tecnologias digitais, gestão da informação e epistemologia aplicada.

No conjunto, o perfil global de Maurício Barcellos Almeida demonstra um pesquisador com alta produtividade, impacto consolidado e atuação interdisciplinar, cuja contribuição é significativa para o avanço de áreas como web semântica, ontologias, terminologia, Ciência da Informação e estudos aplicados em ambientes digitais complexos.

A Figura 11 apresenta o perfil local do pesquisador Maurício Barcellos Almeida, considerando exclusivamente as publicações incluídas no corpus específico desta pesquisa.

Figura 11 - Perfil local Almeida, M. B



Fonte: dados da pesquisa (2025).

O perfil local do autor Maurício Barcellos Almeida, evidencia uma atuação pontual, porém alinhada ao tema da área de ontologias e organização do conhecimento. O autor possui 2 publicações dentro do *corpus* analisado, que somam 3 citações locais, resultando em um índice h local igual a 1, o que significa que pelo menos uma de suas publicações recebeu uma citação dentro da coleção. O indicador Citado/Ano igual a 1 sugere um impacto proporcionalmente moderado, considerando o número reduzido de documentos.

Com 15 anos de atividade científica, observa-se que a sua representatividade em relação ao *corpus* de referência ocorre de forma espaçada, com atividade registrada apenas no último ano, indicando que a restrição de incluir apenas artigos em acesso aberto limitou o escopo de análise bibliométrica do *corpus* referência. O gráfico de tendência das publicações nos últimos anos reforça esse padrão, mostrando apenas um documento publicado em 2025 e nenhuma produção registrada nos anos anteriores.

As palavras-chave associadas ao autor reitera sua inserção em áreas da Ciência da Informação, tais como, sistemas de informação, ontologia, organização do conhecimento, modelos conceituais, teoria da classificação, ontologias e representação do conhecimento. Esses termos revelam que,

embora sua participação quantitativa seja limitada no *corpus*, ela está vinculada a domínios teóricos relevantes e alinhados à discussão epistemológica e metodológica do campo científico.

O perfil local de Maurício Barcellos Almeida mostra uma contribuição concentrada e temática, com presença relevante em tópicos teóricos, metodológicos e aplicados da área.

A Figura 12 apresenta o perfil global do pesquisador Carlos Henrique Marcondes, vinculado à Universidade Federal Fluminense, com base em dados agregados do OpenAlex.

Figura 12 - Perfil global Marcondes, C. H



Fonte: dados da pesquisa (2025).

O perfil global do pesquisador Carlos Henrique Marcondes, docente da Universidade Federal Fluminense (Brasil), destacando sua trajetória científica consolidada no cenário internacional. O autor contabiliza 98 publicações indexadas, que totalizam 328 citações, resultando em um índice h igual a 10, valor que indica a existência de pelo menos dez trabalhos com dez ou mais

citações, evidenciando impacto relevante e continuado em suas áreas de atuação. O índice  $i_{10}$  igual a 11 reforça essa consistência, apontando a presença de um conjunto estável de publicações referenciadas.

A citação média de dois anos (0,2) sugere que, embora o autor possua impacto cumulativo expressivo, sua produção científica apresenta ritmo moderado de absorção citacional, comportamento comum em pesquisadores com trajetória longa e diversificada, cujas contribuições mais influentes podem estar situadas em períodos anteriores.

A análise das tendências de publicação nos últimos anos revela picos produtivos importantes, especialmente nos anos de 2018 (7 publicações) e 2021 (10 publicações), indicando momentos de maior engajamento em projetos e linhas de pesquisa específicas. A produção se mantém consistente ao longo da década, com variações naturais, e apresenta queda gradual a partir de 2022, refletindo mudanças de foco, maturidade temática ou outros fatores institucionais e profissionais.

Os principais tópicos de pesquisa associados ao autor demonstram inserção nas áreas de Ciência da Informação e bibliotecas (46 ocorrências); *web semântica* e ontologias (32 ocorrências); gestão de arquivos digitais e tradicionais (13 ocorrências), mineração de texto biomédico e ontologias (11 ocorrências) e educação e tecnologias digitais (12 ocorrências).

Esses temas revelam uma atuação interdisciplinar que articula fundamentos teóricos da organização e representação da informação com abordagens aplicadas relacionadas à *web semântica*, digitalização de acervos e tecnologias educacionais. A predominância dessas áreas indica que o autor tem contribuído tanto para o avanço dos fundamentos epistemológicos da Ciência da Informação quanto para sua aplicação em contextos digitais e tecnológicos complexos.

O perfil global de Carlos Henrique Marcondes retrata um pesquisador com impacto consolidado e atuação diversificada, cuja produção científica dialoga com temas da Ciência da Informação, especialmente no que diz respeito às ontologias, à representação do conhecimento e à gestão de informação em ambientes digitais.

A Figura 13 apresenta o perfil local do pesquisador Carlos Henrique Marcondes, considerando exclusivamente as publicações incluídas na coleção analisada neste estudo.

Figura 13 - Perfil local Marcondes, C. H



Fonte: dados da pesquisa (2025).

O perfil local do autor Carlos Henrique Marcondes, evidenciando seus indicadores de produtividade e impacto no *corpus* analisado.

O autor possui 4 publicações, que totalizam 14 citações locais, resultando em um índice h local igual a 2, o que indica que pelo menos duas de suas publicações receberam duas ou mais citações dentro da coleção analisada. O indicador Citado/Ano (0,3) revela um ritmo de impacto relativamente baixo ao longo do tempo, o que é coerente com o intervalo temporal e a especificidade temática da coleção.

A figura destaca ainda que o autor possui 8 anos de atividade científica, com uma média de 3,5 citações por unidade de trabalho, embora apresente atividade nula nos últimos 5 anos, sugerindo uma fase de menor engajamento em publicações relacionadas ao tema no período atual.

Em relação às tendências de publicação, observa-se que a sua produção concentra-se em momentos pontuais, com picos nos anos 2008, 2013 e 2015, onde foram registrados os maiores números de publicações, demonstrando a sua contribuição em relação ao *corpus* de referência. Esses períodos parecem coincidir com a sua participação em projetos específicos que exigem envolvimento em linhas de pesquisa voltadas à representação e comunicação científica.

Por fim, as palavras-chave associadas à sua produção científica revelam áreas de atuação vinculadas aos domínios da comunicação científica, ontologia, representação do conhecimento, metodologia científica e informação, entre outros termos relevantes. Esse conjunto evidencia a contribuição do autor para temas da Ciência da Informação, especialmente relacionados à epistemologia, organização do conhecimento, processos comunicacionais e fundamentos ontológicos.

O perfil local indica que Carlos Henrique Marcondes é um autor com participação significativa, contribuindo para eixos teóricos da comunidade científica.

A Tabela 6 apresenta os autores mais produtivos da coleção analisada, considerando tanto o número absoluto de artigos quanto o indicador de artigos fracionados, que pondera a produção individual em função da coautoria.

Tabela 6 - Top 10: autores mais relevantes

Autor	Artigos	Artigos Fracionados
<a href="#">CHUTE, CHRISTOPHER G.</a>	6	1,03
<a href="#">TAO, CUI</a>	6	0,72
<a href="#">MURPHY, SHAWN N.</a>	5	0,73
<a href="#">CUI, LICONG</a>	4	0,66
<a href="#">DARAIO, CINZIA</a>	4	1,01
<a href="#">DENNY, JOSHUA C.</a>	4	0,63
<a href="#">KLANN, JEFFREY G.</a>	4	0,63
<a href="#">MARCONDES, CARLOS HENRIQUE</a>	4	2,70
<a href="#">MUSEN, MARK A.</a>	4	0,47
<a href="#">PATHAK, JYOTISHMAN</a>	4	0,47

Fonte: dados da pesquisa (2025).

Os autores mais relevantes apresentam os pesquisadores que se destacam no *corpus* tanto pela produtividade (número de publicações) quanto pelo impacto acumulado, considerando as citações e a continuidade da produção científica ao longo do tempo. Em geral, esses autores constituem o núcleo ativo do campo científico, desempenhando um papel importante na consolidação teórica, metodológica e aplicada das pesquisas científicas.

Autores como *Christopher G. Chute, Tao Cui, Shawn N. Murphy, Joshua C. Denny, Jeffrey G. Klann, Mark A. Musen, Jyotishman Pathak, Licong Cui e Qian Zhu* sobressaem pela recorrência com que aparecem em diferentes indicadores bibliométricos, seja nas métricas de produtividade, nas citações locais e globais ou nos índices h, g e m. Essa participação consistente confirma a sua posição de liderança intelectual, especialmente no contexto da informática biomédica, área dominante no *corpus* analisado.

A relevância desses autores também é reforçada pela continuidade temporal da produção científica, com contribuições significativas publicadas ao longo de vários anos, frequentemente associadas a instituições de excelência como *Stanford, Harvard, Mayo Clinic, Vanderbilt e University of Texas*. Esses centros de pesquisa sustentam linhas de investigação colaborativas, refletindo no impacto de produtividade dos autores.

A presença de pesquisadores de diferentes países, embora com predominância dos Estados Unidos, evidencia o caráter internacional e interdisciplinar do campo científico, que reúne competências em Ciência da Informação, engenharia do conhecimento, sistemas de saúde, interoperabilidade semântica e ciência de dados.

A tabela demonstra que os autores mais relevantes são aqueles que combinam elevada produtividade, impacto expressivo e atuação contínua, configurando o grupo que contribui para a consolidação e evolução da pesquisa no cenário internacional, evidenciando que a estrutura de autoria do campo combina lideranças individuais com forte dinâmica de colaboração científica, característica de áreas interdisciplinares como os estudos sobre ontologia.

A Tabela 7 apresenta os autores com maior número de citações locais na coleção analisada, indicador que reflete o grau de reconhecimento e influência dos pesquisadores dentro do próprio *corpus* do estudo.

Tabela 7 - Top 10: autores locais mais citados

Autor	Citações locais
CHUTE, CHRISTOPHER G.	10
JIANG, GUOQIAN	6
PATHAK, JYOTISHMAN	6
HUFF, STANLEY M.	5
KLANN, JEFFREY G.	5
MURPHY, SHAWN N.	5
ONIKI, THOMAS A.	5
TAO, CUI	5
ZHU, QIAN	5
BARTOLUCCI, ALESSANDRO	4

Fonte: dados da pesquisa (2025).

Os autores locais mais citados evidenciam os pesquisadores que exercem maior influência dentro do *corpus* analisado, isto é, aqueles cujos trabalhos são frequentemente referenciados pelos demais autores da amostra.

No topo da lista encontra-se *Christopher G. Chute*, com 10 citações locais, destacando-se como o pesquisador de maior impacto interno. Sua liderança reforça a sua relevância no campo da informática biomédica e na construção de ontologias aplicadas à saúde, áreas frequentemente representadas nas publicações analisadas.

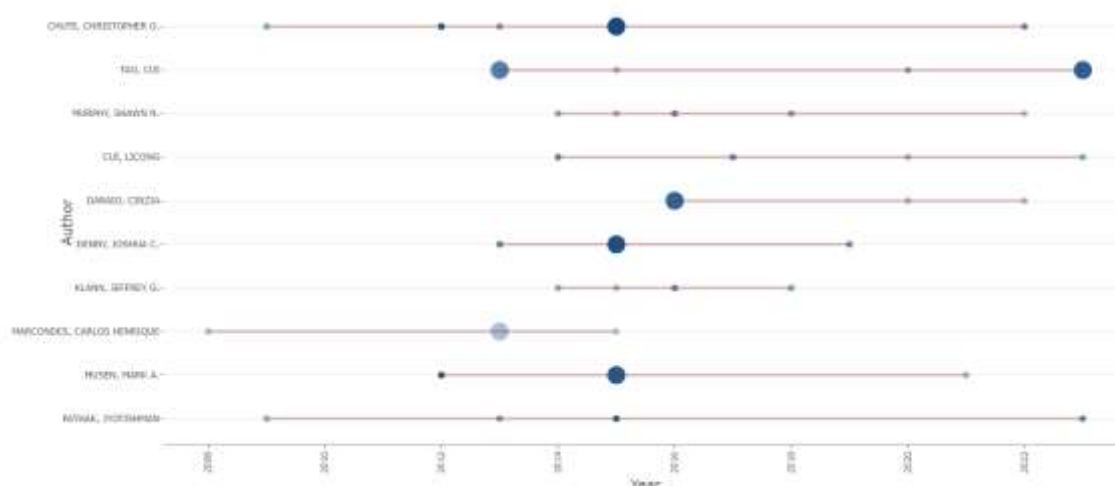
Em seguida, aparecem *Guoqian Jiang* e *Jyotishman Pathak*, ambos com 6 citações locais, indicando sua posição como importantes referências conceituais e metodológicas dentro do subconjunto de estudo. A seguir, há um grupo de autores com 5 citações locais, composto por *Stanley M. Huff*, *Jeffrey G. Klann*, *Shawn N. Murphy*, *Thomas A. Oniki*, *Tao Cui* e *Qian Zhu*. Esse agrupamento demonstra a existência de um núcleo de pesquisadores cujos trabalhos são regularmente utilizados nas discussões científicas relacionadas ao *corpus*.

Finalmente, *Alessandro Bartolucci*, com 4 citações locais, também figura entre os autores mais influentes, embora em menor intensidade. No conjunto, a lista revela um padrão: a predominância de pesquisadores vinculados à informática médica, à engenharia ontológica e à ciência de dados em saúde, indicando que essas áreas exercem um papel importante na sustentação temática da literatura analisada.

Essa distribuição confirma que a influência local concentra-se em um grupo relativamente pequeno de autores, responsáveis por oferecer as bases teóricas e metodológicas que orientam a produção científica incluída na análise bibliométrica, evidenciando uma estrutura intelectual concentrada, na qual poucos autores exercem papel de liderança, ao mesmo tempo em que há um núcleo ampliado de pesquisadores com impacto relevante no desenvolvimento temático do campo.

A Figura 14 apresenta a distribuição temporal da produção científica dos principais autores da coleção, evidenciando a evolução de suas publicações ao longo do período analisado.

Figura 14 - Evolução temporal da produção científica dos autores



Fonte: dados da pesquisa (2025).

A figura apresenta a evolução temporal da produção científica dos autores, evidenciando a trajetória individual e a intensidade produtiva ao longo dos anos. O eixo horizontal representa os anos de publicação, enquanto o eixo vertical lista os autores, sendo que os pontos indicam ocorrências de publicações e o tamanho dos marcadores reflete a intensidade ou relevância da produção científica em determinados anos.

Observa-se que grande parte desses pesquisadores iniciou sua contribuição científica entre 2008 e 2014, o que indica que o desenvolvimento do campo científico é relativamente recente e fortemente associado à expansão das pesquisas em informática biomédica e Ciência da Informação.

Autores como *Christopher G. Chute, Tao Cui, Mark A. Musen, Shawn N. Murphy, Joshua C. Denny e Cui Licong* apresentam continuidade produtiva ao longo dos anos, com pontos específicos de maior intensidade, representados pelas bolhas de maior tamanho, correspondendo a artigos que receberam maior número de citações. Esses padrões sugerem um impacto crescente e recorrente desses autores no desenvolvimento do tema, com momentos de impacto e relevância científica.

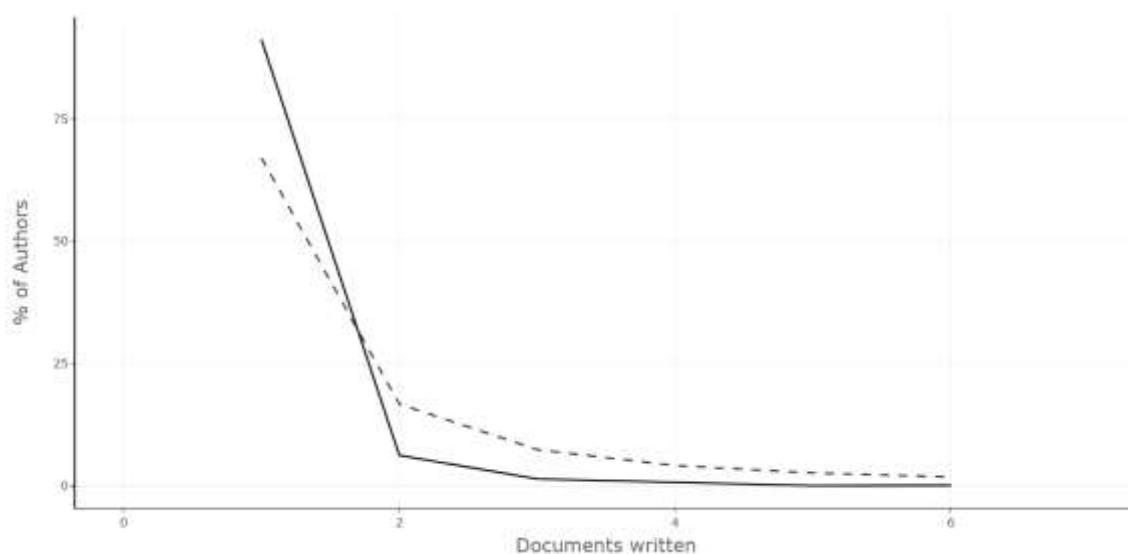
Além disso, o gráfico evidencia que autores como *Cinzia Daraio, Jeffrey G. Klann, Jyotishman Pathak e Carlos Henrique Marcondes* possuem trajetórias produtivas concentradas em determinados períodos, publicando em intervalos específicos, o que pode refletir a participação em projetos colaborativos ou a inserção em linhas de pesquisa delimitadas temporalmente.

A extensão das linhas de cada autor ao longo dos anos indica a sua permanência no campo científico, enquanto a presença e o tamanho relativo das bolhas revelam não apenas a produção anual, mas também o impacto de cada publicação. A proximidade temporal das produções científicas sugere a formação de um núcleo ativo de pesquisadores que contribuíram de maneira significativa para a consolidação conceitual e aplicada do campo científico.

A figura demonstra que a produção dos autores é marcada por trajetórias heterogêneas, mas convergentes em direção a uma expansão contínua do campo científico, com destaque para pesquisadores que apresentam constância, impacto e presença significativa em diferentes momentos da evolução do campo científico.

O Gráfico 9 apresenta a distribuição da produtividade dos autores da coleção analisada, expressa pela relação entre o número de documentos publicados (eixo horizontal) e o percentual de autores correspondentes (eixo vertical).

Gráfico 9 - Indicador da produtividade do autor através da Lei de Lotka



Fonte: dados da pesquisa (2025).

O gráfico apresenta a distribuição da produtividade dos autores através da Lei de Lotka, que prevê que a maioria dos pesquisadores produz apenas um único artigo, enquanto poucos autores são responsáveis por um volume maior de publicações.

Os resultados observados no *corpus* confirmam essa tendência clássica dos estudos bibliométricos. Nota-se que uma proporção extremamente elevada de autores, superior a 70%, contribuiu com apenas um documento, representando a base mais larga da distribuição e indicando que a produção científica é caracterizada por pesquisadores ocasionais ou de baixa produtividade acumulada.

À medida que aumenta o número de documentos publicados, observa-se um declínio abrupto no percentual de autores, coerente com o comportamento esperado pelo modelo inverso quadrático de Lotka. Autores com dois artigos já representam uma parcela significativamente menor, aproximando-se de cerca de 20%. Quando o número de publicações chega a três ou mais documentos, a participação cai para níveis residuais, abaixo de 10%, refletindo a presença de um pequeno grupo de pesquisadores altamente produtivos que sustentam a continuidade e o aprofundamento da área.

A linha sólida do gráfico representa os valores observados no conjunto de dados, enquanto a linha tracejada corresponde aos valores esperados através

da distribuição teórica proposta por Lotka. A proximidade geral entre essas curvas indica boa aderência do *corpus* ao modelo, embora pequenas diferenças possam sugerir particularidades do campo, como o caráter interdisciplinar ou o crescimento recente da área, fatores que podem influenciar a dispersão autoral.

Os resultados demonstram que a produtividade dos autores segue de maneira consistente a Lei de Lotka, evidenciando um padrão no qual muitos pesquisadores contribuem de forma pontual, enquanto poucos especialistas altamente produtivos desempenham um papel relevante na consolidação do conhecimento científico, reforçando a caracterização do campo como altamente concentrado, no qual poucos autores são responsáveis por uma parcela significativa da produção científica, enquanto a maioria contribui de forma episódica e/ou pontual.

A Tabela 8 apresenta os principais autores segundo indicadores de impacto bibliométrico, combinando métricas de produtividade, influência e longevidade científica no corpus analisado.

Tabela 8 - Métricas de impacto local dos autores

Autor	índice h	índice g	m_índice	TC	NP	PY_start
CHUTE CHRISTOPHER G.	6	6	0,353	338	6	2009
TAO CUI	6	6	0,462	135	6	2013
CUI LICONG	4	4	0,333	115	4	2014
DENNY JOSHUA C.	4	4	0,308	204	4	2013
KLANN JEFFREY G.	4	4	0,333	105	4	2014
MURPHY SHAWN N.	4	5	0,333	106	5	2014
MUSEN MARK A.	4	4	0,286	234	4	2012
PATHAK JYOTISHMAN	4	4	0,235	172	4	2009
XU HUA	4	4	0,222	229	4	2008
ZHANG GUO-QIANG	4	4	0,308	161	4	2013

Fonte: dados da pesquisa (2025).

A tabela apresenta o impacto local dos autores que permite identificar quem exerce maior influência no *corpus*, com base em indicadores bibliométricos utilizados para medir produtividade e impacto científico.

Observa-se que *Christopher G. Chute* e *Tao Cui* se destacam como os autores de maior relevância local, ambos com índice h = 6, índice g = 6 e 6 publicações, embora apresentem diferenças em citações totais (338 para *Chute* e 135 para *Cui*) e no m-índice, indicador ajustado pelo tempo de carreira, no qual

*Cui* apresenta valor superior (0,462) em comparação a *Chute* (0,353), sugerindo um impacto proporcionalmente mais rápido, considerando o início da sua produção científica em 2013.

Autores como *Cui Licong*, *Joshua C. Denny*, *Jeffrey G. Klann*, *Shawn N. Murphy*, *Mark A. Musen*, *Jyotishman Pathak*, *Xu Hua* e *Guo-Qiang Zhang*, todos com índice  $h = 4$ , também demonstram influência significativa dentro do *corpus*. Entre eles, destaca-se *Xu Hua*, com o maior número de citações totais (229), seguido de *Musen* (234) e *Denny* (204), indicando visibilidade de suas contribuições ao campo científico. O  $m$ -índice desses autores varia entre 0,222 e 0,333, refletindo uma progressão constante de impacto ao longo de suas trajetórias científicas.

O indicador APP\_start mostra que esses pesquisadores iniciaram sua produção científica entre 2008 e 2014, o que revela uma concentração temporal relativamente recente e alinhada à expansão da pesquisa no campo científico.

A convergência desses autores em instituições de excelência, como *Stanford*, *Harvard* e *Mayo Clinic*, reforça a relevância de suas contribuições para o desenvolvimento conceitual e aplicado da comunidade científica.

Portanto, a tabela evidencia que o impacto local é dominado por autores que articulam a produtividade com elevado número de citações e a inserção institucional, compondo um núcleo intelectual influente no *corpus* analisado, revelando uma estrutura de liderança científica concentrada, sustentada por autores com produtividade moderada, alto impacto relativo e forte inserção temática ao longo do período analisado.

#### 7.4 AFILIAÇÕES

A análise das afiliações institucionais constitui um elemento importante nos estudos bibliométricos, pois permite identificar os centros de pesquisa que desempenham um papel relevante na produção científica. Ao examinar as instituições às quais os autores estão vinculados, torna-se possível reconhecer os pólos de excelência, as áreas disciplinares ativas e as concentrações geográficas de conhecimento que influenciam o desenvolvimento do campo científico.

Essa abordagem também evidencia o nível de especialização das instituições, a força de suas agendas de pesquisa e possíveis redes de colaboração institucional. Deste modo, a análise das afiliações oferece uma visão aprofundada das estruturas institucionais que sustentam e impulsionam a produção científica, contribuindo para compreender a dinâmica organizacional e territorial da pesquisa científica no cenário internacional.

A Tabela 9 apresenta as principais afiliações institucionais associadas aos autores mais produtivos do corpus analisado, evidenciando a distribuição institucional da produção científica sobre ontologia no período estudado.

Tabela 9 - Top 10: afiliações mais relevantes

Afiliações	Artigos
STANFORD UNIV STANFORD CTR BIOMED INFORMAT RES STANFORD CA 94305 EUA	6
HARVARD UNIV SCH MED BOSTON MA EUA	3
MAYO CLIN DEPT HLTH SCI RES ROCHESTER MN EUA	3
PARTNERS HEALTHCARE BOSTON MA EUA	3
DEPARTAMENTO DE RÁDIO DA UNIVERSIDADE DE STANFORD, STANFORD, CA 94305, EUA	3
UNIV GRANADA DEPT BIBLIOTECA E DOCUMENTAÇÃO E-18071 GRANADA ESPANHA	3
UNIV TEXAS AUSTIN SCH INFORMAT AUSTIN TX EUA	3
UNIV TEXAS HLTH SCI CTR HOUSTON SCH BIOMED INFORMAT HOUSTON TX EUA	3
UNIV ZURIQUE DEPTO. DE GEOG. ZURIQUE SUÍÇA	3
FUNDAÇÃO BRITÂNICA DO CORAÇÃO CENTRO DE CIÊNCIA DE DADOS LONDRES INGLATERRA	2

Fonte: dados da pesquisa (2025).

As afiliações mais relevantes revelam as instituições que possuem maior representatividade para a produção científica, indicando os principais centros de pesquisa que lideram o desenvolvimento científico do campo.

Observa-se que a *Stanford University* se destaca de forma significativa, com 6 artigos, reforçando sua posição consolidada como um dos pólos globais de excelência em pesquisas científicas.

Em seguida, aparecem instituições norte-americanas de alto prestígio, cada uma com 3 publicações, entre elas a *Harvard Medical School*, o *Mayo Clinic Department of Health Sciences Research*, o *Partners Healthcare* e o *Departamento de Rádio da Universidade de Stanford*. Essas instituições, vinculadas ao domínio da saúde e da informática médica, evidenciam a forte

relação entre ontologias e aplicações clínicas, sinalizando que grande parte da produção relevante emerge de contextos biomédicos.

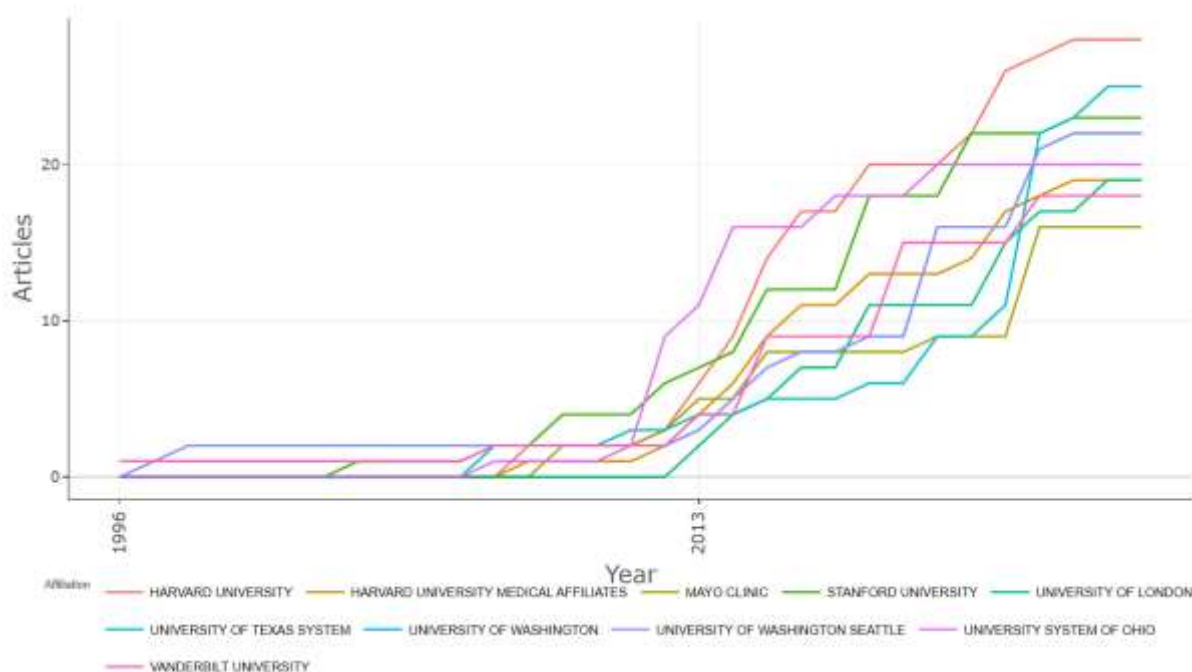
A presença da *Universidad de Granada* (Espanha) também com 3 artigos indica a inserção europeia na pesquisa em ontologias, especialmente nas áreas de organização da informação e gestão do conhecimento. Outras instituições dos Estados Unidos, como a *University of Texas at Austin*, a *University of Texas Health Science Center at Houston*, reforçam a dominância do país na produção científica, igualmente com 3 publicações cada.

Além disso, instituições de outros países, como a Universidade de Zurique (Suíça) e a *British Heart Foundation Data Science Centre* (Reino Unido), aparecem com produção relevante, embora em menor número de artigos (2), indicando a expansão da pesquisa em ontologias para diferentes contextos científicos e geográficos.

Portanto, a tabela evidencia que as afiliações mais produtivas concentram-se predominantemente em instituições norte-americanas de referência internacional, particularmente na intersecção entre informática biomédica, Ciência da Informação e pesquisas clínicas. A presença de instituições europeias complementa o panorama internacional, revelando um cenário colaborativo e interdisciplinar na produção científica relacionada às ontologias.

O Gráfico 10 apresenta a evolução temporal da produção científica das principais instituições vinculadas ao tema analisado, considerando o número acumulado de artigos publicados ao longo dos anos.

Gráfico 10 - Evolução temporal da produção de afiliações



Fonte: dados da pesquisa (2025).

O gráfico apresenta a evolução temporal da produção científica das afiliações, permitindo visualizar como diferentes instituições passaram a integrar, consolidar ou ampliar a sua participação ao longo dos anos. Observa-se um crescimento progressivo e contínuo da produção científica institucional a partir dos anos 2000, com aceleração mais evidente após 2010, movimento que indica a expansão do interesse e da maturação do campo científico.

A *Harvard University* destaca-se como a afiliação com maior trajetória de crescimento, evidenciando a liderança e constância na produção científica. Outras instituições de grande impacto, como *Stanford University*, *University of Washington* e *University of Texas System*, também apresentam curvas ascendentes expressivas, revelando a consolidação de grupos de pesquisa e o fortalecimento de redes colaborativas. O *Mayo Clinic*, *Vanderbilt University* e o *University System of Ohio* acompanham essa tendência, ainda que com volumes de produção menores, mas demonstrando a crescente inserção no tema ao longo do tempo.

As instituições europeias, como a *University of London*, e as norte-americanas associadas aos sistemas universitários mais amplos também exibem evolução constante, embora com variações nas taxas de crescimento entre os períodos.

Portanto, os dados evidenciam que a produção científica em nível institucional não apenas aumenta de forma contínua, mas se diversifica entre diferentes universidades, o que sugere a ampliação da base de pesquisadores, a intensificação das colaborações e o fortalecimento da pesquisa internacional no campo científico, ilustrando a concentração da produção científica em um conjunto restrito de universidades e reforçando a assimetria institucional característica do desenvolvimento científico e da relevância dessas organizações na estrutura social do campo.

## 7.5 PAÍSES

A análise dos indicadores bibliométricos associados aos países permite compreender de maneira abrangente a distribuição geográfica da produção científica, seu impacto e os padrões de colaboração internacional no campo.

As tabelas apresentadas evidenciam, de um lado, o volume de publicações e a liderança de países como Estados Unidos, Reino Unido, Espanha, China e Brasil; de outro, revelam o impacto dessas produções científicas por meio do total de citações e da média de citações por artigo, destacando a predominância dos Estados Unidos e de nações europeias de alta tradição científica.

Complementarmente, as tabelas que tratam da autoria correspondente e dos modos de colaboração (SCP e MCP) permitem identificar diferentes *perfis* de inserção internacional: países como *EUA* e Brasil apresentam maior autonomia na produção científica, enquanto Reino Unido, China, Canadá e diversos países da Europa demonstram integração com redes colaborativas em âmbito global.

Dessa forma, o conjunto dos indicadores bibliométricos apresentam um panorama consistente da atividade científica mundial, evidenciando assimetrias

entre os principais pólos de pesquisa quanto às tendências emergentes que reforçam a internacionalização da ciência e o caráter interdisciplinar do campo.

A Tabela 10 apresenta a distribuição geográfica da produção científica por país, destacando tanto o volume de artigos quanto os padrões de colaboração internacional no corpus analisado.

Tabela 10 - Top 10: países do autor correspondente

Pais	Artigos	Artigos %	SCP	MCP	MCP %
EUA	128	31,5	107	21	16,4
ESPAÑA	33	8,1	25	8	24,2
BRASIL	29	7,1	27	2	6,9
REINO UNIDO	29	7,1	18	11	37,9
CHINA	22	5,4	15	7	31,8
ITÁLIA	18	4,4	12	6	33,3
AUSTRÁLIA	16	3,9	11	5	31,3
FRANÇA	12	3,0	8	4	33,3
HOLANDA	10	2,5	9	1	10,0
CANADÁ	9	2,2	5	4	44,4

Fonte: dados da pesquisa (2025).

Os países do autor correspondente permitem identificar a distribuição geográfica da liderança das publicações científicas, evidenciando quais países concentram a representatividade das pesquisas científicas.

Observa-se que os Estados Unidos ocupam posição de destaque, com 128 artigos (31,5%), dos quais 107 foram produzidos com autoria correspondente única (SCP) e 21 em colaboração internacional (MCP). Embora apresentem o maior volume, a proporção de publicações em colaboração internacional (16,4%) é relativamente moderada, sugerindo forte capacidade de produção autônoma.

A Espanha aparece em segundo lugar, com 33 artigos (8,1%), sendo 25 SCP e 8 MCP, o que resulta em uma proporção mais elevada de colaboração internacional (24,2%), indicando a inserção significativa em redes de pesquisa transnacionais.

O Brasil, com 29 artigos (7,1%), apresenta um perfil distinto: a grande maioria das publicações é de autoria correspondente única (27 SCP), e apenas

2 envolvem a colaboração internacional (6,9%), o que evidencia um padrão de produção predominantemente interna, ainda pouco articulada com redes globais.

O Reino Unido, também com 29 artigos (7,1%), destaca-se por uma elevada proporção de MCP (37,9%), caracterizando um dos níveis mais altos de colaboração internacional entre os países listados.

A China apresenta 22 artigos (5,4%), com 15 SCP e 7 MCP, refletindo um equilíbrio crescente entre produção independente e colaboração global (31,8%).

Já países europeus como Itália, França e Holanda apresentam níveis moderados de produção, mas com porcentagens expressivas de MCP (em torno de 33%), indicando participação ativa em redes colaborativas internacionais.

A Austrália e o Canadá, com 16 e 9 artigos, respectivamente, também se destacam pela alta porcentagem de colaboração internacional, 31,3% para a Austrália e 44,4% para o Canadá, sendo este último o país com a maior proporção relativa de MCP, apesar do menor volume de publicações.

Portanto, a análise de países mostra que, enquanto países como EUA e Brasil mantêm padrões mais autônomos de produção científica, a Europa apresenta a integração em redes internacionais, revelando diferentes estratégias e graus de inserção global no campo científico e indicando que a produção científica sobre ontologia é internacionalizada, com diferenças pontuais entre os países quanto ao equilíbrio entre a produção nacional e a cooperação transnacional.

A Tabela 11 apresenta a frequência absoluta de ocorrência dos países associados à produção científica analisada, evidenciando a concentração e a distribuição geográfica das afiliações ou contribuições no corpus.

Tabela 11 - Top 10: produção científica dos países

Pais	Frequência
EUA	463
Reino Unido	75
ESPAÑA	71
CHINA	67
BRASIL	58
ALEMANHA	36
AUSTRÁLIA	34
FRANÇA	32
ITÁLIA	31
HOLANDA	24

Fonte: dados da pesquisa (2025).

A produção científica dos países evidencia a distribuição da produção científica por país, indicando quais países concentram maior número de publicações.

Os Estados Unidos destacam-se como o principal produtor científico da área, com 463 publicações, o que confirma sua liderança histórica e sua consolidação como o maior pólo de pesquisa internacional.

Em seguida, o Reino Unido ocupa a segunda posição, com 75 publicações, refletindo a sua tradição em Ciência da Informação, modelagem conceitual e sistemas de informação.

A Espanha aparece como o terceiro país mais produtivo, com 71 publicações, demonstrando uma atuação expressiva sobretudo nas áreas de organização da informação e aplicações na área de saúde.

A China, com 67 publicações, apresenta crescimento significativo e consolidado, indicando sua inserção crescente na pesquisa global em tecnologia e engenharia do conhecimento.

O Brasil, com 58 publicações, também se destaca como um importante ator na produção científica latino-americana, evidenciando a expansão da pesquisa em ontologias e da Ciência da Informação no contexto brasileiro.

Na sequência, países europeus como Alemanha (36), Austrália (34), França (32), Itália (31) e Holanda (24) apresentam contribuições relevantes, ainda que em menor escala, compondo um bloco consistente de produção científica que reforça a dominância da Europa Ocidental na pesquisa

interdisciplinar de ontologias. Esse conjunto de países evidencia que a produção científica é relativamente distribuída, embora permaneça concentrada majoritariamente entre Estados Unidos e nações europeias.

Os dados revelam que o campo científico de estudo de ontologia possui forte índice de internacionalização, mas com assimetrias claras entre os países, refletindo diferenças em capacidade de investimento, tradição científica e consolidação institucional. Esses resultados corroboram com o papel dominante dos EUA e da Europa Ocidental, ao mesmo tempo em que indicam o crescimento de países emergentes como China e Brasil no cenário global.

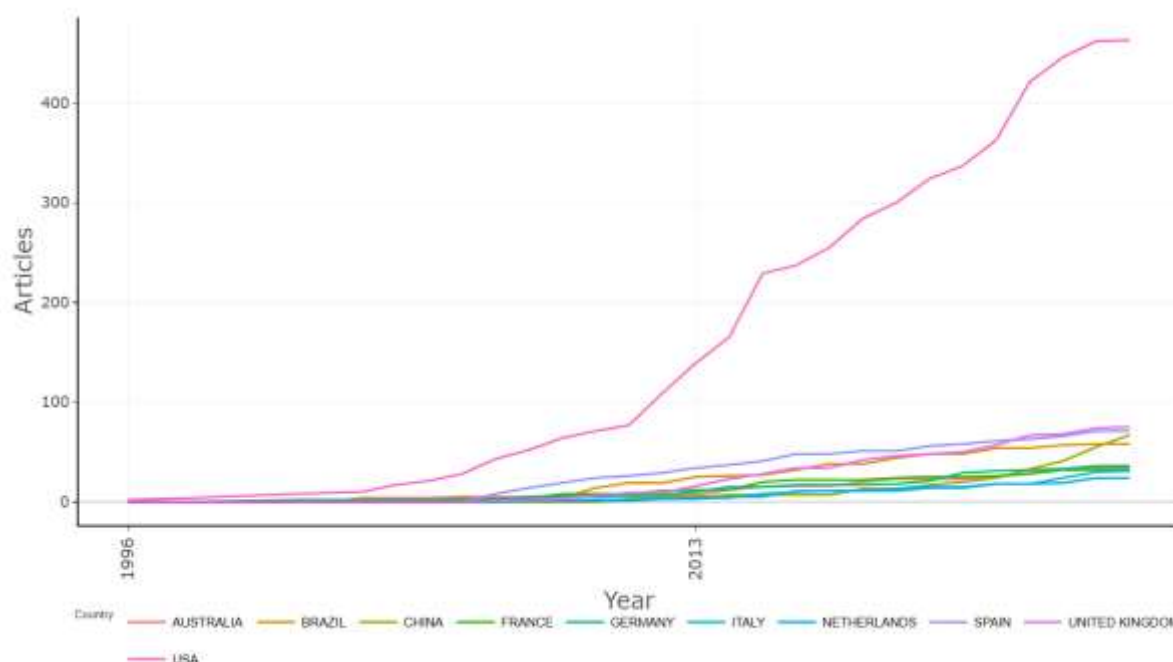
A análise da produtividade científica por país permite compreender como a produção de conhecimento está distribuída geograficamente e quais nações desempenham um papel dominante no desenvolvimento do campo científico.

Ao examinar o volume de publicações ao longo do tempo, torna-se possível identificar tendências de crescimento, assimetrias regionais e a emergência de novos pólos de pesquisa, bem como relacionar tais dinâmicas aos contextos institucionais, econômicos e científicos de cada país.

Assim, a análise da produtividade científica oferece uma visão abrangente da contribuição global para o avanço do campo científico, revelando os países que lideram historicamente a produção científica quanto os países que apresentam trajetórias emergentes e em expansão, revelando uma distribuição assimétrica, caracterizada por forte concentração em poucos países dominantes e uma participação relevante, ainda que menor, de países que atuam como pólos intermediários nas redes internacionais de pesquisa.

O Gráfico 11 ilustra a evolução temporal da produção científica por país, considerando o número acumulado de artigos publicados ao longo do período analisado.

Gráfico 11 - Evolução temporal da produção científica dos países



Fonte: dados da pesquisa (2025).

O Gráfico apresenta a evolução temporal da produção científica dos países, evidenciando diferenças significativas entre os países analisados.

Observa-se que os Estados Unidos mantêm, de forma destacada, a liderança em número de publicações, com um crescimento contínuo e acentuado a partir de meados dos anos 2000. Esse aumento se intensifica após 2010, culminando em um volume superior a 450 artigos nos anos mais recentes, o que confirma o papel dominante do país como principal pólo global de pesquisa científica no tema.

Os demais países apresentam trajetórias de crescimento moderadas, porém consistentes. O Reino Unido, a Espanha e a Holanda despontam como os países europeus com maior número de publicações, mostrando uma curva ascendente especialmente após 2010.

Em seguida, Alemanha, França e Itália apresentam crescimento contínuo, ainda que em menor escala, refletindo participação significativa na consolidação teórica e metodológica do campo.

Entre os países fora do eixo Europa–Estados Unidos, destaca-se o crescimento gradual de China, Austrália e Brasil, que, embora apresentem

produção quantitativamente menor, demonstram expansão constante ao longo das últimas duas décadas.

No caso específico do Brasil, o aumento expressivo ocorre após 2013, acompanhando a consolidação da pesquisa em Ciência da Informação no país.

Portanto, o gráfico revela que a produção científica em nível global tem crescido de forma contínua e geograficamente distribuída, mas ainda é altamente concentrada nos Estados Unidos, cuja trajetória se distancia amplamente dos demais países.

Observa-se que a tendência destaca o caráter internacional e interdisciplinar da área, ao mesmo tempo em que confirma a persistência de assimetrias entre os principais pólos de produção científica, evidenciando a concentração geográfica da produção científica, ao mesmo tempo em que revela a gradual ampliação da participação de países emergentes, refletindo processos de internacionalização e diversificação do campo ao longo do tempo.

A Tabela 12 apresenta a distribuição do impacto científico por país, considerando o total de citações (TC) e as citações médias por artigo, permitindo uma análise comparativa entre volume de impacto acumulado e intensidade média de reconhecimento das publicações.

Tabela 12 - Top 10: países mais citados

País	TC	Citações médias do artigo
EUA	4196	32,80
REINO UNIDO	673	23,20
HOLANDA	364	36,40
FRANÇA	322	26,80
ESPAÑA	259	7,80
ITÁLIA	251	13,90
SUÉCIA	223	27,90
CANADÁ	218	24,20
AUSTRÁLIA	177	11,10
ALEMANHA	169	18,80

Fonte: dados da pesquisa (2025).

A tabela apresenta os países mais citados e evidencia a distribuição geográfica da produção científica de maior impacto, permitindo identificar os países que desempenham um papel relevante no desenvolvimento e na disseminação das pesquisas científicas, evidenciando que o impacto científico não se distribui de forma proporcional ao volume de produção e reforçando a importância de análises que considerem simultaneamente os indicadores quantitativos e qualitativos da produção científica.

Os Estados Unidos (EUA) destacam-se de forma expressiva, com 4.196 citações e média de 32,80 citações por artigo, sinalizando sua posição de liderança global e sua tradição em pesquisas relacionadas à Ciência da Informação, informática médica e tecnologias semânticas.

Em seguida, o Reino Unido aparece como o segundo país mais citado, com 673 citações e média de 23,20 citações por artigo, confirmando sua relevância histórica em estudos sobre organização da informação, sistemas de informação e modelagem conceitual.

A Holanda, com 364 citações, apresenta a maior média de citações por artigo (36,40), o que sugere que, embora produza um volume menor de publicações, seus trabalhos alcançam impacto internacional, especialmente nas áreas de ontologias formais e interoperabilidade semântica.

Países como França (322 citações) e Suécia (223 citações) também demonstram expressiva representatividade no campo, com médias de 26,80 e 27,90 citações por artigo, respectivamente, refletindo contribuições significativas em áreas como bioinformática, linguística computacional e sistemas de informação geográfica.

A Espanha (259 citações) e a Itália (251 citações) aparecem com níveis de impacto moderados, embora consistentes, destacando-se por suas tradições consolidadas em Ciência da Informação e engenharia do conhecimento.

O Canadá e a Austrália, com 218 e 177 citações, respectivamente, também compõem o grupo de países influentes, confirmando uma participação ativa em pesquisas multidisciplinares que envolvem ontologias, tecnologias emergentes e saúde digital.

A Alemanha (169 citações), embora com menor volume total, mantém relevância devido à sua contribuição em modelagem formal, engenharia ontológica e aplicações em sistemas médicos.

Portanto, a distribuição das citações evidencia um campo concentrado em países do Norte Global, com destaque para os Estados Unidos e a Europa Ocidental, regiões que lideram tanto em volume quanto em impacto das publicações, configurando os principais pólos de produção e difusão científica de ontologias.

## 7.6 DOCUMENTOS










A análise dos documentos citados, tanto em nível local quanto global, constitui uma etapa importante nos estudos bibliométricos, pois permite compreender a estrutura de influência que orienta o desenvolvimento do campo.

Enquanto as citações globais revelam os trabalhos de maior reconhecimento internacional, evidenciando contribuições disseminadas e consolidadas na literatura científica, as citações locais destacam os documentos que exercem impacto direto no *corpus* analisado, sustentando de modo imediato a base teórica e metodológica da produção científica.

Deste modo, a análise desses indicadores possibilita identificar não apenas os textos seminalmente reconhecidos no cenário internacional, mas também aqueles que desempenham um papel relevante na articulação interna do campo, oferecendo uma visão abrangente da dinâmica de circulação, apropriação e influência das publicações relacionadas às ontologias na Ciência da Informação.

A Tabela 13 apresenta os artigos mais citados do corpus, destacando indicadores de impacto absoluto e relativo, como total de citações, taxa média de citações por ano e citações normalizadas, permitindo uma análise comparativa entre obras publicadas em diferentes períodos.

Tabela 13 - Top 10: documentos mais citados globalmente

Resumo	Papel	DOI	Total de citações	TC por ano	TC normalizado
	NICKERSON RC, 2013, EUR J INFORM SYST	<a href="https://doi.org/10.1057/ejis.2012.26">10.1057/ejis.2012.26</a>	738	56,77	15,01
	MUSEN MA, 1996, JAM MED INFORM ASSN	<a href="https://doi.org/10.1136/jamia.1996.97084511">10.1136/jamia.1996.97084511</a>	228	7,60	2,94
	SCOTT SV, 2014, MIS QUART	<a href="https://doi.org/10.25300/MISQ/2014/38.3.11">10.25300/MISQ/2014/38.3.11</a>	194	16,17	6,69
	COECKELBERGH M, 2010, ÉTICA EM TECNOLOGIA DA INFORMAÇÃO	<a href="https://doi.org/10.1007/s10676-010-9235-5">10.1007/s10676-010-9235-5</a>	194	12,13	9,73
	SMITH B, 2001, INT J GEOGR INF SCI	<a href="https://doi.org/10.1080/13658810110061199">10.1080/13658810110061199</a>	165	6,60	1,74
	WORBOYS M, 2005, INT J GEOGR INF SCI	<a href="https://doi.org/10.1080/13658810412331280167">10.1080/13658810412331280167</a>	160	7,62	2,46
	WAGNER G, 2022, J INF TECHNOL-UK	<a href="https://doi.org/10.1177/02683962211048201">10.1177/02683962211048201</a>	136	34,00	9,07
	RODRÍGUEZ MA, 2004, INT J GEOGR INF SCI	<a href="https://doi.org/10.1080/13658810310001629592">10.1080/13658810310001629592</a>	134	6,09	1,72
	CHEN ES, 2008, J AM MED INFORM ASSN	<a href="https://doi.org/10.1197/jamia.M2401">10.1197/jamia.M2401</a>	133	7,39	6,05
	MUSEN MA, 2012, J AM MED INFORM ASSN	<a href="https://doi.org/10.1136/amiajnl-2011-000523">10.1136/amiajnl-2011-000523</a>	131	9,36	4,82

Fonte: dados da pesquisa (2025).

Os documentos mais citados globalmente apresentam os trabalhos de maior impacto internacional, evidenciando aqueles que alcançaram ampla repercussão na comunidade científica e contribuíram de forma decisiva para o avanço teórico e metodológico da área.

O estudo de *Nickerson et al.* (2013), publicado no *European Journal of Information Systems*, ocupa a primeira posição, com expressivas 738 citações e um índice de 56,77 citações por ano, além de apresentar o mais elevado valor de TC normalizado (15,01). Esses indicadores demonstram sua influência significativa na área de sistemas de informação, consolidando-o como um dos trabalhos mais reconhecidos no *corpus* de referência.

Em seguida, destaca-se o artigo de *Musen* (1996), publicado no *Journal of the American Medical Informatics Association*, com 228 citações, representando um marco pioneiro na interface entre ontologias e informática médica.

A publicação de *Scott* (2014) na *MIS Quarterly* também figura entre os trabalhos de maior impacto, com 194 citações, assim como o estudo de *Coeckelbergh* (2010), que igualmente registra 194 citações, refletindo debates éticos relacionados às tecnologias da informação.

Outros documentos com forte repercussão internacional incluem os artigos de *Smith* (2001), *Worboys* (2005) e *Rodríguez* (2004), todos oriundos da

área de *Geographic Information Science*, demonstrando o papel das ontologias na modelagem espacial e no desenvolvimento de sistemas de informação.

O trabalho de Wagner (2022), apesar de recente, apresenta um elevado índice de TC por ano (34,00) e TC normalizado (9,07), apontando para sua rápida ascensão e relevância emergente na área de tecnologia da informação.











Também se destacam os estudos de *Chen* (2008) e *Musen* (2012), ambos vinculados à informática médica, cujas contribuições reforçam a maturidade e a consolidação do uso de ontologias em contextos clínicos e biomédicos.

A presença recorrente de autores associados à medicina e às ciências da informação geográfica indica que essas áreas constituem dois dos principais eixos de difusão global das ontologias no cenário científico internacional.

Portanto, os documentos mais citados globalmente evidenciam um campo interdisciplinar e consolidado, influenciado por pesquisas em sistemas de informação, ética e tecnologias emergentes, informática médica e ciências geoespaciais, revelando a amplitude e a relevância internacional do tema, revelando a coexistência de obras clássicas consolidadas e contribuições recentes de alto impacto, reforçando a dinâmica evolutiva do campo e a importância de considerar métricas normalizadas para comparações mais equilibradas entre períodos distintos de publicação.

A Tabela 14 apresenta os artigos mais citados localmente no corpus, articulando indicadores de impacto local e global e permitindo avaliar a relevância relativa das obras para o conjunto específico da pesquisa, em contraste com sua visibilidade internacional.

Tabela 14 - Top 10: documentos locais mais citados

Resumo	Papel	DOI	Ano	Citações locais	Citações globais	Relação LC/GC (%)	Citações locais normalizadas	Citações globais normalizadas
	TAO C, 2013, J AM MED INFORM ASSN	<a href="https://doi.org/10.1136/amiajn-2012-001326">10.1136/amiajn-2012-001326</a>	2013	4	42	9,52	6,40	0,85
	BEYNON-DAVIES P, 2016, INFORM ORGAN-UK	<a href="https://doi.org/10.10165.informorg.2016.04.001">10.10165.informorg.2016.04.001</a>	2016	4	16	25,00	5,43	0,76
	PEDRAZA-JIMENEZ R, 2007, PROF INFORM	<a href="https://doi.org/10.3145/epi.2007.nov.04">10.3145/epi.2007.nov.04</a>	2007	3	11	27,27	5,25	0,65
	HYUN S, 2009, J AM MED INFORM ASSN	<a href="https://doi.org/10.1197/jamia.M2821">10.1197/jamia.M2821</a>	2009	3	15	20,00	5,25	1,22
	MUSEN MA, 2012, J AM MED INFORM ASSN	<a href="https://doi.org/10.1135/amiajn-2011-000523">10.1135/amiajn-2011-000523</a>	2012	3	131	2,29	10,80	4,82
	WEI WQ, 2013, J AM MED INFORM ASSN	<a href="https://doi.org/10.1136/amiajn-2012-001431">10.1136/amiajn-2012-001431</a>	2013	3	70	4,29	4,80	1,42
	KAHN CE, 2015, J AM MED INFORM ASSN	<a href="https://doi.org/10.1093/jamia/ocv020">10.1093/jamia/ocv020</a>	2015	3	8	37,50	9,75	0,32
	DARAIO C, 2016, SCIENTOMETRICS-a	<a href="https://doi.org/10.1007/s11192-015-1814-0">10.1007/s11192-015-1814-0</a>	2016	3	32	9,38	4,07	1,51
	ERIKSSON O, 2018, J INF TECHNOL-UK	<a href="https://doi.org/10.1057/s41265-018-0053-2">10.1057/s41265-018-0053-2</a>	2018	3	12	25,00	12,60	0,67
	KUHN W, 2001, INT J GEOGR INF SCI	<a href="https://doi.org/10.1080/13658810110061180">10.1080/13658810110061180</a>	2001	2	119	1,68	2,00	1,26

Fonte: dados da pesquisa (2025).

Os documentos locais mais citados permitem identificar os artigos que exercem maior influência interna no conjunto documental analisado, revelando aqueles que contribuíram de maneira significativa para a consolidação do campo, evidenciando a complementaridade entre métricas locais e globais, reforçando a importância da análise contextualizada do impacto científico.

Observa-se que o estudo de *Tao* (2013), publicado no *Journal of the American Medical Informatics Association*, figura entre os documentos mais citados localmente, acumulando quatro citações internas e 42 citações globais, com uma relação LC/GC de 9,52%, o que indica impacto consistente tanto no *corpus* quanto na literatura da área biomédica.

O trabalho de *Beynon-Davies* (2016) também se destaca, apresentando quatro citações locais e 16 globais, posicionando-se como uma contribuição relevante para a compreensão das estruturas informacionais e organizacionais.

A presença do estudo de *Pedraza-Jiménez* (2007) com três citações locais reforça o papel das pesquisas dedicadas à gestão da informação e à organização do conhecimento na sustentação teórica do campo.

Outros documentos influentes incluem os trabalhos de *Hyun* (2009), *Musen* (2012) e *Wei* (2013), todos publicados na área de Informática Médica, demonstrando a *interface* aplicada entre ontologias e sistemas de saúde.

Em particular, o artigo de *Musen* (2012) obtém 131 citações globais, revelando amplo reconhecimento internacional enquanto a presença de *Kahn* (2015) e *Daraio* (2016) entre os documentos com três citações locais indica diversificação temática, abrangendo aplicações em saúde e métricas de produção científica.

O trabalho de *Eriksson* (2018) destaca-se por apresentar a maior pontuação de citações locais normalizadas (12,60), evidenciando impacto proporcionalmente elevado dentro do *corpus* analisado.

Por fim, o estudo clássico de *Kuhn* (2001) demonstra relevância global (119 citações), ainda que apresente menor incidência de citações locais, sugerindo que sua influência se dá sobretudo em perspectivas amplas, relacionadas à ciência da informação.

Portanto, a distribuição dos documentos mais citados confirma a predominância de estudos situados na *interface* entre ontologias, informática médica e gestão da informação, revelando um campo interdisciplinar cuja base empírica e teórica se apoia em aplicações no domínio biomédico e na modelagem de sistemas informacionais complexos.

## 7.7 REFERÊNCIAS CITADAS

A análise das referências citadas constitui um procedimento importante nos estudos bibliométricos, pois permite identificar os trabalhos que exercem maior influência intelectual no *corpus* analisado.

Ao examinar as obras que recebem o maior número de citações dentro do conjunto documental, torna-se possível compreender quais autores, teorias e contribuições metodológicas fundamentam a produção científica do campo.

Essa abordagem evidencia não apenas os textos clássicos e seminalmente reconhecidos, mas também as correntes de pensamento que estruturam a evolução conceitual e aplicada das ontologias na Ciência da Informação.

Desse modo, a apresentação das referências locais mais citadas oferece uma visão aprofundada da base intelectual que sustenta as pesquisas científicas

e auxilia na identificação dos pilares teóricos que orientam o desenvolvimento do campo.

A Tabela 15 apresenta as referências mais citadas no corpus analisado, evidenciando os trabalhos que exercem maior influência na estrutura intelectual do campo de pesquisa sobre ontologias.

Tabela 15 - Top 10: referências locais mais citadas

Referências citadas	Citações
GRUBER TR, 1993, AQUISIÇÃO DE CONHECIMENTO, V5, P199, DOI 10.1006/KNAC.1993.1008	38
BERNERS-LEE T, 2001, SCI AM, V284, P34, DOI 10.1038/SCIENTIFICAAMERICAN0501-34	32
GRUBER TR, 1995, INT J HUM-COMPUT ST, V43, P907, DOI 10.1006/IJHC.1995.1081	19
BODENREIDER O, 2004, NUCLEIC ACIDS RES, V32, PD267, DOI 10.1093/NAR/GKH061	16
ARONSON AR, 2001, J AM MED INFORM ASSN, P17	14
GUARINO N, 1998, FR ART INT, V46, P3	14
USCHOLD M, 1996, KNOWL ENG REV, V11, P93, DOI 10.1017/S0269888900007797	13
ASHBURNER M, 2000, NAT GENET, V25, P25, DOI 10.1038/75556	12
LINDBERG DAB, 1993, METHOD INFORM MED, V32, P281, DOI 10.1055/S-0038-1634945	11
CIMINO JJ, 1998, METHOD INFORM MED, V37, P394	10

Fonte: dados da pesquisa (2025).

As referências locais mais citadas evidenciam os principais trabalhos que fundamentam a produção científica, revelando os autores e as publicações que constituem o núcleo intelectual do campo de ontologias no domínio da Ciência da Informação, revelando um núcleo teórico robusto, composto por autores e obras que estruturam conceitualmente o desenvolvimento e a evolução das pesquisas analisadas.

Observa-se que o autor *Thomas R. Gruber* aparece em posição de destaque, com três publicações entre as mais citadas, principalmente seu trabalho de 1993, com 38 citações, considerado seminal para a definição de ontologia no contexto da Ciência da Informação. Essa proeminência confirma o papel de Gruber na consolidação conceitual e metodológica do tema.

Outro autor identificado é *Tim Berners-Lee*, cujo artigo publicado em 2001 apresenta 32 citações e reforça a conexão entre ontologias e a evolução da *web*

*semântica*, destacando a importância das tecnologias voltadas à interoperabilidade e ao enriquecimento semântico dos dados na *web*.

A presença recorrente de autores como *Bodenreider* (2004), com 16 citações, e *Aronson* (2001), com 14 citações, evidencia a *interface* do campo com a área biomédica, refletindo a influência de referenciais como *UMLS* e vocabulários controlados na construção de ontologias aplicadas à saúde.

A lista inclui ainda contribuições fundamentais de *Nicola Guarino* (1998), reconhecido por sua abordagem ontológica formal; *Uschold* (1996), importante para a sistematização de metodologias de modelagem ontológica; e *Ashburner* (2000), associado ao *Gene Ontology*, um dos projetos mais influentes no uso de ontologias aplicadas às ciências biológicas.

Por fim, autores como *Lindberg* (1993) e *Cimino* (1998) reforçam a predominância de aplicações em sistemas médicos, confirmando o papel dessa área como uma das primeiras e mais consolidadas a adotar ontologias em larga escala.

Portanto, as referências mais citadas revelam que o campo científico se apóia em trabalhos seminais da engenharia ontológica, da *web semântica* e das aplicações biomédicas, delineando um núcleo teórico e metodológico que sustenta o campo científico.

## 7.8 PALAVRAS

A análise das palavras permite identificar os termos que estruturam o discurso científico e evidenciam os principais focos temáticos do campo.

Ao examinar a incidência dos termos presentes nos títulos, resumos e palavras-chave do conjunto documental, torna-se possível compreender como determinadas palavras se consolidam, articulam-se entre si e orientam a agenda de pesquisa ao longo do tempo.

Portanto, a apresentação das palavras mais recorrentes no *corpus* analisado oferece uma visão das tendências conceituais que caracterizam a produção científica.

A Tabela 16 apresenta a frequência das palavras mais recorrentes no corpus analisado, evidenciando os principais eixos conceituais que estruturam o campo de pesquisa.

Tabela 16 - Top 10: palavras mais frequentes

Palavras	Ocorrências
ontologia	103
ontologias	54
conhecimento	33
web semântica	33
Informação	32
sistemas	26
classificação	20
modelo	19
estrutura	18
processamento de linguagem natural	17

Fonte: dados da pesquisa (2025).

As palavras mais frequentes permitem identificar os conceitos predominantes na produção científica de ontologia, evidenciando a interdisciplinaridade e a evolução tecnológica das pesquisas em ontologia no contexto da Ciência da Informação e áreas correlatas.

O termo “ontologia” apresenta a maior frequência (103 ocorrências), confirmando sua posição como eixo estruturante do campo científico.

Em seguida, observa-se a presença de “ontologias” (54 ocorrências), indicando que o debate se articula tanto em torno da definição conceitual quanto de suas múltiplas aplicações e modelos.

Os termos “conhecimento” e “*web semântica*”, ambos com 33 ocorrências, evidenciam a relação entre ontologias e processos de representação, organização e interoperabilidade semântica.

A recorrência de “informação” (32) e “sistemas” (26) reforça a inserção das ontologias em ambientes informacionais e tecnológicos, destacando sua função mediadora entre dados, estruturas informacionais e sistemas computacionais.

As palavras de frequência intermediária, como “classificação” (20), “modelo” (19) e “estrutura” (18), revelam a importância atribuída à modelagem conceitual e aos mecanismos de organização da informação, tradicionais pilares teóricos da área.

Já o termo “processamento de linguagem natural” (17) indica a presença de discussões que relacionam ontologias a métodos computacionais, ampliando o escopo interdisciplinar do campo científico.

Portanto, a distribuição das ocorrências demonstra que a literatura se concentra em temas que articulam ontologias, representação do conhecimento, *web semântica*, sistemas de informação e tecnologias linguísticas, configurando um panorama coerente com a evolução das pesquisas internacionais na área.

A Figura 15 apresenta uma nuvem de palavras construída a partir das palavras-chave e termos mais recorrentes na produção científica analisada, oferecendo uma visão da estrutura conceitual do campo.

Figura 15 - Nuvem de palavras



Fonte: dados da pesquisa (2025).

A nuvem de palavras evidencia a distribuição de frequência dos termos presentes no corpus analisado e permite identificar os conceitos mais

recorrentes no debate científico sobre ontologia no domínio da Ciência da Informação.

Observa-se que o termo '*ontology*' se destaca com maior dimensão tipográfica, indicando sua centralidade temática e confirmando sua posição como núcleo conceitual do campo.

Termos como '*ontologies*', '*semantic web*', '*information*', '*knowledge*' e '*systems*' também aparecem com elevado grau de proeminência, revelando alta co-ocorrência e sugerindo que a literatura concentra-se em discussões relativas à representação do conhecimento, web semântica, sistemas de informação e processos de organização e gestão da informação. Esses conceitos delineiam os eixos estruturantes que orientam as investigações sobre ontologias.

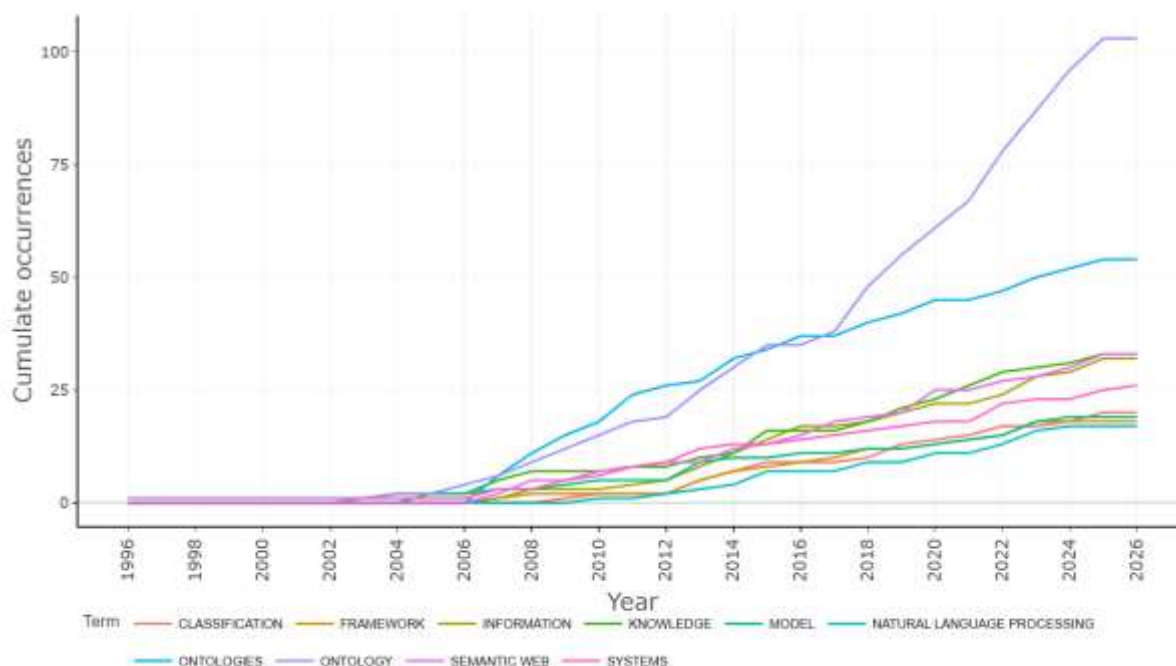
Além desses, a presença de termos de frequência intermediária como '*classification*', '*terminology*', '*metadata*', '*framework*', '*model*', '*integration*', '*linked data*' e '*natural language processing*' indica a existência de subtemáticas que, embora menos expressivas, contribuem para a compreensão multidimensional do campo, enquanto a ocorrência de expressões como '*electronic health records*' e '*informatics*' aponta para aplicações práticas, voltadas à saúde.

Portanto, a nuvem de palavras confirma a predominância de três eixos analíticos na produção científica examinada: fundamentos teórico-conceituais sobre ontologias; interfaces tecnológicas, sobretudo aquelas relacionadas à web semântica e à interoperabilidade; e processos informacionais, incluindo organização do conhecimento, terminologia e metadados.

Logo, a nuvem de palavras permite compreender que o campo se estrutura na interseção entre ontologias, tecnologias semânticas e gestão da informação, refletindo a maturidade e a diversidade temática da área.

O Gráfico 12 apresenta a evolução temporal cumulativa da frequência de termos-chave ao longo do período analisado, permitindo observar a dinâmica de consolidação e crescimento dos principais conceitos do campo.

Gráfico 12 - Indicador de frequência das palavras



Fonte: dados da pesquisa (2025).

O gráfico ilustra a frequência acumulada das palavras-chave mais recorrentes no corpus de pesquisa, demonstrando a evolução e a proeminência dos principais conceitos ao longo do tempo, evidenciando a maturação do campo e a diversidade temática ao longo do tempo.

O eixo vertical representa o número acumulado de ocorrências de cada termo, enquanto o eixo horizontal indica o período de análise, de 1996 a 2026. Este tipo de gráfico é importante para identificar os termos que não apenas são frequentes, mas que também mantiveram uma relevância crescente e consolidada no campo científico.

A análise do gráfico evidencia a centralidade indiscutível dos termos 'ONTOLOGY' (em roxo) e 'ONTOLOGIES' (em azul claro), que apresentam as curvas de crescimento mais acentuadas e atingem os maiores valores de frequência acumulada. O crescimento exponencial de ambos, especialmente a partir de 2006, confirma que eles constituem o núcleo temático da pesquisa. A trajetória ascendente e contínua desses dois termos indica um interesse consolidado e crescente no conceito de ontologia como objeto de estudo principal.

Um segundo grupo de termos, embora com crescimento menos expressivo, demonstra uma importância consistente para o campo científico. Destaca-se o termo 'NATURAL LANGUAGE PROCESSING', cuja curva ascendente, principalmente após 2014, o posiciona como um dos conceitos aplicados mais relevantes, sugerindo uma conexão entre ontologias e o processamento de linguagem natural.

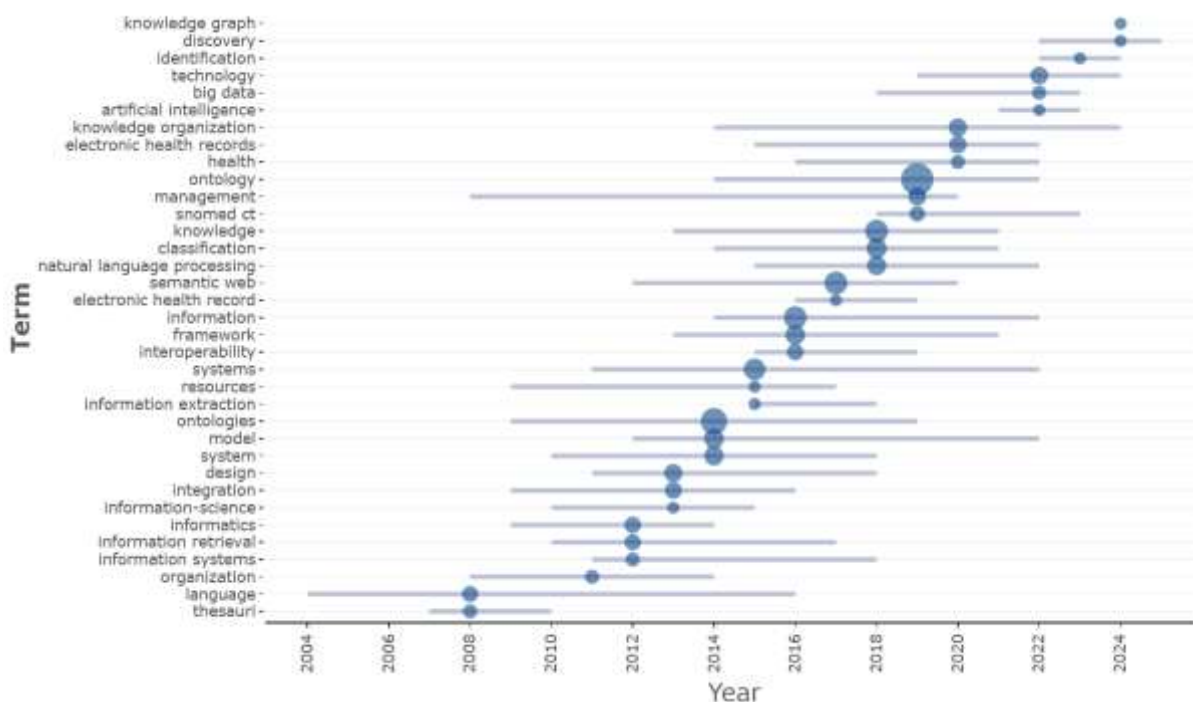
Em um patamar similar, termos como '*SEMANTIC WEB*', '*INFORMATION*' e '*KNOWLEDGE*' também exibem um crescimento constante, formando a base conceitual e tecnológica sobre a qual os estudos de ontologias se desenvolvem.

Por fim, um terceiro grupo de palavras, como '*SYSTEMS*', '*FRAMEWORK*', '*MODEL*' e '*CLASSIFICATION*', apresenta um crescimento moderado, mas estável. A presença desses termos indica uma preocupação persistente com a aplicação prática e a estruturação do conhecimento, refletindo no uso de ontologias como modelos e frameworks para a construção de sistemas de informação.

Portanto, o gráfico demonstra a consolidação do campo de pesquisa que, partindo de conceitos basilares da Ciência da Informação, se solidificou em torno do tema 'ontologia' e evoluiu para incorporar abordagens aplicadas.

O Gráfico 13 apresenta a linha do tempo dos tópicos emergentes e consolidados no campo analisado, evidenciando, para cada termo, o período de ocorrência, a intensidade e a duração de sua presença na literatura científica, evidenciando a transição gradual do campo que articula fundamentos consolidados com novas agendas temáticas orientadas à inovação e à interdisciplinaridade.

Gráfico 13 - Indicador de tópicos em alta



Fonte: dados da pesquisa (2025).

O indicador de tópicos em alta apresenta a análise da evolução temporal dos termos-chave, gerada a partir do *corpus* de pesquisa sobre ontologias na Ciência da Informação.

Este tipo de visualização, conhecida como mapa de tendências (*trend topics*), ilustra a dinâmica e a longevidade dos principais conceitos abordados na literatura.

No gráfico, o eixo horizontal representa o tempo (anos), o eixo vertical lista os termos, o tamanho de cada círculo indica a frequência de um termo em um determinado ano, e as linhas horizontais marcam o período de vida de cada conceito.

A primeira fase, anterior a 2012, é marcada por termos seminais da Ciência da Informação, como '*thesauri*', '*language*', '*organization*', '*information systems*' e '*information retrieval*'. Isso sugere que as discussões iniciais sobre ontologias emergiram a partir das preocupações clássicas com a organização e recuperação da informação.

A segunda fase, entre 2012 e 2018, representa a consolidação do tema. O termo '*ontology*' atinge seu pico de frequência por volta de 2014-2015, indicando o auge de sua popularidade como objeto de estudo. Nesse período, surgem e se fortalecem termos técnicos e aplicados, como '*semantic web*', '*natural language processing*', '*interoperability*' e '*information extraction*', além do domínio de aplicação '*health*'. Essa fase demonstra um amadurecimento do campo científico, com um deslocamento de conceitos seminais para a aplicação prática e a solução de problemas de integração e processamento de dados.

A terceira e mais recente fase, a partir de 2018, aponta para as frentes de pesquisa atuais e emergentes. Observa-se uma ascensão de termos como '*knowledge graph*', '*big data*', '*artificial intelligence*' e '*electronic health records*'. Essa tendência indica uma integração das ontologias com os paradigmas da ciência de dados e da inteligência artificial, onde os grafos de conhecimento se tornam uma aplicação proeminente.

A consolidação do domínio da saúde também é evidente, com a especificidade do termo '*electronic health records*'. Portanto, o gráfico ilustra a transição de um campo focado na organização conceitual para um domínio dinâmico e integrado às tecnologias de dados em larga escala.

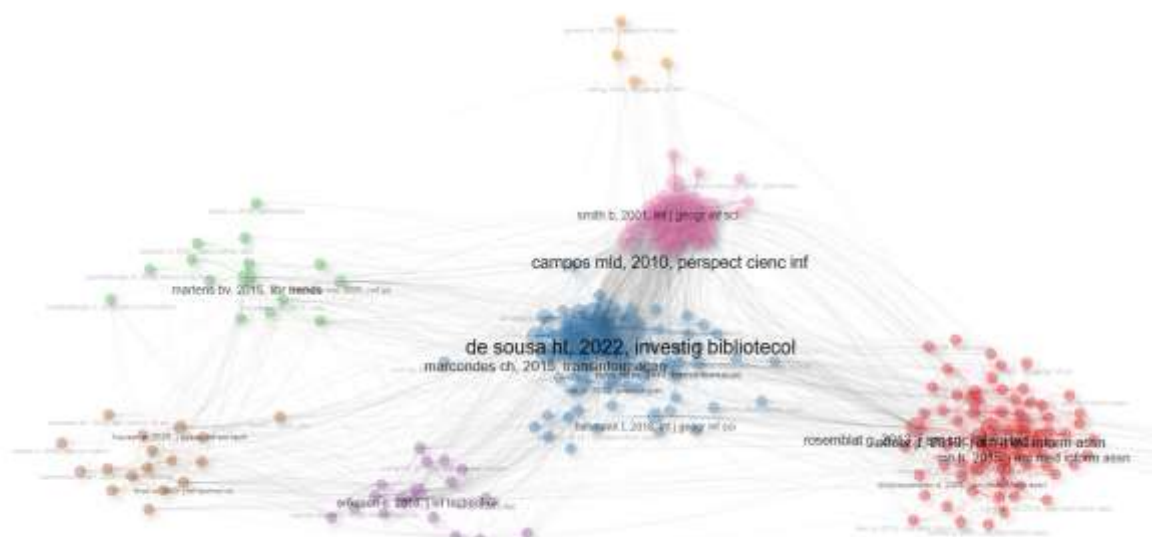
## 7.9 AGRUPAMENTO POR ACOPLAMENTO BIBLIOGRÁFICO

O agrupamento por acoplamento bibliográfico permite identificar os principais núcleos de produção científica, formados por documentos que compartilham bases teóricas comuns. Esse método evidencia as frentes de pesquisa emergentes, uma vez que se fundamenta nas referências citadas pelos documentos atuais, revelando a similaridade temática e a proximidade intelectual entre os estudos.

Deste modo, os *clusters* formados representam comunidades científicas com potencial para indicar novas tendências de pesquisa de ontologia em Ciência da Informação.

A Figura 16 apresenta a rede de acoplamento bibliográfico da produção científica analisada, evidenciando a organização estrutural do campo a partir do compartilhamento de referências entre os documentos.

Figura 16 - Agrupamento por acoplamento bibliográfico



Fonte: dados da pesquisa (2025).

A Figura apresenta uma rede de agrupamento por acoplamento bibliográfico, evidenciando como as publicações científicas se organizam em *clusters* temáticos a partir do compartilhamento de referências bibliográficas comuns. Nesse tipo de análise, quanto mais dois trabalhos citam as mesmas fontes, maior é sua proximidade intelectual e, conseqüentemente, sua ligação na rede, permitindo compreender a estrutura cognitiva e a dinâmica de articulação do campo, evidenciando tanto sua coesão interna quanto sua diversidade temática.

Observa-se a formação de múltiplos *clusters*, representados por diferentes cores, cada qual articulado em torno de autores (nós maiores). Esses nós funcionam como núcleos temáticos ou centros de pesquisa, destacando-se por alta frequência de conexões, indicando uma atuação na construção teórica do campo científico.

No *cluster* azul, localizado na região central inferior, há predominância de estudos associados à Ciência da Informação e ontologias aplicadas à

organização do conhecimento, com destaque para trabalhos de *Sousa (2022)* e *Marcondes (2015)*, que exercem um papel estruturante na rede. Esse núcleo representa uma base consolidada com ampla articulação interdisciplinar.

O *cluster* rosa, na área superior central, reúne estudos voltados às discussões epistemológicas e fundamentos teóricos da informação, com ênfase em autores como *Smith (2001)* e *Campos (2010)*, que exploram os aspectos conceituais das ontologias, suas aplicações e limitações.

O *cluster* verde, à esquerda, corresponde a trabalhos voltados à ética, métricas da informação e tendências emergentes no campo, destacando autores como *Martens (2015)* e *Bawden (2020)*, indicando um eixo teórico-metodológico com foco em políticas, implicações éticas e impacto da informação.

Já o *cluster* vermelho, à direita, mostra integração entre pesquisas de saúde, medicina e sistemas de informação biomédica, com trabalhos influentes como *Rosemlat (2013)* e *Mo (2015)*. Esse grupo representa um subcampo aplicado, com produção ligada à interoperabilidade semântica e ontologias médicas.

Por fim, o *cluster* marrom/lilás, na parte inferior esquerda, concentra estudos emergentes e recentes, com menor densidade, mas ligados a abordagens tecnológicas e inovação, consolidando-se como fronteiras em desenvolvimento no campo científico.

Portanto, a rede demonstra uma estrutura intelectual diversificada, madura e com alto grau de especialização, onde os *clusters* revelam subcomunidades temáticas bem definidas, articuladas por meio de bases teóricas compartilhadas. Ela evidencia a existência de núcleos consolidados, temas emergentes e zonas de intermediação que favorecem a conectividade entre os diferentes campos de aplicação das ontologias na Ciência da Informação.

## 7.10 ESTRUTURA CONCEITUAL

A análise da estrutura conceitual possibilita identificar os principais eixos teóricos que sustentam o domínio estudado, bem como a formação de *clusters*



Observa-se que os termos '*ontology*' e '*ontologies*' ocupam posições centrais na rede, funcionando como nós de alta centralidade. Eles estabelecem conexões com múltiplos *clusters* temáticos, indicando que o conceito de ontologia constitui o núcleo conceitual do campo científico, articulando diferentes abordagens, aplicações e perspectivas teóricas.

O *cluster* vermelho representa a aplicação de ontologias em sistemas de informação, padronização terminológica e interoperabilidade que aponta para uma vertente técnico-aplicada, voltada à representação do conhecimento.

O *cluster* azul representa pesquisas sobre ontologias enquanto mecanismos de apoio à organização do conhecimento, interoperabilidade informacional e estruturação semântica de dados na *web*.

O *cluster* verde reflete estudos teóricos e metodológicos voltados à modelagem do conhecimento, extração semântica, processamento de linguagem natural e padrões conceituais.

O *cluster* roxo indica discussões voltadas à integração de sistemas de informação, gestão do conhecimento e construção de infraestruturas informacionais baseadas em ontologias.

Portanto, '*Ontology*' é o nó central, articulando todos os *clusters*, representando o conceito nuclear do campo científico.

*Knowledge* aparece como ponte semântica entre os campos teóricos (representação) e práticos (organização e sistemas).

*Semantic web* funciona como eixo de convergência entre Ciência da Informação, Ciência da Computação e *Web Semântica*.

A presença de termos de saúde, ciência, tecnologia e gestão revela o caráter interdisciplinar do campo científico.

A rede de co-ocorrência evidencia que os estudos sobre ontologia na Ciência da Informação se estruturam em torno de duas grandes dimensões: dimensão teórico-conceitual com foco na representação do conhecimento, teoria da informação e modelagem conceitual; e a dimensão aplicada e tecnológica que faz uso de ontologias em sistemas de informação, interoperabilidade, organização de dados, *web semântica* e saúde.

Essa configuração demonstra a evolução do campo científico rumo a abordagens interdisciplinares, tecnológicas e orientadas à gestão do

conhecimento, consolidando a ontologia como instrumento estratégico para a organização, representação e circulação de informações digitais em ambientes complexos.

## 7.11 ESTRUTURA INTELECTUAL

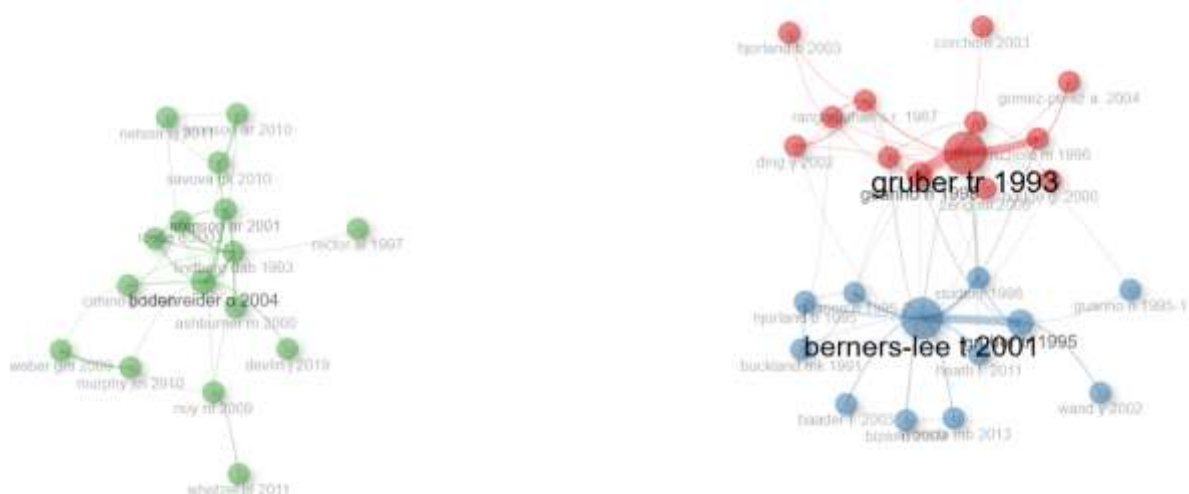
A análise da estrutura intelectual permite identificar os fundamentos epistemológicos que sustentam o desenvolvimento da investigação científica sobre ontologia no domínio da Ciência da Informação.

Por meio da rede de co-citação, foi possível mapear os autores mais influentes, além de delimitar as correntes teóricas que estruturam o campo científico.

Os agrupamentos identificados revelam diferentes escolas intelectuais, representadas por abordagens voltadas à modelagem de conhecimento, interoperabilidade semântica, construção de ontologias e fundamentos teóricos da representação da informação.

A Figura 18 apresenta a rede de co-citação de autores que representa a estrutura intelectual do campo investigado a partir das referências citadas conjuntamente nos documentos analisados.

Figura 18 - Rede de co-citação de documentos



Fonte: dados da pesquisa (2025).

Nessa rede, os nós correspondem a autores e obras fundamentais, cujo tamanho indica o grau de centralidade ou frequência de co-citação, enquanto as arestas representam a força das relações intelectuais estabelecidas entre eles.

Observam-se agrupamentos bem definidos, diferenciados por cores, que indicam escolas teóricas ou tradições conceituais distintas, permitindo identificar os principais referenciais teóricos, as convergências intelectuais e a organização cognitiva que sustenta a evolução científica da área analisada.

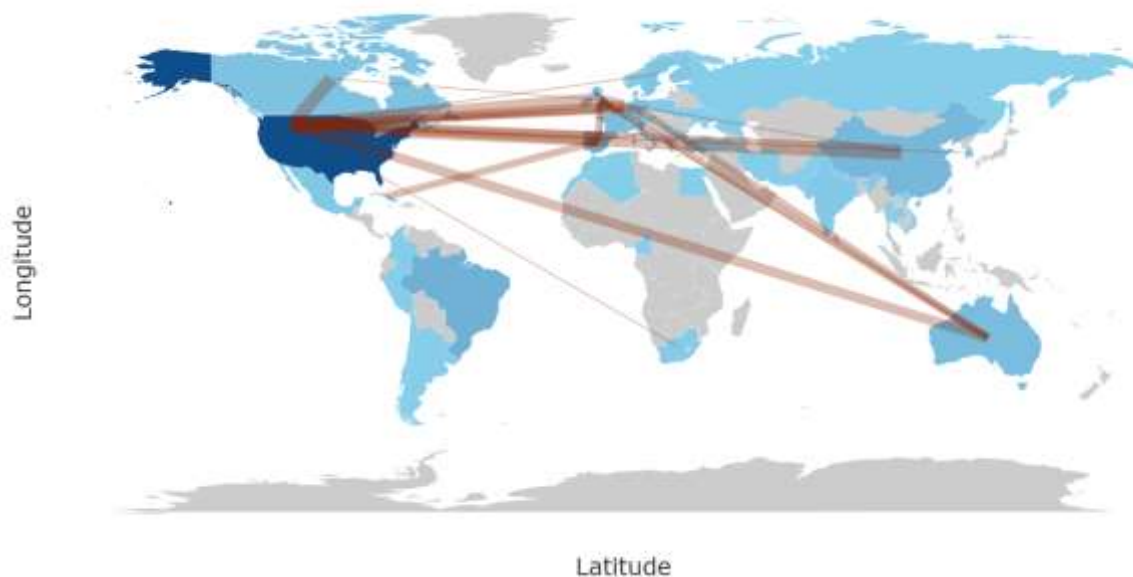
A rede de co-citação evidencia três correntes intelectuais no campo científico: uma vertente fundacional, que estabelece os princípios teóricos e metodológicos da ontologia (*cluster Gruber*); uma vertente tecnológica, voltada à integração das ontologias na *Web Semântica* e nos sistemas digitais (*cluster Berners-Lee*); e uma vertente aplicada, que incorpora as ontologias como instrumento de representação e organização da informação, especialmente em contextos biomédicos (*cluster Bodenreider*). Essa estrutura revela a evolução epistemológica do campo científico, desde a conceituação teórica até as aplicações práticas e tecnológicas.

## 7.12 ESTRUTURA SOCIAL

A análise da estrutura social constitui uma das dimensões da bibliometria que permite compreender como o conhecimento científico é produzido de forma colaborativa por meio das interações entre autores, instituições e países. Essa abordagem revela os padrões de cooperação científica que caracterizam um determinado domínio de pesquisa, possibilitando identificar os principais atores, suas relações e o grau de internacionalização do campo científico.

A Figura 19 apresenta o mapa mundial de colaboração científica entre países, evidenciando a dimensão geográfica e relacional da produção científica analisada e permitindo compreender a estrutura social e geográfica da colaboração internacional, ressaltando os principais hubs de pesquisa e as dinâmicas globais que sustentam a circulação do conhecimento no campo investigado.

Figura 19 - Mapa de colaboração científica global



Fonte: dados da pesquisa (2025).

A Figura apresenta um mapa mundial que representa a colaboração científica em âmbito internacional.

O mapa utiliza uma codificação cromática para representar a intensidade da produção por país, enquanto as linhas de conexão indicam relações de coautoria internacional, cuja espessura reflete a força ou frequência dessas colaborações, utilizando diferentes graduações de tonalidade azul para indicar a intensidade de participação de cada país na rede de colaboração internacional, sendo que as tonalidades mais escuras correspondem aos países com maior volume de produção científica e maior envolvimento em parcerias internacionais, enquanto os tons mais claros indicam participação menos expressiva. Os países que não apresentam registros de colaboração ou de publicações associadas aparecem representados em cinza.

Observam-se linhas interligando os países, as quais representam as relações de coautoria e cooperação científica. A espessura dessas conexões varia proporcionalmente à frequência ou intensidade das colaborações, de modo que linhas mais espessas indicam vínculos fortes ou recorrentes.

Nota-se uma centralidade dos Estados Unidos, que estabelecem múltiplas parcerias com países europeus, asiáticos e da Oceania. No contexto europeu,

destacam-se países como Reino Unido, Alemanha, França, Espanha e Itália, os quais também apresentam expressivas conexões entre si e com outras regiões.

A Austrália figura como um ator relevante no hemisfério sul, estabelecendo conexões significativas com os Estados Unidos, Europa e Ásia. Em contraste, alguns países da América Latina, África e partes da Ásia apresentam menor inserção na rede científica internacional.

Portanto, a rede de colaboração científica em âmbito internacional evidencia um padrão de colaboração predominantemente concentrado entre países da América do Norte, Europa Ocidental e Oceania, os quais se configuram como pólos estratégicos de produção e circulação do conhecimento científico.

O mapa revela a existência de centros de pesquisa na rede científica, especialmente os Estados Unidos e determinadas nações europeias, que desempenham papel de influência na articulação das redes globais de coautoria.

A Tabela 17 apresenta o fluxo de colaboração internacional a partir do Brasil, evidenciando as conexões estabelecidas com outros países no contexto do corpus analisado.

Tabela 17 - Colaboração Brasil - Austria - Colombia – Portugal

From	To	Frequency
BRAZIL	All	All
BRAZIL	AUSTRIA	14.1264760996
BRAZIL	COLOMBIA	-73.0811458241
BRAZIL	PORTUGAL	-8.50104361268

Fonte: dados da pesquisa (2025).

No que se refere às relações de colaboração científica brasileira no cenário internacional, observa-se a existência de vínculos com três países: Áustria, Colômbia e Portugal. Essas relações evidenciam a inserção internacional do Brasil na rede global de coautoria, ainda que de forma pontual e menos expressiva quando comparada às principais potências científicas observadas no mapa global de colaboração científica.

A colaboração mais significativa ocorre com a Áustria, com frequência aproximada de 14,13, indicando uma interação relevante, possivelmente

associada a projetos específicos ou a parcerias institucionais consolidadas entre universidades e centros de pesquisa dos dois países. Esse resultado sugere um alinhamento temático entre os pesquisadores brasileiros e austríacos no domínio da Ciência da Informação, com ênfase no tema ontologia.

A conexão com Portugal apresenta frequência negativa (-8,50), possivelmente refletindo limitações na mensuração de dados ou baixa intensidade de coautorias registradas. Ainda assim, essa relação evidencia a proximidade linguística e histórica entre os países, que costuma favorecer redes de colaboração científica, embora, neste conjunto de dados, a interação apareça de forma reduzida.

Da mesma forma, a relação com a Colômbia apresenta frequência negativa (-73,08), indicando uma colaboração incipiente ou pouco expressiva no recorte temático. Essa baixa interação sugere que, apesar de ambos integrarem o contexto latino-americano, ainda não há uma articulação científica consolidada entre Brasil e Colômbia nesse campo do conhecimento científico.

Portanto, os resultados revelam que, embora o Brasil participe da rede internacional de colaboração científica, sua atuação ainda é limitada e concentrada em poucos países, com destaque para a Áustria. A baixa intensidade nas relações com países latino-americanos indica um potencial a ser explorado para o fortalecimento de redes regionais de pesquisa, sobretudo, considerando a relevância crescente da ontologia e da Ciência da Informação no contexto da internacionalização da ciência, evidenciando que, embora o Brasil esteja inserido em redes internacionais de pesquisa, essas interações apresentam comportamentos heterogêneos, refletindo diferentes graus de consolidação e intensidade das parcerias científicas.

## 8 CONSIDERAÇÕES FINAIS

Ao longo da dissertação, propôs-se uma reflexão crítica sobre o papel da comunicação científica como vetor para a democratização do acesso à informação e para a promoção da equidade na disseminação do conhecimento científico. Logo, a análise das inter-relações nesse processo permitiu evidenciar como determinadas estratégias e políticas científicas estão reconfigurando o cenário da ciência.

O estudo demonstrou que o acesso aberto se consolida como uma estratégia fundamental para a democratização do conhecimento científico, ao remover as barreiras financeiras que historicamente marginalizam pesquisadores e instituições com menor capacidade de investimento.

Em paralelo, a internacionalização da ciência foi discutida como uma política de difusão que, ao fomentar a colaboração científica e a visibilidade global, acelera as descobertas e o avanço do conhecimento no campo científico.

Argumentou-se, ainda, que a articulação dessas estratégias culmina no movimento pela ciência aberta e inclusiva. Esta não é apenas uma prática, mas uma mudança cultural que transforma a gestão da informação, promovendo maior transparência, produtividade e reprodutibilidade da pesquisa científica.

Ao valorizar o protagonismo da pesquisa e estimular a adoção de boas práticas científicas, a ciência se torna mais confiável, acessível e responsiva aos desafios sociais existentes no campo científico.

Por último, cabe destacar a contribuição da análise bibliométrica da produção científica de ontologia em Ciência da Informação no cenário internacional, além de apresentar as limitações da pesquisa e as recomendações para investigações futuras.

### 8.1 CONTRIBUIÇÃO DO ESTUDO À CIÊNCIA DA INFORMAÇÃO

A dissertação apresentada contribui com a área da Ciência da Informação em três níveis principais: empírico, metodológico e conceitual.

Primeiramente, em um nível empírico, o estudo oferece um mapeamento detalhado e abrangente da produção científica sobre ontologia, revelando a

estrutura intelectual, as redes de colaboração e as principais frentes de pesquisa deste domínio no âmbito internacional. Ao fornecer este panorama, o trabalho serve como um guia para novos pesquisadores e um diagnóstico do estado da arte do campo.

Em segundo lugar, em um nível metodológico, a pesquisa detalha um percurso rigoroso para a condução de análises bibliométricas, desde a coleta e tratamento de dados da *Web of Science* até a aplicação de técnicas de mapeamento científico, servindo como um modelo replicável para futuros estudos sobre a dinâmica de outros domínios do conhecimento.

Por fim, em um nível conceitual, a investigação transcende a análise descritiva ao conectar os padrões de publicação científica com debates pertinentes da área, como o papel do acesso aberto na democratização do conhecimento e a dinâmica da comunicação científica, reforçando a importância dos estudos métricos como ferramenta para a própria ciência se auto-observar e compreender enquanto campo científico.

## 8.2 LIMITAÇÕES DA PESQUISA BIBLIOMÉTRICA

É importante reconhecer que os resultados apresentados na dissertação, embora obtidos por meio de uma metodologia rigorosa, estão sujeitos a limitações inerentes aos estudos bibliométricos e às escolhas metodológicas realizadas. Deste modo, a consciência dessas limitações é fundamental para uma interpretação crítica e contextualizada dos indicadores bibliométricos apresentados.

Assim, as principais limitações da pesquisa bibliométrica são:

- Viés da Base de Dados: a escolha da *Web of Science* (*WoS*) como única fonte de dados, apesar de justificada por sua qualidade e rigor, introduz um viés de cobertura. A *WoS* é conhecida por sua seletividade, com um foco predominante em periódicos de alto impacto e publicações em língua inglesa. Consequentemente, a produção científica publicada em outros idiomas, em periódicos de escopo regional ou em formatos alternativos pode estar sub-

representada, e os indicadores de colaboração internacional e impacto de citação devem ser interpretados à luz desse contexto.

- **Qualidade e Completude dos Metadados:** embora a *WoS* seja reconhecida pela alta qualidade de seus metadados, a análise ainda está sujeita a inconsistências ou dados incompletos. Variações na padronização de nomes de autores ou afiliações institucionais, por exemplo, podem afetar a precisão das redes de coautoria. Da mesma forma, a sensibilidade dos cálculos da taxa de crescimento anual aos anos de início e término da coleta de dados pode fazer com que picos ou quedas em anos específicos distorçam a tendência geral observada.
- **Restrição ao Acesso Aberto:** o foco exclusivo em publicações de acesso aberto, embora relevante para o objetivo deste estudo sobre a democratização do conhecimento científico, representa um recorte específico do universo total da produção científica. Isso significa que trabalhos relevantes sobre o tema, publicados sob modelos de assinatura, não foram incluídos na análise, de modo que os resultados retratam o panorama do movimento de acesso aberto, e não a totalidade da literatura sobre o assunto.
- **Especificação do Tipo Documental:** a análise foi concentrada no tipo documental "artigo", por ser o principal veículo de comunicação científica revisada por pares. No entanto, essa escolha exclui outras formas de produção de conhecimento, como livros, capítulos de livros e anais de eventos, que também são relevantes para a disseminação do conhecimento em diversas áreas científicas. Portanto, a visão apresentada reflete a comunicação científica em periódicos, mas não abrange a totalidade dos formatos de publicação.
- **Atraso na Acumulação de Citações (*Citation Lag*):** a análise do impacto dos documentos, medida pelo número de citações, é sensível ao tempo. Logo, as publicações mais recentes, especialmente aquelas com menos de dois anos, não tiveram tempo suficiente para acumular um número significativo de

citações. Portanto, a média de citações por documento pode subestimar o impacto dos trabalhos recentes no *corpus* de referência, e as tendências de citação devem ser analisadas com cautela para os anos finais do período investigado.

### 8.3 RECOMENDAÇÕES PARA INVESTIGAÇÕES FUTURAS

Como sugestão para futuras investigações, recomenda-se a ampliação do escopo dos estudos métricos por meio de análises comparativas, que é reconhecida por ser uma abordagem que permite transcender a descrição de um único campo para gerar insights explicativos sobre as dinâmicas científicas. A bibliometria demonstra que o poder do método científico é potencializado ao se contrastar diferentes conjuntos de dados, revelando como os padrões de crescimento, colaboração e impacto são influenciados por contextos distintos.

Nesse sentido, abrem-se diversas frentes de pesquisa futura, tais como:

- Realizar um estudo contrastivo entre a produção científica de ontologias, indexadas na *Web of Science*, representando o cenário global, e, a produção registrada na base BRAPCI, focada no Brasil. Tal análise permitiria identificar as especificidades da pesquisa nacional, suas redes de colaboração e particularidades temáticas em relação às tendências internacionais.
- Investigar como o conceito de "ontologia" é tratado em diferentes áreas do conhecimento, comparando, por exemplo, o *corpus* da Ciência da Informação com o da Ciência da Computação. Isso poderia revelar se as bases intelectuais, autores e trabalhos mais citados, as frentes de pesquisa e os padrões de autoria diferem significativamente entre os campos científicos.
- Conduzir um estudo metodológico que compare os resultados obtidos para a mesma consulta de busca nas bases de dados *Web of Science (WoS)*, *Library, Information Science & Technology Abstracts (LISTA)* e a Base de Dados em Ciência da Informação

(BRAPCI). Essa investigação evidenciaria os vieses de cobertura de cada plataforma e seu impacto sobre os indicadores bibliométricos, contribuindo para uma compreensão abrangente das ferramentas de pesquisa bibliométrica.

## REFERÊNCIAS

- ALMEIDA, C. H. M.; CAMPOS, M. L. A. Ontologia e web semântica: o espaço da pesquisa em ciência da informação. **Ponto de Acesso**, v. 2, n. 1, 2008.
- ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, v. 32, n. 3, 2003.
- ALMEIDA, M. B.; MENDONÇA, F. M.; FREITAS, E. C. A. Interfaces entre ontologias e conceitos seminais da Ciência da Informação: em busca de avanços na organização do conhecimento. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 14., 2013, Florianópolis. **Anais** [...]. Florianópolis: ANCIB, 2013.
- ALMEIDA, M. B.; OLIVEIRA, V. N. P.; COELHO, K. C. Estudo exploratório sobre ontologias aplicadas a modelos de sistemas de informação: perspectivas de pesquisa em ciência da informação. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 15, n. 30, 2010.
- ALMEIDA, M. B. **Ontologia em ciência da informação: teoria e método**. Curitiba: CRV, 2020. (Coleção Representação do Conhecimento em Ciência da Informação - volume 1).
- ALMEIDA, M. B. Teorias ontológicas para modelagem. **Fronteiras da Representação do Conhecimento**, v. 1, n., 2021.
- ALMEIDA, M. B. Uma abordagem integrada sobre ontologias: ciência da informação, ciência da computação e filosofia. **Perspectivas em Ciência da Informação**, v. 19, n. 3, 2014.
- ARAÚJO, A. C.; FERNANDES, L. Internacionalização e pós-graduação: a política de editais da Capes (2005-2018). **Avaliação** - Revista da Avaliação da Educação Superior, Campinas, SP, v. 26, n. 02, maio-ago. 2021.
- ARAÚJO, C. A. A. A ciência da informação como uma ciência social. **Ciência da Informação**, v. 32, n. 3, 2003.
- ARAÚJO, C. A. A.; ARAUJO, C. A. A. Bibliometria: evolução histórica e questões atuais. **Em Questão**, v. 12, n. 1, 2006.
- ARAUJO, K. M.; BARCELLOS, C.; SACRAMENTO, I. Consolidação de uma política brasileira de ciência aberta e melhores práticas de avaliação da ciência: um caminho necessário para a concretização de uma ciência inclusiva e global. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 18, n. 4, 2024.

ARIA, M.; CUCCURULLO, C.. Bibliometrix: an R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, Amsterdam, v. 11, n. 4, p. 959–975, 2017.

BORKO, H. Information science: what is it? **American Documentation**, Washington, v. 19, n. 1, p. 3–5, jan. 1968.

BUDAPEST OPEN ACCESS INITIATIVE (BOAI). **Read the Budapest Open Access Initiative**. Budapest, 2002. Disponível em: <https://www.budapestopenaccessinitiative.org/read>. Acesso em: 22 dez. 2025

BÖRNER, K. ; CHEN, C. ; BOYACK, K. W. Visualizing knowledge domains. **Annual Review of Information Science and Technology**, Hoboken, v. 37, n. 1, p. 179–255, 2003.

BORTOLI, A. F.; OLIVEIRA, M. R.; PEZARICO, G. Estratégia nacional de ciência, tecnologia e inovação (2016-2019): uma análise das políticas de CTI no Brasil. **Colóquio - Revista do Desenvolvimento Regional**, FACCAT, v. 17, n. 4. 2020.

BRAGA, G. M. Relações bibliométricas entre a frente de pesquisa (research front) e revisões da literatura. **Ciência da Informação**, Rio de Janeiro, v. 2, n. 1, p. 9-73, 1973.

BRASCHER, M.; CARLAN, E. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, J.; BRASCHER, M. (Org.). **Passeios pelos bosques da informação**: estudos sobre representação e organização da informação e do conhecimento. Brasília: IBICT, 2010. p. 147-176.

BUCKLAND, M. K. Information as thing. **Journal of the American Society for Information Science**, Hoboken, v. 42, n. 5, p. 351–360, 1991.

BUFREM, L. S.; PRATES, Y. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, v. 34, n. 2, 2005.

CAFE, L. M. A.; SANTOS, R. N. M.; BARROS, C. M. Os estudos de gruber e guarino sobre ontologias na ciência da informação e nas ciências da computação. **DataGramZero**, v. 16, n. 3, 2015.

CALLON, M.; COURTIAL, J. P. PENAN, H. **Cientometría: el estudio cuantitativo de la actividad científica**: de la bibliometría a la vigilancia tecnológica. Gijón: Ediciones Trea, 1993.

CARVALHO, S. B. R.; ARAÚJO, G. C. Gestão da internacionalização das instituições de ensino superior. **Avaliação - Revista da Avaliação da Educação Superior**, v. 25, n. 1. 2020.

CASTRO, F. F.; ARAKAKI, A. C. S.; SIMIONATO, A. O. C. Revisitando ontologia e metadados à luz dos ambientes informacionais digitais. **Perspectivas em Ciência da Informação**, v. 25, n. 4, 2020.

CEDRO, L. F. A.; GOTTSCHALG DUQUE, C. A disseminação da informação científica como garantia da legitimidade dos resultados de ensaios clínicos. RDBCI: **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 18, n. 00, p. e020010, 2020.

COBO, M. J.; LÓPEZ-HERRERA, A. G.; HERRERA-VIEDMA, Enrique; HERRERA, Francisco. 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**, Amsterdam, v. 5, n. 1, p. 146–166, 2011.

COLE, F. J.; EALES, N. B. The history of comparative anatomy: part I: a statistical analysis of the literature. **Science Progress**, London, v. 11, n. 41, p. 578–596, 1917.

COSTA, M. P.; LEITE, F. C. L. Fatores que exercem influência na comunicação dos dados de pesquisa: uma revisão sistematizada da literatura no campo da Ciência da Informação. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 19., 2018. Londrina. **Anais [...]**. Londrina: ANCIB, 2018.

COSTA, M. P.; LEITE, F. C. L. Princípios e recomendações basilares para a comunicação dos dados de pesquisa. **Em Questão**, v. 23, n. 1, 2017.

COSTA, R. S. **Estudo bibliométrico e de ontologia**: descrição de panoramas em nanocompósitos poliméricos com argila, grafeno, nanocelulose e energia nuclear. 2024. Tese (Doutorado em Tecnologia Nuclear - Materiais) - Instituto de Pesquisas Energéticas e Nucleares, Universidade de São Paulo, São Paulo, 2024. Acesso em: 2025-03-06.

CURTY, R. G.; DELBIANCO, N. R. As diferentes metrias dos estudos métricos da informação: evolução epistemológica, inter-relações e representações. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da Informação, v. 25, n., 2020.

Crane, D. **Faculdades invisíveis**: Difusão de conhecimento em comunidades científicas. Chicago: University of Chicago Press, 1972.

EMYGDIO, J. L.; ALMEIDA, M. B.; TEIXEIRA, L. M. D. Ensaio sobre ontologia aplicada na recuperação da informação para a ciência da informação. **Ponto de Acesso**, v. 15, n. 3, 2021.

FALANGOLA, R. F.; SANTOS, E. S.; RODRIGUES, J.S. O papel das comunicações acadêmicas na construção e disseminação do conhecimento científico: uma revisão bibliográfica. **LATTICE** - Revista Científica Multidisciplinar, v. 2, n. 2, 2025.

FERREIRA, A. G. C. Bibliometria na avaliação de periódicos científicos. **DataGramZero**, v. 11, n. 3, 2010.

FIGUEIREDO, F. C.; ALMEIDA, F. G. Ontologias em ciência da informação: um estudo bibliométrico no Brasil. **Ciência da Informação**, v. 46, n. 1, 2017.

FIORAVANTE, K. E.; ROBAINA, I. M. M.; NABOZNY, A.. Internacionalização e produção do conhecimento científico: algumas reflexões a partir da atuação dos bolsistas de produtividade em pesquisa em Geografia do CNPq. **Ateliê Geográfico**, v. 17, n. 3, 2023.

GAAL, L. P. M.; MARTINS, M. S. Acesso aberto no contexto da pesquisa em ciência da informação. **Transinformação**, v. 34, n., 2022.

GHENO, E. M. *et al.* Impacto da internacionalização na visibilidade da produção científica do programa de pós-graduação em ciências biológicas: bioquímica/UFRGS (2007-2016). **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 25, n., 2020.

GLÄNZEL, Wolfgang. National characteristics in international scientific co-authorship relations. **Scientometrics**, Dordrecht, v. 51, n. 1, p. 69–115, 2001.

GLÄNZEL, Wolfgang. **Bibliometrics as a research field: a course on theory and application of bibliometric indicators**. Leuven: Course Handouts, 2003.

GLÄNZEL, Wolfgang; SCHUBERT, Andrés. Analyzing scientific networks through co-authorship. In: MOED, Henk F.; GLÄNZEL, Wolfgang; SCHMOCH, Ulrich (org.). Handbook of quantitative science and technology research. Dordrecht: **Kluwer Academic Publishers**, 2004. p. 257–279.

GRUBER, T.R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, p. 199-220, 1993.

GUARINO, N. **Formal Ontology in Information Systems**. 1998.

GUARINO, N.; GIARETTA, P. **Ontologies and Knowledge Bases: towards a terminological clarification**. 1995.

GUEDES, V. L. S. A bibliometria e a gestão da informação e do conhecimento científico e tecnológico: uma revisão da literatura. **Ponto de Acesso**, v. 6, n. 2, 2012.

GUEDES, V. L. S.; BORSCHIVER, S. Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento em sistemas de informação, de comunicação e de avaliação científica e tecnológica. In: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO (CINFORM), 6., 2005, Salvador. **Anais [...]**. Salvador: CINFORM, 2005. p. 1–18.

HEINZ, M.; MIRANDA, A. Ciência aberta: argumentos e desafios para sua legitimação científica. **Em Questão**, Porto Alegre, v. 30, e-135618, 2024.

HJØRLAND, B. Fundamentals of knowledge organization. **Knowledge Organization**, [S.l.], v. 30, n. 2, p. 87-111, 2003.

HJØRLAND, Birger. The concept of information. **Annual Review of Information Science and Technology**, Hoboken, v. 37, p. 343–411, 2003.

HULME, E. W. **Statistical bibliography in relation to the growth of modern civilization**: two lectures delivered in the University of Cambridge in May, 1922. London: Butler & Tanner, 1923.

LEÃO, B. L. F.; NOGUEIRA, F. A.; CASTRO, A. M. D. A. Internacionalização da pós-graduação no Brasil: o Programa CAPES-PRINT (2018–2022). **Revista da FAEEBA – Educação e Contemporaneidade**, Salvador, v. 33, n. 73, p. 91–107, jan./mar. 2024.

MACIAS-CHAPULA, C. A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciência da Informação**, v. 27, n. 2, 1998.

MARICATO, J. M.; MENDES, M. M. Conceitos, tendências e atores envolvidos na divulgação científica: considerações a partir das pesquisas apresentadas na intercom | concepts, tendencies and actors involved in scientific publicizing: considerations from researches presented at intercom. **Liinc em revista**, v. 11, n. 2, 2015.

MARIN NETO, A. **Ontologias na representação do conhecimento**: uma ferramenta semântica para a ciência da informação. 2018. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2018.

MARQUES, Nelson Luiz Reyes. Acesso Aberto: Democratização do Conhecimento e Impactos na Ciência e na Sociedade. **Revista Educar Mais**, v. 9, p. 1-4. 2025.

MEDEIROS, C. B.; SILVEIRA, L.; REZENDE, L.; ARAÚJO, P. C. de; FREITAS CAMPOS, F. de; SENA, P.; DRUCKER, D.; BARATA, G.; COSTA, M. M.; SIMIONATO ARAKAKI, A. C.; GARCIA DE SOUZA, M.; SHEARER, Kathleen; PACKER, A. L.; MELLO RODE, S. de; AMARO, B.; UMPIERRE, D.; TURBA DE PAULA, R.; CARVALHO SEGUNDO, W. R. de. **Promoção e implementação da ciência aberta no Brasil**: resposta ao texto “Ciência Aberta: uma visão desapaixonada”. Zenodo, 2025.

MEDEIROS, J. M. G.; VITORIANO, M. A. V. A evolução da bibliometria e sua interdisciplinaridade na produção científica brasileira. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 13, n. 3, 2015.

MENEZES, D. P.; GALVÃO, R.. **Promoção e implementação da ciência aberta no Brasil**: resposta ao texto “Ciência aberta: uma visão desapaixonada”. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Brasília, 2025.

MUCK, F. A. L.; CAREGNATO, S. E. A produção científica em acesso aberto da odontologia brasileira: uma análise da publicação e do impacto na coleção principal da *web of science*. **Transinformação**, v. 35, n., 2023.

MORAL-MUÑOZ, J. A.; LÓPEZ-HERRERA, A. G.; HERRERA-VIEDMA, E.; COBO, M. J. **Software tools for analyzing scientific mapping**: a review. *In*: GLÄNZEL, Wolfgang; MOED, Henk F.; SCHMOCH, Ulrich; THELWALL, M. (org.). Springer handbook of science and technology indicators. Cham: Springer, 2019. p. 159–185.

NORONHA, D. P.; MARICATO, J. M. Estudos métricos da informação: primeiras aproximações. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v., n. esp., 2008.

NOYONS, E.C.M., MOED, H.F. ; VAN RAAN, A.F.J. Integrating research performance analysis and science mapping. **Scientometrics**, Dordrecht, v. 46, n. 3, p. 591–604, 1999.

NOYONS, E.C.M., MOED, H.F. ; LUWEL, M. Combining mapping and citation analysis for evaluative bibliometric purposes: a bibliometric study. **Journal of the American Society for Information Science**, Hoboken, v. 50, n. 2, p. 115–131, 1999.

OTLET, P. **Traité de documentation**. Bruxelles: Éditions du Mundaneum, 1934.

PEREIRA, D.R; M. Os impactos da ciência aberta na divulgação científica. **Leitura: Teoria & Prática**, v. 40, n. 86, 2022.

PERSSON, O. DANELL, R.; WIBORG SCHNEIDER, J.. How to use Bibexcel for various types of bibliometric analysis. *In*: ÅSTRÖM, Fredrik; DANELL, Rickard; LARSEN, Birger; WIBORG SCHNEIDER, Jesper (org.). **Celebrating scholarly communication studies**: a Festschrift for Olle Persson at his 60th birthday. v. 5. Leuven: International Society for Scientometrics and Informetrics, 2009. p. 9–24.

PINTO, A. L. *et al.* Alguns métodos estatísticos voltados às unidades de informação. **Biblios** (Peru), v., n. 46, 2012.

PRICE, Derek J. de Solla. **Little science, big science**. New Cork : Columbia University Press, 1963.

PRICE, D.; GÜRSEY, S. Studies in scientometrics: I. Transience and continuance in scientific authorship. **Ciência da Informação**, Rio de Janeiro, v. 4, n. 1, p. 27–40, 1975.

PRITCHARD, A. Statistical bibliography or bibliometrics. **Journal of Documentation**, London, v. 25, p. 348–349, 1969.

RANGANATHAN, S. R. **Classification and international documentation**. The Hague: Fédération Internationale de Documentation, 1948.

R CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2016.

SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, M. D. P. B. **Metodologia de pesquisa**. 5. ed. Porto Alegre: Penso, 2013.

SANTOS, A. C. G. **Ciência aberta e gestão da informação científica institucional: modelo proposto para gestão de dados científicos na Universidade Federal Rural da Amazônia**. Tese (Doutorado em Ciência da Informação) - Universidade Fernando Pessoa, Porto, 2022.

SANTOS, M. T. **Estudo do processo de apropriação da ontologia pela Ciência da Informação no Brasil**. 2014. Dissertação (Mestrado em Ciência da Informação) - Centro de Artes e Comunicação, Universidade Federal de Pernambuco, Recife, 2014. Acesso em: 25.03.06.

SARACEVIC, T. A natureza interdisciplinar da ciência da informação. **Ciência da Informação**, Brasília, v. 24, n. 1, 1995.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**. Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996

SHINTAKU, M.; VIOLA, C. M. M.; COSTAL, M.; SCHIESSL, I. T.; MACEDO, D. J. Popularização da ciência e suas essências taxonômicas: vertentes possíveis. In: Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB, 24., 2024, Vitória. **Anais [...]**. Vitória: Universidade Federal do Espírito Santo, 2024.

SILVA, Talita Guimarães da; SILVA, Tarcisio Torres. Estratégias de internacionalização brasileiras: algumas lições do Ciência sem Fronteiras. **Arquivos Analíticos de Políticas Educativas**, v. 29, n. 129.

SMALL, H. Visualizing science by citation mapping. **Journal of the American Society for Information Science**, Hoboken, v. 50, n. 9, p. 799–813, 1999.

SMITH, B. **Ontology and Informations Systems**. 2004.

TRINCA, T. P.; ALBAGLI, S. Avaliação da pesquisa científica no âmbito das políticas nacionais de ciência aberta. **RDBCI - Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 21, e023021, 2023.

SOERGUEL, D. **Functions of a thesaurus, classification and ontological knowledge bases**. 1997.

SOWA, J. F. **Building, sharing and merging ontologies**. Tutorial. [S. 1.: s. n.], 1999.

SOUZA, C. L. P. **Mapeamento da produção científica de organização e representação da informação e do conhecimento de pós-graduação em ciência da informação das universidades públicas do sudeste brasileiro: 2017-2022**. 2023. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2023.

TOMOYOSE, K.; *et al.* Indicadores bibliométricos no estudo da ontologia. **Biblionline**, v. 14, n. 1, 2018.

USCHOLD, M; GRUNINGER, M. Ontologies: principles, methods an applications. **Knowledge Engineering Review**, v. 11, n. 2, 1996.

UNESCO - Organização das Nações Unidas para a Educação, a **Ciência e a Cultura**. Recomendação da UNESCO sobre Ciência Aberta. Disponível em: [https://unesdoc.unesco.org/ark:/48223/pf0000379949\\_por](https://unesdoc.unesco.org/ark:/48223/pf0000379949_por). Acesso em: 19 set. 2025.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, v. 31, n. 2, 2002.

VAN RAAN, A. F. J. Advances in bibliometric analysis: research performance assessment and science mapping. *In*: BLOCKMANS, Wim; E., L.; WEAIRE, D. (org.). **Bibliometrics: use and abuse in the review of research performance**. Wenner-Gren International Series, v. 87. London: Portland Press, 2014. p. 17–28.

VELHO, L. Conceitos de ciência e a política científica, tecnológica e de inovação. **Sociologias**, Porto Alegre, ano 13, n. 26, jan./abr. 2011, p. 128-153.

VICKERY, B. C. Ontologies. **Journal of Information Science**, London, v. 23, n. 4, p. 227-286, 1997.

WALTMAN, L.; VAN ECK, N. J.; NOYONS, E. C. M. A unified approach to mapping and clustering of bibliometric networks. **Journal of Informetrics, Amsterdam**, v. 4, n. 4, p. 629–635, 2010.

ZUPIC, I.; ČATER, T. Bibliometric methods in management and organization. **Organizational Research Methods**, Thousand Oaks, v. 18, n. 3, p. 429–472, 2015.

### APÊNDICE A - CRONOGRAMA DE PESQUISA

ATIVIDADES	PRAZOS (MESES)																													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Fundamentos em Ciência da Informação	X	X	X	X																										
Metodologia da Pesquisa em Ciência da Informação	X	X	X	X																										
Fundamentos em Organização da Informação						X	X	X	X	X																				
Tópicos Especiais em Organização da Informação: Arquitetura da Informação e Blockchain					X	X	X	X	X																					
Tópicos Especiais em Organização da Informação: Arquitetura da Informação e Big Data										X	X	X	X	X																
Pesquisa Orientada em Ciência da Informação										X	X	X	X	X																
Submissão de Artigo ONTOBRAS													X	X	X	X														
Estágio Docência														X	X	X	X	X												
Pesquisa Orientada em Ciência da Informação																				X	X	X	X	X						

Fonte: elaborado pela autora (2025).

ATIVIDADES	PRAZOS (MESES)																															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
Prorrogação de Prazo																											X	X	X	X	X	X
Trancamento de Matrícula: Tratamento Médico																											X	X	X	X	X	X
Relatório de Qualificação																										X	X	X				
Agendamento: Banca de Qualificação																													X	X		
Submissão de Artigo																													X	X		
Qualificação																														X		
Revisão, Ajustes e Formatação																													X	X	X	
Agendamento: Banca de Defesa																														X	X	
Defesa da Dissertação																															X	

Fonte: elaborado pela autora (2025).

## **ANEXO A - DECLARAÇÃO DE FINANCIAMENTO**

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (CAPES) — Código de Financiamento 001 — *This study was financed in part by the* Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - *Finance Code* 001.