



UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

ÉRIKA DEMACHKI AGUIAR

**COBERTURAS DE FONTES DE DADOS E CORRELAÇÕES ENTRE
INDICADORES ALTMÉTRICOS E CITAÇÕES EM PORTAIS DE PERIÓDICOS
BRASILEIROS DE ACESSO ABERTO**

Brasília
2022

ÉRIKA DEMACHKI AGUIAR

COBERTURAS DE FONTES DE DADOS E CORRELAÇÕES ENTRE INDICADORES
ALTMÉTRICOS E CITAÇÕES EM PORTAIS DE PERIÓDICOS BRASILEIROS DE
ACESSO ABERTO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação (PPGCINF), da Faculdade de Ciência da Informação (FCI), da Universidade de Brasília (UnB), como requisito para a obtenção do título de Mestra em Ciência da Informação.

Área de concentração: Gestão da Informação

Linha de pesquisa: Comunicação e Mediação da Informação

Orientador: Prof. Dr. João de Melo Maricato.

Brasília
2022

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

AA283c Aguiar, Erika Demachki
Coberturas de fontes de dados e correlações entre indicadores altmétricos e citações em portais de periódicos brasileiros de acesso aberto / Erika Demachki Aguiar; orientador João de Melo Maricato. -- Brasília, 2022. 171 p.

Dissertação(Mestrado em Ciência da Informação) -- Universidade de Brasília, 2022.

1. Comunicação Científica. 2. Almetria. 3. Indicadores bibliométricos. 4. Acesso Aberto. 5. Portais de Periódicos. I. Maricato, João de Melo, orient. II. Título.

FOLHA DE APROVAÇÃO

Título: “COBERTURAS DE FONTES DE DADOS E CORRELAÇÕES ENTRE INDICADORES ALTMÉTRICOS E CITAÇÕES EM PORTAIS DE PERIÓDICOS BRASILEIROS DE ACESSO ABERTO ”

Autor (a): ERIKA DEMACHKI AGUIAR

Área de concentração: Gestão, Organização e Comunicação da Informação e do Conhecimento

Linha de pesquisa: Produção, socialização e Usos da Informação e do Conhecimento

Dissertação submetida à Comissão Examinadora designada pelo Colegiado do Programa de Pós-graduação em Ciência da Informação da Faculdade em Ciência da Informação da Universidade de Brasília como requisito parcial para obtenção do título de **MESTRE** em Ciência da Informação.

Dissertação aprovada em: 12 de dezembro 2022.

Presidente (UnB/PPGCINF): Prof. Dr. João de Melo Maricato

Membro Externo (IBICT/UFRJ): Prof. Dr. Fabio Castro Gouveia

Membro Interno (UnB/PPGCINF): Prof. Dr. Fernando César Lima Leite

Suplente (UnB/PPGCINF): Prof. Dr. Dalton Lopes Martins

Em 15/12/2022.



Documento assinado eletronicamente por **Joao de Melo Maricato, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 15/12/2022, às 17:14, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Fernando César Lima Leite, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 16/12/2022, às 11:22, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Fabio Castro Gouveia, Usuário Externo**, em 20/12/2022, às 17:14, 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, informando o código verificador **9094096** e o código CRC **F69DF4B2**.

Referência: Processo nº 23106.130297/2022-78

SEI nº 9094096

AGRADECIMENTOS

Agradecer primeiramente a Deus pela sabedoria que me foi concedida para chegar até aqui, assim como pela minha vida e saúde. Sem Ele, eu nada seria, nada alcançaria.

Agradecer ao meu orientador, João Maricato, pelo incentivo a seguir em frente, por me fazer manter o foco em meio às dificuldades, por exigir qualidade em cada produção científica desenvolvida, me ensinando a fazer pesquisa, e por ter lutado junto comigo nessa jornada que foi o mestrado, sendo sempre muito paciente e atencioso. Obrigada, professor João, pela oportunidade, disponibilidade e pelos ensinamentos.

Agradecer ao meu marido Divino, pelo apoio nos momentos difíceis, pela paciência nos dias em que parecia que nada daria certo, e por sempre me incentivar. Seu amor e cuidado fez toda a diferença para que eu pudesse chegar até aqui, e seu amparo contribuiu para que eu alcançasse essa conquista. Obrigada por segurar minha mão tantas vezes. Obrigada por me apoiar nos momentos de medo e fraqueza e me lembrar que sou capaz.

Ao meu querido pai, Semi, que sempre sabe exatamente o que dizer para me incentivar e me trazer confiança, e que nunca duvida do meu potencial, mesmo quando até eu mesma duvido.

À minha mãe, Kátia, a quem eu dedico todos os esforços para concluir esta pesquisa. Sua alegria foi me ver iniciar este mestrado, e eu gostaria que ela estivesse aqui para me ver concluí-lo. A vida não espera uma fase passar para iniciar outra, e perder minha mãe durante o mestrado deixou uma sensação de incompletude ao concluí-lo.

Aos meus sogros, Elenir e Odemy, pelo carinho, apoio e cuidado.

Aos professores que participaram da minha banca de qualificação, pois deixaram suas pertinentes contribuições.

E a todos os amigos e familiares que me incentivaram e me apoiaram ao longo dessa jornada e que contribuíram de forma direta ou indiretamente com este trabalho.

RESUMO

Os periódicos científicos e seus respectivos artigos são canais de comunicação científica de reconhecida importância. A qualidade destas publicações pode ser medida, monitorada e avaliada quantitativamente por meio de indicadores, sendo tradicionalmente utilizados indicadores bibliométricos baseados em citação. Entretanto, o uso de indicadores altmétricos permite captar formas diversas de se mensurar o impacto acadêmico e social das publicações científicas. Nesse contexto, nesta pesquisa busca-se contribuir com o debate existente sobre as relações, características e complementaridades dos indicadores altmétricos e bibliométricos. Assim, o objetivo desta pesquisa é analisar as coberturas de fontes de dados e relações entre indicadores bibliométricos e altmétricos, no âmbito dos portais de periódicos de acesso aberto de universidades brasileiras. Com abordagem quantitativa e descritiva, foi feita a análise de coberturas de citação e indexação de artigos nas fontes de dados bibliográficas *Crossref*, *Scopus* e *Web of Science*, bem como as coberturas e menções aos artigos nas fontes de dados altmétricos *Mendeley*, *Twitter*, *Facebook*, de visualização e download de artigos dos portais. Além disso, são analisadas as correlações entre indicadores bibliométricos, altmétricos e entre eles, para todos os artigos, nas diferentes áreas do conhecimento, e de acordo com o ano de publicação do artigo. Os resultados das análises apontaram que as coberturas de citação e indexação de revistas e artigos na fonte de dados *Crossref*, em geral, apresentam-se mais altas, indicando que esta fonte tem grande potencial para o monitoramento e avaliação da ciência, especialmente em países com produção sub-representada em bases de dados comerciais internacionais. Em relação à cobertura das fontes altmétricas, destacaram-se os leitores no *Mendeley*, e sua relevância para determinadas áreas do conhecimento, como a Ciências da Vida. O *Mendeley* apresentou as maiores correlações com os dados de citação, nas diversas fontes, indicando que esta mídia tem maior capacidade de prever citações, e que os indicadores altmétricos desta fonte são menos complementares aos bibliométricos, podendo serem vislumbrados como alternativos para algumas áreas com menor inserção internacional. Indicadores altmétricos provenientes das mídias sociais *Twitter* e *Facebook* demonstraram exercer baixa ou nula influência nas citações de artigos no âmbito de portais de periódicos de universidades brasileiras de acesso aberto, sugerindo que há baixa sobreposição aos indicadores de citação, sendo os resultados alinhados aos argumentos que defendem que tais indicadores são complementares. As relações temporais entre os indicadores altmétricos e de citação analisadas, diferentemente de outros estudos, se apresentaram próximas a nulidade. As coberturas e correlações encontradas indicam haver diferenças nas práticas acadêmicas entre as áreas do conhecimento, portanto, há diferentes potencialidades dos indicadores para monitoramento e avaliação da ciência em cada uma delas. A partir destes resultados, conclui-se que novas fontes de dados, a exemplo da *Crossref* (citação) e *Mendeley* (altmétrico) são bastante promissoras para a avaliação da ciência publicada nos portais de periódicos. O uso de indicadores altmétricos de outras mídias sociais merece ser mais bem compreendido para a sua efetiva utilização no monitoramento e avaliação da ciência, porém demonstram ser um tipo de métrica complementar às citações. Ressalta-se por fim, a importância da ciência desenvolvida no Brasil e dos portais de periódicos gerenciados pelas universidades do país, havendo necessidade de maior presença destes em mídias sociais e melhor organização e descrição dos dados, das revistas e dos artigos por elas indexadas, em fontes de dados de grande potencial, a exemplo da *Crossref*.

Palavras-chave: Comunicação Científica. Altmetria. Indicadores bibliométricos. Acesso Aberto. Portais de Periódicos.

ABSTRACT

Scientific journals and their respective articles are channels of scientific communication of recognized importance. The quality of these publications can be quantitatively measured, monitored, and evaluated through indicators, traditionally using citation-based bibliometric indicators. However, the use of altmetric indicators allows capturing different ways of measuring academic and social impact of scientific publications. In this context, this research seeks to contribute to the existing debate on the relationships, characteristics, and complementarities of altmetric and bibliometric indicators. Thus, the objective of this research is to analyze the coverage of data sources and relationships between bibliometric and altmetric indicators within the scope of open access journal portals of Brazilian universities. With a quantitative and descriptive approach, has been made the analysis of citation coverage and indexing of articles in the bibliographic data sources *Crossref*, *Scopus* and *Web of Science*, as well as the coverage and mentions of articles in the altmetric data sources *Mendeley*, *Twitter*, *Facebook*, viewing and downloading articles from the portals. In addition, correlations are analyzed between bibliometric and altmetric indicators and between them, for all articles, in different areas of knowledge, and according to the year of publication of the article. The results of the analyzes showed that the citation and indexing coverage of journals and articles in the *Crossref* data source, in general, are higher, indicating that this source has great potential for monitoring and evaluating science, especially in countries with production underrepresented in international commercials databases. Regarding the coverage of altmetric sources, *Mendeley* readers were highlighted, and their relevance to certain areas of knowledge, such as Life Sciences. *Mendeley* presented the highest correlations with the citation data, in the different sources, indicating that this media has a greater capacity to predict citations, and that altmetric indicators from this source are less complementary to the bibliometric ones, and can be glimpsed as alternatives for some areas with less international insertion. Altmetric indicators from the social media *Twitter* and *Facebook* have shown to have little or no influence on citations of articles within the scope of journal portals from Brazilian open access universities, suggesting that there is low overlapping of citation indicators, with the results being aligned with the arguments that argue that such indicators are complementary. The temporal relationships between the analyzed altmetric and citation indicators, unlike other studies, were close to zero. The coverage and correlations found indicate that there are differences in academic practices between the areas of knowledge, therefore, there are different potentialities of the indicators for monitoring and evaluating science in each one of these areas. From these results, it is concluded that new data sources, such as *Crossref* (citation) and *Mendeley* (altmetric) are very promising for the evaluation of science published in journal portals. The use of altmetric indicators from other social media deserves to be better understood for their effective use in monitoring and evaluating science, but they prove to be a complementary type of metric to citations. Finally, the importance of the science developed in Brazil and the portals of journals managed by universities in the country are highlighted, with the need for a greater presence of these in social media and better organization and description of data, journals and articles indexed by them, in data sources with great potential, such as *Crossref*.

Keywords: Scientific Communication. Altmetrics. Bibliometric indicators. Open Access. Journal Portals.

LISTA DE FIGURAS

Figura 1 - Evolução do periódico científico ao longo do tempo	24
Figura 2 - Modelo de comunicação científica para as comunidades de Ciências Sociais e Humanas	25
Figura 3 - Distribuição dos documentos por idioma na base de dados <i>SciELO</i> (2021)	31
Figura 4 - Esquema de relações entre os campos da Informetria, Bibliometria, Cientometria, Cibermetria, Webometria, Webmetria e Altmtria	45
Figura 5 - Legenda de cores de acordo com a fonte que compõem o <i>donut</i> do <i>Altmtric Attention Score</i>	66
Figura 6 - Distribuição geral das publicações com alométricas	72
Figura 7 - Porcentagens de citações encontradas por área de conhecimento	73
Figura 8 - Relevância dos <i>tweets</i> de comunicação científica por disciplina	86
Figura 9 - Estatísticas descritivas de citação e alométricas dos artigos da PLOS divididas por ano	93
Figura 10 - Correlações entre citações e outras métricas dos artigos divididas por ano	94
Figura 11 - Distribuição das médias de citação (MCS) e leitores no <i>Mendeley</i> (MRS) ao longo do tempo	95
Figura 12 - Interface da ferramenta Odisseia	103
Figura 13 - Matriz de correlação entre leitores do <i>Mendeley</i> e citações na Crossref	126
Figura 14 - Matriz de correlação entre <i>tweets</i> e citações na Crossref	126
Figura 15 - Matriz de correlação entre downloads e citações na Crossref	126
Figura 16 - Matriz de correlação entre menções no <i>Facebook</i> e citações na Crossref	126
Figura 17 - Matriz de correlação entre acessos ao resumo e citações na Crossref	126
Figura 18 - Matriz de correlação entre citações Scopus e citações na Crossref	126
Figura 19 - Matriz de correlação entre citações na Web of Science e citações na Crossref ..	127
Figura 20 - Matriz de correlação entre citações na WoS e citações na Scopus	127
Figura 21 - Matriz de correlação entre leitores no <i>Mendeley</i> e downloads	127
Figura 22 - Matriz de correlação entre leitores no <i>Mendeley</i> e menções no <i>Facebook</i>	127
Figura 23 - Matriz de correlação entre leitores no <i>Mendeley</i> e acesso ao resumo	127
Figura 24 - Matriz de correlação entre downloads e <i>Facebook</i>	127
Figura 25 - Matriz de correlação entre leitores no <i>Mendeley</i> e citações na Scopus	127
Figura 26 - Matriz de correlação entre leitores no <i>Mendeley</i> e <i>tweets</i>	127
Figura 27 - Matriz de correlação entre leitores no <i>Mendeley</i> e citações na WoS	127

Figura 28 - Matriz de correlação entre acessos ao resumo e downloads	128
Figura 29 - Matriz de correlação entre acessos ao resumo e menções no <i>Facebook</i>	128
Figura 30 - Matriz de correlação entre downloads e citações na Scopus	128
Figura 31 - Matriz de correlação entre menções no <i>Facebook</i> e citações na Scopus	128
Figura 32 - Matriz de correlação entre acessos ao resumo e citações na Scopus	128
Figura 33 - Matriz de correlação entre downloads e <i>tweets</i>	128
Figura 34 - Matriz de correlação entre <i>tweets</i> e menções no <i>Facebook</i>	128
Figura 35 - Matriz de correlação entre acessos ao resumo e <i>tweets</i>	128
Figura 36 - Matriz de correlação entre citações na Scopus e <i>tweets</i>	128
Figura 37 - Matriz de correlação entre citações na WoS e <i>tweets</i>	129
Figura 38 - Matriz de correlação entre downloads e citações na WoS	129
Figura 39 - Matriz de correlação entre menções no <i>Facebook</i> e citações na WoS	129
Figura 40 - Matriz de correlação entre acessos ao resumo e citações na WoS	129
Figura 41 - Matriz de correlação entre citações na WoS e citações na Scopus	129

LISTA DE QUADROS

Quadro 1 - Definição dos tipos de acesso aberto	33
Quadro 2 - Indicadores bibliométricos mais usados para avaliar pesquisadores individuais e produção científica	49
Quadro 3 - Colégios e grandes áreas de conhecimento definidos pela CAPES para avaliação de Programas de Pós-graduação, e número de revistas por grande área de conhecimento.	102
Quadro 5 - Intervalo numérico e testes de normalidade para as variáveis da totalidade dos artigos analisados	108
Quadro 6 - Intervalo numérico e testes de normalidade para as variáveis dos artigos analisados, por área de conhecimento	108
Quadro 7 - Intervalos para a interpretação e análise dos Coeficientes de Correlação de Spearman	110

LISTA DE TABELAS

Tabela 1 - Número de revistas com ISSN e número de revistas	100
Tabela 2 - Total de menções, coberturas, média e mediana dos artigos, por fonte de dados alométricos	114
Tabela 3 - Coberturas de indexação, médias e medianas de citações nas fontes de dados bibliométricas	118
Tabela 4 - Correlações de Spearman entre indicadores alométricos e bibliométricos para todos os artigos	121
Tabela 5 - Indicadores alométricos de menções, coberturas, média e mediana de menções por artigos por área de conhecimento dos portais de periódicos	132
Tabela 6 - Total de citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais por área de conhecimento, de acordo com as fontes de dados bibliométricas	137
Tabela 7 - Correlações de Spearman entre indicadores alométricos e bibliométricos para os artigos dos portais por área de conhecimento	141
Tabela 8 - Total de citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais com algum indicador alométrico de acessos ao resumo ou downloads, de acordo com as fontes de dados bibliométricas nas quais as revistas estão indexadas	144
Tabela 9 - Total de citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais por ano de publicação, de acordo com as fontes de dados bibliométricas nas quais estão indexados	146
Tabela 10 - Total de menções, coberturas, médias e medianas de menções para artigos dos portais com algum indicador alométrico de acessos ao resumo ou downloads ..	147
Tabela 11 - Correlações de Spearman entre indicadores alométricos de acessos e downloads, e indicadores bibliométricos para os artigos dos portais por ano de publicação ...	149

LISTA DE ABREVIATURAS E SIGLAS

<i>A&HCI</i>	<i>Arts & Humanities Citation Index</i>
<i>AAS</i>	<i>Altmetric Attention Score</i>
<i>ABNT</i>	<i>Associação Brasileira de Normas Técnicas</i>
<i>APC</i>	<i>Article Processing Charge</i>
<i>API</i>	<i>Application Programming Interface</i>
<i>BASE</i>	<i>Bielefeld Academic Search Engine</i>
<i>BOAI</i>	<i>Budapest Open Access Initiative</i>
<i>CAPES</i>	<i>Coordenação de Aperfeiçoamento de Pessoal de Nível Superior</i>
<i>CD-ROM</i>	<i>Compact Disc Read Only Memory</i>
<i>CETM</i>	<i>Ciências Exatas, Tecnológicas e Multidisciplinar</i>
<i>CH</i>	<i>Ciências Humanas</i>
<i>COCI</i>	<i>Citações abertas de DOI para DOI da Crossref</i>
<i>CPP</i>	<i>Citation Per Publication</i>
<i>CS</i>	<i>Dados de Citação da Scopus</i>
<i>CSA</i>	<i>Ciências Sociais Aplicadas</i>
<i>CSV</i>	<i>Comma Separated Values</i>
<i>CTC-ES</i>	<i>Conselho Técnico-Científico da Educação Superior</i>
<i>CV</i>	<i>Ciência da Vida</i>
<i>CW</i>	<i>Dados de Citação da Web of Science</i>
<i>DOAJ</i>	<i>Directory of Open Access Journals</i>
<i>DOI</i>	<i>Digital Object Identifier</i>
<i>FI</i>	<i>Fator de Impacto</i>
<i>HTML</i>	<i>HyperText Markup Language</i>
<i>IBICT</i>	<i>Instituto Brasileiro de Informação em Ciência e Tecnologia</i>
<i>ID</i>	<i>Identificador</i>
<i>IES</i>	<i>Instituições de Ensino Superior</i>
<i>ISBN</i>	<i>International Standard Book Number</i>
<i>ISSN</i>	<i>International Standard Serial Number</i>
<i>JCR</i>	<i>Journal Citation Reports</i>
<i>JIF</i>	<i>Journal Impact Factor</i>
<i>JMIR</i>	<i>Journal of Medical Internet Research</i>
<i>JSON</i>	<i>Java Script Object Notation</i>

<i>LISA</i>	<i>Library & Information Science Abstracts</i>
<i>LLA</i>	<i>Linguística, Letras e Artes</i>
<i>MEDLINE</i>	<i>Medical Literature Analysis and Retrieval System Online</i>
<i>OA</i>	<i>Open Access</i>
<i>OAAA</i>	<i>Open Access Altmetrics Advantage</i>
<i>OACA</i>	<i>Open Access Citation Advantage</i>
<i>OAI</i>	<i>Open Archives Initiative</i>
<i>OAI-PMH</i>	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>
<i>OD</i>	<i>Odisseia Metrics</i>
<i>OE</i>	<i>Objetivo Específico</i>
<i>OJS</i>	<i>Open Journal Systems</i>
<i>ORCID</i>	<i>Open Researcher and Contributor</i>
<i>PDF</i>	<i>Portable Document Format</i>
<i>PKP</i>	<i>Public Knowledge Project</i>
<i>PLOS</i>	<i>Public Library of Science</i>
<i>POS</i>	<i>Public Library of Science</i>
<i>RAA</i>	<i>Relatório de Acesso por Ano</i>
<i>RDA</i>	<i>Relatório de Downloads por ano</i>
<i>RWWR</i>	<i>Ranking Web of World Repositories</i>
<i>SCI</i>	<i>Science Citation Index</i>
<i>SciELO</i>	<i>Scientific Electronic Library Online</i>
<i>SJR</i>	<i>SCImago Journal Rank</i>
<i>SSCI</i>	<i>Social Sciences Citation Index</i>
<i>TIC</i>	<i>Tecnologias da Informação e Comunicação</i>
<i>UFG</i>	<i>Universidade Federal de Goiás</i>
<i>UFPR</i>	<i>Universidade Federal do Paraná</i>
<i>UFRN</i>	<i>Universidade Federal do Rio Grande do Norte</i>
<i>UFSC</i>	<i>Universidade Federal de Santa Catarina</i>
<i>UnB</i>	<i>Universidade de Brasília</i>
<i>USP</i>	<i>Universidade de São Paulo</i>
<i>WoS</i>	<i>Web of Science</i>

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Problemas de Pesquisa e justificativa	14
1.2 Objetivos	19
2 REFERENCIAL TEÓRICO-CONCEITUAL	20
2.1 Ciência, produção e comunicação científica	20
2.1.1 Periódicos Científicos Brasileiros	28
2.1.2 Periódicos de Acesso Aberto e sua indexação	32
2.2 Indicadores e Métricas de Produção Científica	43
2.3 Indicadores de Qualidade de Periódicos Científicos	46
2.4 Bibliometria e estudos de citação	50
2.4.1 Fonte de dados bibliométricos	53
2.5 Almetria e a mensuração na Web 2.0	56
2.5.1 Fontes de dados alométricos	65
2.6 Estudos sobre coberturas de fontes de informação bibliométricas e alométricas	69
2.7 Estudos sobre correlações entre indicadores bibliométricos e alométricos	74
2.8 Estudos sobre indicadores alométricos e bibliométricos em diferentes áreas de conhecimento	83
2.9 Estudos sobre a influência do fator temporal entre indicadores alométricos e bibliométricos	91
3 METODOLOGIA	97
3.1 Caracterização da pesquisa	97
3.2 Coleta e tratamento dos dados	98
3.2.1 Etapa 1: Seleção dos Portais de Periódicos	98
3.2.2 Etapa 2: Identificação de DOIs de publicações	99
3.2.3 Etapa 3: Identificação das áreas de conhecimento	101
3.2.4 Etapa 4: Coleta dos dados alométricos e bibliométricos	102
3.2.5 Etapa 5: Tratamento dos dados coletados	105
3.2.6 Etapa 6: Testes estatísticos e de cobertura	107
3.3 LIMITAÇÕES DA PESQUISA	111
4 RESULTADOS E DISCUSSÃO	113
4.1 COBERTURAS E CORRELAÇÕES DE INDICADORES ALTMÉTRICOS E BIBLIOMÉTRICOS PARA OS ARTIGOS DOS PORTAIS	113

4.1.1 Coberturas de indicadores alométricos e bibliométricos	113
4.1.2 Correlações entre indicadores alométricos e bibliométricos	120
4.2 COBERTURAS E CORRELAÇÕES DE INDICADORES ALMÉTRICOS E BIBLIOMÉTRICOS PARA OS ARTIGOS DOS PORTAIS POR ÁREA DE CONHECIMENTO	130
4.2.1 Coberturas de indicadores alométricos e bibliométricos para artigos dos portais por área de conhecimento	130
4.2.2 Correlações entre indicadores alométricos e bibliométricos para artigos dos portais por área de conhecimento	138
4.3 COBERTURAS E CORRELAÇÕES DE INDICADORES ALMÉTRICOS E BIBLIOMÉTRICOS PARA OS ARTIGOS DOS PORTAIS DE ACORDO COM A DATA DE PUBLICAÇÃO	143
4.3.1 Coberturas de indicadores alométricos e bibliométricos para artigos dos portais de acordo com a data de publicação	143
4.3.2 Correlações entre indicadores alométricos e bibliométricos para artigos dos portais de acordo com a data de publicação	148
5 CONSIDERAÇÕES FINAIS.....	152
REFERÊNCIAS	155

1 INTRODUÇÃO

A motivação de um pesquisador em publicar suas descobertas está embasada em três motivos principais: reconhecimento pela comunidade; proteção da propriedade intelectual; e para declarar a descoberta de novos conhecimentos (PRICE, 1976). É através da comunicação científica em seus canais formais e informais que estas descobertas são comunicadas, sendo a comunicação tão importante para a ciência quanto a própria descoberta em si (MEADOWS, 1999). Quando falamos no canal oficial de comunicação científica mais utilizado e reconhecido, os periódicos científicos, um complexo sistema de publicação entra em cena para garantir a proteção da autoridade e prestígio da ciência através de critérios rígidos de avaliação das pesquisas e regulação de todos os elementos envolvidos no fluxo editorial.

Na busca pelo reconhecimento de seu trabalho e prestígio no meio acadêmico, o pesquisador tende a dar preferência para publicar suas descobertas em periódicos que agreguem qualidade ao seu trabalho. Mas como saber quais periódicos são de qualidade? Através do sistema de avaliação da produção científica, que utiliza métricas que permitem quantificar diversos aspectos das publicações e periódicos científicos, e construir indicadores a serem utilizados na avaliação da produção. Tradicionalmente, indicadores bibliométricos baseados em citação vêm sendo utilizados para atribuir qualidade a periódicos científicos e, conseqüentemente, aos trabalhos publicados nestes.

Nesse cenário, outra questão relevante a ser ponderada pelo pesquisador é o tipo de acesso que será concedido à sua pesquisa ao ser publicada. Periódicos científicos podem cobrar, por meio do pagamento de assinatura ou taxas, para que os leitores acessem o conteúdo publicado, o chamado acesso restrito, ou podem disponibilizar as publicações gratuitamente, prática conhecida como acesso aberto.

Dentro do escopo dos estudos de acesso aberto, hipóteses que apontam para uma possível vantagem de publicações disponibilizadas em acesso aberto receberem mais citações e menções online do que aquelas publicadas em acesso restrito ganharam força e comprovação científica ao longo do tempo (CINTRA; FURNIVAL; MILANEZ, 2017). Dessa forma e, baseado nessas hipóteses, analisar indicadores de publicações de periódicos de acesso aberto se mostra uma abordagem promissora em captar o impacto acadêmico e social dentro do escopo destes periódicos.

Com o surgimento da internet e a migração dos periódicos científicos do meio impresso para o meio eletrônico, surge a possibilidade de estruturar vários periódicos em um único site, organizando e disponibilizando as publicações científicas em portais de periódicos.

Os periódicos hospedados nestes portais geralmente estão sujeitos a uma gestão unificada que provê apoio técnico, financeiro, e garante a preservação da informação científica publicada, sendo comum que instituições de ensino superior adotem portais de periódicos para estruturar sua produção acadêmica. O crescimento da importância dos periódicos científicos como instrumento de avaliação dos programas de pós-graduação das universidades também contribui para aumentar o interesse das instituições de ensino na manutenção e qualificação dos seus periódicos, especialmente os de acesso aberto (RODRIGUES; FACHIN, 2010).

Com o avanço da tecnologia e o surgimento da Web 2.0, além da possibilidade de estruturar periódicos em portais, o alcance da comunicação científica rompeu barreiras geográficas e tornou as publicações científicas mais facilmente acessíveis para além da comunidade acadêmica. Publicações online passam a ser compartilhadas na web nos mais diversos ambientes virtuais, incluindo nas redes sociais.

Nesse novo cenário, surgem indicadores baseados em métricas provenientes das atividades e ações, relacionadas a publicações científicas, nas redes sociais digitais, como compartilhamento, downloads, acessos, menções e leituras. Esses novos indicadores, chamados indicadores alométricos, possibilitam uma visão mais ampla do impacto gerado pelas publicações científicas e permitem quantificar este impacto na sociedade como um todo, e não apenas no meio acadêmico, como é o caso dos indicadores bibliométricos.

Ao considerarmos a almetria como uma forma alternativa de medir o impacto, seus indicadores podem complementar e melhorar as limitações das métricas tradicionais, tais como as observadas pela bibliometria, e possibilitar novos *insights* na análise de impacto (GALLIGAN; DYAS-CORREIA, 2013), inclusive expandindo o conceito de impacto acadêmico para outros tipos de impacto, como o social.

Dessa forma, a complementaridade entre indicadores alométricos e bibliométricos deve ser investigada. Para isso, é necessário realizar análises comparativas entre os indicadores, como o estudo da cobertura de artigos indexados em bases de dados bibliográficas e a ocorrência destes artigos nas fontes de dados alométricos.

Diversos estudos analisaram a cobertura das fontes de dados alométricos, como os de Costas, Zahedi e Wouters (2014), Haustein, Costas e Larivière (2015), Robinson-García *et al.* (2014), Wang *et al.* (2014) e Zahedi e Haustein (2018), e identificaram o *Mendeley* e o *Twitter* como fontes de dados alométricos com maior cobertura. Outras pesquisas, como as de Haustein e Larivière (2014), Ouchi *et al.* (2019), Priem, Piwowar e Hemminger (2011), Thelwall (2019), Thelwall *et al.* (2013) e Li, Thelwall e Giustini (2012), analisaram as correlações entre os indicadores alométricos e bibliométricos. Os resultados de Wang *et al.*

(2014) e Zahedi, Costas e Wouters (2017) indicam que as correlações entre o indicador altmétrico de leitores do *Mendeley* e o indicador bibliométrico de citação são mais fortes ao considerarmos os primeiros 5 anos de publicação do artigo. Já Mohammadi *et al.* (2015), Haustein e Larivière (2014), e Mohammadi e Thelwall (2013) afirmam que as correlações entre os indicadores são mais fortes para publicações de revistas de áreas de conhecimento específicas, como as áreas de Saúde, como Ciências Médicas e Biologia; e ciências duras, como Engenharia e Bioquímica.

Tendo em vista o contexto apresentado, a presente pesquisa procura contribuir com as análises de coberturas de fontes de dados e de correlações entre os indicadores altmétricos e bibliométricos, estudando o cenário de portais de periódicos de acesso aberto de universidades brasileiras e a ocorrência de seus artigos em uma variedade de bases de dados, em busca de compreender as dinâmicas e relações entre indicadores bibliométricos e indicadores altmétricos no âmbito de portais de periódicos de acesso aberto. Para isso, são necessárias análises com foco nas coberturas de fontes de dados; temáticas; correlacionais; e, temporais dos artigos científicos.

1.1 PROBLEMAS DE PESQUISA E JUSTIFICATIVA

Acredita-se que os indicadores bibliométricos podem ser utilizados em conjunto com os altmétricos, trazendo benefícios para a avaliação do impacto acadêmico das publicações científicas. Para entender melhor as possibilidades e limitações de uso desses indicadores, é necessário a realização de análises de suas coberturas e análises comparativas entre eles. A partir destas análises, é possível identificar se há influência ou relações entre indicadores altmétricos e bibliométricos de citação e sobre as suas potencialidades para a avaliação da ciência. Devido ao campo da altmetria ser relativamente recente, existe a necessidade de mais estudos que analisem as correlações entre os indicadores altmétricos e os indicadores bibliométricos de citação nas diferentes fontes de dados.

As análises que se debruçam sobre os artigos publicados em revistas brasileiras, fortemente vinculadas a portais de periódicos de acesso aberto vinculados às universidades, merecem atenção especial. A maioria dos estudos com abordagens semelhantes são feitas a partir da ciência indexada em grandes bases de dados internacionais, havendo pouco conhecimento sobre as potencialidades de uso e das relações entre indicadores bibliométricos e altmétricos para a avaliação da ciência desenvolvida no país.

A produção, disponibilização e uso de periódicos em portais levanta novas questões, uma vez que até pouco tempo, nas Instituições de Ensino Superior (IES), os periódicos eram considerados apenas de forma individual (GARRIDO; RODRIGUES, 2010). Os portais de periódicos compõem uma estrutura unificada de gestão por uma mesma instituição editora, reunindo uma coleção multidisciplinar de periódicos, com seus diferentes editores, políticas e iniciativas.

A estruturação dos periódicos de IES em portais causa reflexos diretos nas atividades de busca e disseminação da pesquisa, já que essa estrutura permite uma ampliação das possibilidades e estratégias de divulgação científica. Uma vez que o periódico passa a pertencer a uma coleção maior e unificada de periódicos, o foco das atividades de busca e disseminação da pesquisa passa a se concentrar a nível de artigo (SOUTO, 2007). Sendo o Brasil um país que conta com um forte sistema de publicação em revistas de acesso aberto, grande parte dos periódicos científicos estão hospedados em portais de periódicos de acesso aberto vinculados às universidades e instituições de pesquisa.

A maioria dos estudos altmétricos são desenvolvidos analisando publicações de periódicos internacionais, indexados em grandes bases de dados como *Web of Science* (COSTAS; ZAHEDI; WOUTERS, 2014, HAUSTEIN; COSTAS; LARIVIÈRE, 2015, MOHAMMADI; THELWALL, 2013) e *Scopus* (THELWALL; SUD, 2016), ou provenientes de revistas de reputação reconhecida na comunidade científica, como a revista *Nature* (OUCHI *et al.*, 2019), e revistas de grandes portais de periódicos internacionais, como o portal da *PLOS* (LIU *et al.*, 2013, WANG *et al.*, 2014). Dessa forma, a seleção das fontes de dados da maioria dos estudos altmétricos é direcionada para coleta de dados de publicações de língua inglesa, e provenientes de periódicos de maior expressão internacional.

A possibilidade de estudar a cobertura dos indicadores altmétricos para publicações de periódicos brasileiros de acesso aberto abre caminho para obtenção de informações estratégicas sobre o impacto social dessas publicações, sendo informações extremamente relevantes que podem auxiliar na tomada de decisão pelas instituições e editores responsáveis pelos periódicos dos portais das universidades, influenciando na distribuição de fomento entre os periódicos de um mesmo portal institucional, por exemplo.

Enquanto o estudo da cobertura das fontes de dados se mostra importante para entender quais destas fontes estão reunindo maiores informações das publicações e sendo mais utilizadas pela comunidade, a análise das correlações entre os indicadores altmétricos e bibliométricos provindos destas fontes nos permite observar as dinâmicas e influências entre elas. Por exemplo, é possível compreender melhor sobre as complementaridades entre os

indicadores altmétricos e bibliométricos e suas possibilidades de uso para avaliação de áreas do conhecimento específicas.

Alguns problemas que esta pesquisa busca confrontar são relacionados à limitação das coberturas da ciência brasileira em bases de dados internacionais, uma vez que a maioria das revistas científicas brasileiras de universidades não estão indexadas nas grandes bases de dados, havendo limitações para a sua avaliação. De acordo com Farias (2017),

Grande parte das revistas brasileiras, e nesse caso tanto revistas científicas de acesso aberto como de outros modelos de negócios, ainda não alcançou posições de destaque nas chamadas "grandes bases de dados indexadoras internacionais", tais como a JCR e SJR. Aumentar a presença brasileira nesses indexadores internacionais é muito importante, já que, ao estar neles presente, o efeito esperado será o crescimento da visibilidade da ciência nacional, aumentando o seu número de citações e, assim, possibilitando a criação de novas redes de colaboração e produção de conhecimento científico. (FARIAS, 2017, p. 403).

Em complementação à afirmação de Farias (2017), não apenas o número de citações crescerá com o aumento da visibilidade da ciência nacional, mas também se espera um crescimento das menções em redes sociais que a produção científica irá receber, impactando diretamente na cobertura dos indicadores altmétricos provenientes de várias fontes de dados. Alperin (2014) identificou em sua pesquisa que as métricas mais prevalentes de redes sociais digitais são provenientes do *Twitter* e *Mendeley*, inclusive no contexto brasileiro. Porém, as coberturas obtidas a partir de publicações brasileiras indexadas na *Scientific Electronic Library Online (SciELO)* são bem inferiores às identificadas de publicações provindas de bases como *PubMed*¹ e *Web of Science*, confirmando a deficiência em cobertura da ciência brasileira em bases de dados internacionais.

Melo, Trinca e Maricato (2021) confirmam a baixa cobertura de indicadores bibliométricos de bases de dados internacionais para artigos brasileiros de diferentes áreas de conhecimento, ao levantarem dados da produção bibliográfica dos 4.377 Programas de Pós-Graduação cadastrados na Plataforma Sucupira nos anos de 2017 e 2018 e analisarem a indexação de artigos e periódicos destes programas na *Web of Science*. Os autores identificaram que “[...] 9.700 (46,56%) periódicos utilizados para publicação de artigos por pesquisadores vinculados aos Programas de Pós-Graduação brasileiros estão indexados na base, enquanto 11.135 (53,44%) periódicos utilizados não estão indexados” (MELO, TRINCA, MARICATO, 2021, p. 6).

¹ Disponível em: <https://pubmed.ncbi.nlm.nih.gov/>. Acesso em: 12 dez. 2021

É justamente em busca de analisar este recorte de publicações tão carente de estudos mais profundos que este estudo se direciona. Analisar coberturas de fontes de dados e correlações entre indicadores altmétricos e citações em portais de periódicos brasileiros de acesso aberto se mostra um escopo de estudo que contribuirá para entender melhor a possibilidade do uso dos indicadores altmétricos para avaliação da ciência. Para Alperin (2014), a diferença de resultados de cobertura obtidos entre seu estudo e estudos anteriores que utilizaram bases de dados internacionais ressalta a importância de considerar mais estudos altmétricos no contexto de países emergentes e em desenvolvimento, como o Brasil.

Outro problema, além da cobertura, é que não se sabe quais as melhores bases de dados para se avaliar a ciência brasileira. Quando falamos sobre avaliação da qualidade da ciência, nos referimos à qualidade da pesquisa, que nos leva a focar também na qualidade do periódico responsável por sua publicação (MINGERS; YANG, 2016).

Pesquisas publicadas em periódicos brasileiros têm, assim como pesquisas internacionais, sua qualidade avaliada seguindo critérios que utilizam métricas baseadas em números de citação. Dessa forma, seria relevante analisar indicadores bibliométricos de citação de diversas fontes de dados, como feito Martín-Martín *et al.* (2021). Porém, aqui entra uma discussão crescente desde o início dos estudos de complementaridade entre a altmetria e a bibliometria: o uso de indicadores altmétricos na composição dos atributos de avaliação da ciência. O potencial da altmetria em complementar a bibliometria nesta tarefa, atende à necessidade, também, de captar o impacto da ciência além da comunidade científica. Para Maricato e Martins (2017),

[...] apesar de haver a necessidade de reconhecer qualidades (e limitações) dos indicadores utilizados até então, reconhece-se que os baseados em citação, sozinhos, não captam o espectro completo do “impacto” de uma pesquisa científica e que metodologias complementares devem ser consideradas. Diante dessa premissa, uma nova área emergente é a altmetria. (MARICATO; MARTINS, 2017, p.49).

Dessa forma, estudar a correlação entre indicadores altmétricos e bibliométricos é relevante ao contribuir no debate existente de que estes indicadores podem ser úteis em compor novos indicadores para a avaliação da qualidade da produção científica, justamente por captarem novas dimensões do impacto científico. Apesar desse potencial, ainda não temos certeza em quais áreas do conhecimento a altmetria pode ser utilizada para essa avaliação da qualidade. Isso porque estudos anteriores obtiveram resultados diversos quanto à força das correlações entre indicadores altmétricos e bibliométricos de acordo com a área de

conhecimento das publicações estudadas (LI; THELWALL; GIUSTINI, 2012, HAUSTEIN; LARIVIÉRE, 2014, MOHAMMADI *et al.*, 2014, THELWALL *et al.*, 2013).

Assim, no contexto brasileiro, é relevante estudar as correlações dos indicadores altmétricos e bibliométricos dos artigos dos portais de periódicos de acesso aberto nas diferentes áreas do conhecimento. Isso propicia, por um lado, maior conhecimento sobre esses indicadores e suas potencialidades para a avaliação da ciência brasileira e, por outro lado, estimula discussões sobre o papel dos portais de periódicos de acesso aberto na produção, organização e divulgação do conhecimento científico, ao ofertar indicadores consistentes para a tomada de decisões por parte dos gestores desses portais.

A correlação entre os indicadores altmétricos e bibliométricos por área de conhecimento é importante de ser abordada para compreendermos melhor as dinâmicas entre estes indicadores nas diferentes áreas. Afinal, assim como as práticas de comunicação científica e as preferências de publicação variam de acordo com a área de conhecimento (MUELLER, 2005), é de se esperar que essas práticas afetem as relações entre os indicadores que quantificam exclusivamente métricas de publicações em periódicos científicos.

A análise das dinâmicas entre os indicadores por área permite identificar se há variação na força das correlações entre os indicadores altmétricos e bibliométricos e se esta variação é reflexo das práticas de comunicação científica de cada área, contribuindo no debate de se a complementaridade entre esses indicadores para avaliação da qualidade da produção científica é adequada para cada área de conhecimento.

Finalmente, o fator tempo também deve ser mais bem compreendido nas análises dos indicadores e, para isso, a coleção de artigos analisados deve ser dividida por fatias temporais, de acordo com a data de publicação, em busca de identificar a influência do tempo nos resultados obtidos, e se a seleção de pesquisas mais recentes afeta a força das correlações entre indicadores e a cobertura destes nas fontes de dados.

Diversos estudos de correlações entre indicadores altmétricos e bibliométricos identificaram um aumento na força da correlação para publicações de anos mais recentes (COSTAS; ZAHEDI; WOUTERS, 2014; HAUSTEIN; LARIVIÉRE, 2014). Um dos problemas a serem resolvidos nesta pesquisa diz respeito à confirmação ou não desta hipótese encontrada em estudos internacionais, porém voltado ao escopo de publicações de portais de periódicos brasileiros de acesso aberto.

De forma geral, a compreensão de todas estas dinâmicas entre indicadores e coberturas para este escopo contribui para o crescimento como um todo da área da altmetria, bem como

para um melhor entendimento das dinâmicas das relações entre citação e atividades em mídias sociais digitais a nível nacional.

Dito isso e, considerando a circulação do conhecimento científico na web social como uma nova possibilidade para se avaliar o impacto acadêmico, faz-se então necessário conhecer as ferramentas, mapear e analisar como os artigos científicos são utilizados nesse ambiente, e entender os efeitos que essa circulação gera nos próprios artigos. Por essa perspectiva, e dentro da temática de indicadores aplicados a portais de periódicos, é possível levantar algumas questões a serem exploradas, quais sejam: qual a cobertura encontrada para cada fonte de dados? Existe relação entre o fator tempo e a força das correlações entre os indicadores altmétricos e bibliométricos? A altmetria é mais adequada para avaliar determinadas áreas do conhecimento? No contexto brasileiro, qual o impacto dos indicadores altmétricos no número de citações? A partir de tais questionamentos propõem-se os seguintes objetivos a esta pesquisa.

1.2 OBJETIVOS

O objetivo geral desta pesquisa é analisar as coberturas de fontes de dados e relações entre indicadores bibliométricos e indicadores altmétricos no âmbito dos portais de periódicos de acesso aberto de universidades brasileiras.

A partir deste, os seguintes objetivos específicos foram traçados:

- Verificar a cobertura das fontes de dados altmétricos e bibliométricos de artigos dos portais nas diferentes áreas do conhecimento;
- Investigar as correlações dos indicadores altmétricos e bibliométricos dos artigos dos portais nas diferentes áreas do conhecimento;
- Analisar o fator temporal nas correlações entre indicadores bibliométricos e altmétricos dos artigos dos portais.
- Analisar o fator temporal na cobertura das fontes de dados altmétricos e bibliométricos dos artigos dos portais.

2 REFERENCIAL TEÓRICO-CONCEITUAL

Neste capítulo é apresentado o referencial teórico, onde serão abordados aspectos fundamentais para o embasamento desta pesquisa. Inicialmente apresenta-se o contexto geral de conceituação da ciência e os processos de comunicação científica, chegando ao principal canal de publicação da ciência, os periódicos científicos, e as mudanças pelas quais passaram com o advento da Web 2.0.

Considerando o escopo desta pesquisa, são apresentadas as definições, características, estrutura e atores de periódicos e portais de periódicos de acesso aberto brasileiros. Posteriormente, é abordado o sistema de avaliação da produção científica e a qualidade de periódicos e suas publicações, bem como as métricas utilizadas para avaliação. Finalmente, apresenta-se os indicadores altmétricos e bibliométricos, sua aplicação a portais de periódicos de acesso aberto, a importância da cobertura das fontes de informação para construção de indicadores, as correlações entre indicadores altmétricos e bibliométricos, inclusive em diferentes áreas de conhecimento, e a influência do fator temporal nas correlações entre indicadores altmétricos e bibliométricos.

2.1 CIÊNCIA, PRODUÇÃO E COMUNICAÇÃO CIENTÍFICA

Dizer o que é ciência depende do ponto de vista. Para Ziman (1987) ela pode ser um conhecimento organizado. Do ponto de vista metodológico, segundo o autor, ciência é um método de obter informação confiável sobre o mundo, ou, em seu aspecto profissional, algo descoberto por pessoas com vocação para pesquisa. Ciência também é o produto da pesquisa, uma instituição social, um recurso cultural. Cada disciplina considera um modelo diferente de ciência, uma vez que cada disciplina tem suas características e interesses específicos de estudo, o que reflete no modo de fazer ciência.

A comunicação formal da observação do mundo natural pelo cientista é composta por registros públicos do conhecimento obtido pelo pesquisador, que do ponto de vista documental, se caracteriza como ciência. O princípio básico da ciência é que os resultados de pesquisas devem ser comunicados publicamente. Sendo assim, a instituição social fundamental da ciência é a comunicação. Nas palavras de Meadows (1999), a comunicação situa-se no próprio coração da ciência. Os resultados das pesquisas devem ser públicos e as descobertas só serão consideradas parte do conhecimento científico ao serem relatadas e colocadas em registros permanentes.

Antes do surgimento dos periódicos científicos, as notícias sobre o conhecimento científico (novas descobertas e invenções, por exemplo) eram comunicadas através de jornais e folhetins cotidianos, que são meios de comunicação destinados ao público em geral. Até então, o conhecimento especializado era comunicado por correspondências realizadas entre os cientistas ou enviadas às academias (FREITAS, 2006). Os primeiros periódicos científicos surgiram da necessidade de intelectuais em publicar de maneira impressa cartas importantes sobre novas ideias e pesquisas de grupos intelectuais e filosóficos em meados do século XVII (MEADOWS, 1999). Desde então, diversos esforços foram feitos pelas sociedades, que abrangiam diversas áreas do saber, em busca da expansão do conhecimento e distribuição dessas publicações impressas.

As publicações impressas de teor científico especializado passaram então a ser o principal meio de comunicação da ciência, funcionando como um arquivo da ciência, uma coleção de contribuições que vão somando-se e que servem de ponto de partida para novas investigações, uma conexão que se torna explícita com o uso das citações de referências. Os documentos publicados possuem conhecimento confiável, importante característica da ciência que separa o que é conhecimento científico do que é popular, não científico (BARBA, 2003).

Para obter confiabilidade, é utilizada uma metodologia científica rigorosa para a produção e publicação do conhecimento, e é importante que os resultados obtidos pelas pesquisas de um cientista sejam divulgados e submetidos ao julgamento de outros cientistas, seus pares (MUELLER, 2005). Essa avaliação por pares é parte essencial do sistema de publicação científica, composto por elementos que sustentam, regulam e perpetuam o processo de fazer acessível de modo oficial ao resto da comunidade científica as pretensões do pesquisador de contribuir ao acervo científico (BARBA, 2003).

A revisão por pares possui duas características primordiais: a paridade, onde se mostram necessários dois ou mais avaliadores para minimizar erros na avaliação; e o anonimato, mantendo oculta a identidade tanto do autor do trabalho para os avaliadores, quanto dos avaliadores para o autor (BARBA, 2003). É o funcionamento desse sistema de avaliação por pares que dá o veredicto se um trabalho é cientificamente aceitável ou não. Ao ser aceito na revisão por pares, o trabalho é encaminhado para publicação e passa a fazer parte de uma coleção de resultados científicos.

A publicação de artigos revisados por pares em periódicos científicos é apenas o primeiro passo de um longo processo de reconhecimento do cientista e do seu trabalho, no qual o sistema de comunicação tem um papel vital. Um pesquisador é levado a comunicar sua produção por três motivos primordiais: na busca por reconhecimento pela comunidade

científica; para a proteção do conhecimento e de sua propriedade intelectual, e para servir como veículo de informação para o bem da humanidade e para compartilhar com os demais cientistas (PRICE, 1976). Entre os canais de comunicação da ciência, o periódico é considerado a fonte primária mais importante para a comunidade científica, e é através dele que a pesquisa é formalizada e torna-se de conhecimento público.

Esses periódicos científicos passaram a ser percebidos por pesquisadores profissionais como uma forma de assegurar a prioridade e conservação de contribuições passadas, como um acervo oficial que será utilizado no futuro por novos pesquisadores, garantindo a transmissão do conhecimento científico oficializado. Para lidar com o volume crescente de material, as publicações passaram a ser mais frequentes, se tornando periódicas e de publicação contínua, se especializando por áreas do conhecimento e temáticas específicas.

Quanto ao histórico do sistema de publicação, Barba (2003) cita um dos primeiros obstáculos enfrentados no processo de publicação científica: o ocultamento dos descobrimentos pelos investigadores por medo de plágio ou roubo filosófico. Para que um pesquisador torne públicas suas descobertas, é necessário que estas passem pelo sistema de publicação, que deve ser confiável e imparcial na avaliação das pesquisas, mitigando a prática de ocultamento das descobertas por parte do pesquisador.

Através da publicação, o pesquisador encontra uma motivação para comunicar suas descobertas: o reconhecimento por parte de outros pesquisadores e o prestígio atribuído a ele pelo resultado alcançado. Levando em consideração que o reconhecimento gera prestígio e que é mais fácil ser escutado e valorizado quem tem certo prestígio, o reconhecimento é, portanto, a principal explicação da avidez dos cientistas em fazer descobertas originais e em comunicá-las (BARBA, 2003). O reconhecimento é útil para o desenvolvimento da carreira individual, e em situações de recursos escassos, o prestígio facilitaria o acesso a recursos, instalações e até mesmo financiamentos para o pesquisador.

Para Latour e Woolgar (2013), os cientistas se interessam uns pelos outros não porque se veem obrigados pelo sistema a reconhecer as conquistas dos outros, mas porque cada um necessita do outro para aumentar a própria produção de informação. A noção de que o reconhecimento será obtido através da publicação de resultados faz com que, para Barba (2003), seja improvável um pesquisador não publicar suas descobertas formalmente, apesar das várias regras e normas envolvidas no processo de publicação científica. Neste ponto, o autor esclarece que, “se um cientista deixar algo interessante se perder nos canais informais, este corre um alto risco de perder boa parte do reconhecimento que poderia obter através de sua descoberta” (BARBA, 2003, p. 37, tradução nossa).

Dessa forma, o reconhecimento é, portanto, a valorização por outros da competência científica de um investigador. Para Latour e Woolgar (2013), a necessidade de reconhecimento está vinculada à continuidade do investigador como tal. O reconhecimento também pode ser visto como forma de admissão, e a publicação é o principal modo de alcançá-la. Sendo assim, esse entendimento do reconhecimento é uma das forças que sustentam o costume de publicar novas descobertas. Assim, se apresenta necessário atender a três sentidos agora apresentados do reconhecimento: como admissão na comunidade científica; como avaliação da capacidade de relação com outros colegas; e como validação das contribuições. A publicação formal prossegue sendo a única forma de alcançar o reconhecimento em qualquer desses sentidos.

O caminho para o reconhecimento é a consecução de resultados científicos que infere na capacidade de um pesquisador em fazer ciência e conseguir resultados. A importância dos resultados para a ciência é vista de duas perspectivas: segundo o grau de novidade e relevância; e segundo a transformação que este resultado produziu efetivamente no conhecimento científico. Algo obtido através de atividades científicas será um resultado científico quando é tido como novo e relevante para comunidade científica, o que completa o sentido de reconhecimento na ciência (BARBA, 2003).

As práticas para a publicação de descobertas em periódicos científicos, bem como os demais processos do sistema de avaliação da produção científica, até então veiculados de forma impressa, necessitaram se adaptar a uma nova forma de comunicação com o avanço das tecnologias de informação e comunicação. Entre 1980 e 1990, antes do surgimento da Web, o periódico eletrônico se desenvolveu se adaptando aos recursos e suportes emergentes, como a distribuição das edições de periódicos utilizando a tecnologia do *CD-ROM*; o armazenamento dos periódicos de maneira local em computadores, fazendo bom uso da tecnologia que possibilitava o salvamento do que era considerando, na época, um grande volume de dados; o acesso online através de redes locais e proprietárias, com o surgimento da internet (OLIVEIRA, 2008).

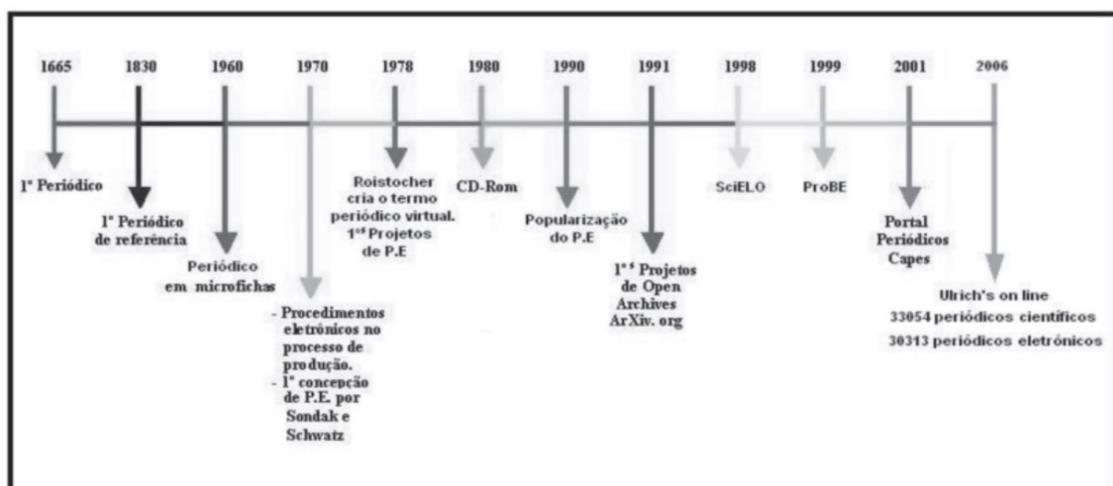
Com o avanço da internet e o surgimento da Web, a partir de 1990, a publicação científica em revistas eletrônicas representa a comunicação formal da ciência online e amplia o acesso deste conhecimento não apenas por cientistas, mas pela sociedade em geral, possibilitando também o acesso não apenas à informação, mas também ao documento que a contém. Cada vez mais, os computadores funcionam como canais para os meios de comunicação em massa e facilitam o acesso à informação disponível na web, incluindo as publicações científicas de revistas eletrônicas e, cada vez menos, edições de revistas

científicas impressas são publicadas. A consequência dessa migração do canal de comunicação impresso para o digital com disponibilidade online é o aumento do volume e do alcance das informações científicas em circulação (MEADOWS, 1999).

As barreiras geográficas foram rompidas e a colaboração entre cientistas situados em diferentes lugares no mundo se tornou possível, com facilidade e agilidade na troca de conhecimento. Além disso, outros agentes chegaram com este novo ambiente online, surgindo assim bases de dados que oferecem indicadores para o monitoramento e avaliação da qualidade da produção científica disponível.

Essa qualidade pode ser representada quantitativamente através da análise destes indicadores, que podem medir, monitorar e avaliar a qualidade a nível de periódicos ou de publicações científicas. Oliveira (2006) pontua cronologicamente (Figura 1) os momentos mais marcantes da evolução do periódico científico, desde o formato impresso até o eletrônico.

Figura 1 - Evolução do periódico científico ao longo do tempo



Fonte: Oliveira (2006, p. 46).

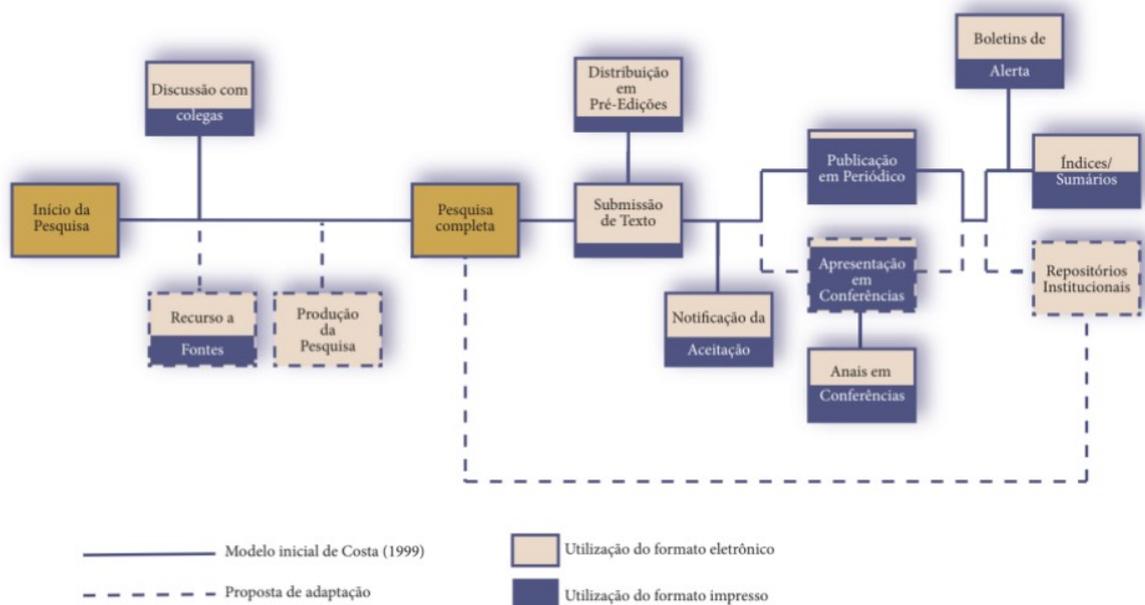
De acordo com Lancaster (1995), o desenvolvimento pelo qual as publicações passaram pode ser dividido em quatro etapas, que podem ser simultâneas: o uso de computadores na geração da publicação impressa tradicional; a distribuição da publicação por meio eletrônico, com as versões eletrônicas e impressas iguais; a publicação eletrônica com formato da impressa, mas agregando alguns diferenciais; e a geração de uma publicação completamente nova, que verdadeiramente explore as possibilidades do meio eletrônico (OLIVEIRA, 2008).

Essas etapas descrevem bem as variadas práticas de editores de periódicos científicos que ainda são adotadas atualmente, onde percebemos revistas eletrônicas com características

residuais da época em que eram apenas impressas, como a formatação do texto em duas colunas, por exemplo. Outras, ainda que tenham se adequado ao meio eletrônico no quesito formatação e visualização do conteúdo do documento, ainda não exploram ao máximo as possibilidades que o suporte eletrônico oferece, como o uso de recursos audiovisuais. Algumas revistas seguem publicando suas edições em versões impressas e eletrônicas, enquanto novas revistas nasceram no meio eletrônico e se adequam melhor ao uso de seus recursos. As práticas dos agentes envolvidos no processo de publicação de periódicos são variadas, mas independente destas, a importância e o valor científico do periódico fazem com que este permaneça inegável como o principal canal oficial de comunicação da ciência.

Existem diversos modelos de comunicação científica que foram adaptados ao longo do tempo e que permitem a visualização do fluxo da comunicação. O modelo de comunicação científica de Pinto e Costa (2018) para as áreas de Ciências Sociais e Humanas ilustrado na Figura 2 é baseado no modelo de Costa (1999), com adaptações necessárias para complementação e atualização, demonstrando ser o modelo que mais se aproxima da realidade (PINTO; COSTA, 2018).

Figura 2 - Modelo de comunicação científica para as comunidades de Ciências Sociais e Humanas



Fonte: Pinto e Costa (2018, p. 155).

No modelo acima é possível observar que as adaptações possuem relação com o avanço da tecnologia e seu impacto no processo de produção do conhecimento, onde as fontes de informação podem ser impressas ou eletrônicas, assim como o contato informal entre

pesquisadores, que passa a ser majoritariamente por meio eletrônico. Após a publicação em periódico impresso ou eletrônico, a pesquisa também pode ser depositada em repositórios de acesso aberto.

Embora o modelo de Pinto e Costa (2018) ainda apresente a publicação em periódicos sendo em sua maioria no formato impresso, com a Web surgem projetos de disponibilização dos periódicos em formato eletrônico e editoras comerciais interessadas neste formato de publicação online. Com isso, esperava-se que o preço das assinaturas de periódicos diminuísse, o que não ocorreu de fato.

Diante dessa situação e motivados pelo descontentamento com os altos preços praticados pelas editoras comerciais e pela demora entre o processo de submissão e publicação dos artigos, pesquisadores enxergaram o potencial da Web em facilitar o acesso ao conhecimento e a possibilidade de este ser distribuído online de forma gratuita. Assim surgiram os repositórios abertos, ou *open archives*, que permitem livre acesso a artigos e coleções de artigos online.

Além das questões relacionadas ao acesso aos periódicos científicos, a própria estrutura dos periódicos se desenvolveu para garantir credibilidade a eles. Quanto à estrutura dos periódicos, além das especificidades do seu suporte, é fundamental a definição de componentes, como o conselho editorial, política de acesso, periodicidade, avaliadores, foco e escopo, e recursos humanos qualificados para contribuir no processo de editoração científica, como na prestação de suporte técnico e operacional aos sistemas eletrônicos, editoração de texto, revisão e tradução. Ao falarmos da estrutura editorial de um periódico, estabelecer um conselho editorial é a garantia de perenidade para a revista, de modo que a existência da mesma não seja dependente unicamente da responsabilidade de um único editor. Para Trzesniak (2009), existem três requisitos para uma revista ser considerada científica:

A primeira e principal preocupação de um periódico científico deve ser levar aos seus leitores-pesquisadores (e pesquisadoras) o conhecimento novo e relevante dentro de sua área temática. As outras duas, sem precedência de uma sobre a outra, são fazê-lo a tempo e hora e perenizar-se, ou seja, existir para sempre. Somente atendendo a essas condições, a revista fará efetivamente jus ao adjetivo científica. Nenhuma delas pode faltar (TRZESNIAK, 2009, p. 88).

O conselho editorial deve ser presente nas tomadas de decisões, como uma retaguarda institucional na qual a revista pode se apoiar. Para assegurar a credibilidade de um periódico é também necessário a criação de um corpo editorial científico, constituído por pesquisadores especialistas e responsáveis pela seleção dos textos a serem publicados na revista, bem como

pela seleção dos avaliadores/pareceristas para executar a avaliação por pares (TRZESNIAK, 2009). Outras atividades necessárias para o bom funcionamento de um periódico científico envolvem recursos humanos diversos, reafirmando a importância de uma instituição que dê apoio ao periódico.

Em suma, a importância dos periódicos científicos reside em sua função de armazenar, preservar e disseminar as descobertas científicas. No processo de desenvolvimento das Tecnologias da Informação e Comunicação (TICs) e da Web, e com a migração dos periódicos científicos do formato impresso para o digital, os periódicos científicos enfrentaram desafios derivados do grande volume de publicações online, desordenadas na Web, provindas de diferentes sites e plataformas, de recuperação complexa e dependentes de indexadores.

A falta de padronização existente nos títulos de periódicos e nomes de autores (CAFÉ; BRÄSCHER, 2008, p. 68), por exemplo, dificulta a coleta e análise de dados bibliométricos, seja por pesquisadores ou por bases de dados indexadoras de conteúdo científico. De acordo com Rodrigues e Fachin,

São muitos os recursos que despontam como soluções para as informações científicas digitais. Ênfase se dá para as TICs que eclodiram nos últimos anos, possibilitando as diversas áreas do conhecimento: criar, organizar e circular periódicos, nas mais diferentes formas, padrões e suportes, promovendo um aumento exponencial de publicações, muitas das quais de forma isolada e que não permitem buscas integradas entre si, tornando a recuperação da informação ainda mais ineficaz. (RODRIGUES; FACHIN, 2010, p. 36).

Nesse cenário de difícil recuperação da informação científica, surgem modelos alternativos para organização de publicações científicas de modo coletivo em portais de periódicos e repositórios digitais, podendo estes serem temáticos ou institucionais. Um portal de periódicos científicos é uma estrutura online que funciona como agregador e índice (GARRIDO; RODRIGUES, 2010), com o objetivo de ajudar pesquisadores a encontrar informações específicas acerca da coleção de publicações científicas disponíveis nas revistas hospedadas. A vantagem da agregação de periódicos em um único portal reside justamente na possibilidade de acesso através de uma interface única que funciona como ponto de acesso direto aos serviços e informações disponibilizados, permitindo a filtragem das informações por tema, revista, e área de conhecimento.

Tendo em vista este panorama geral da evolução, estrutura e funcionamento do periódico científico, trata-se a seguir do cenário brasileiro.

2.1.1 Periódicos Científicos Brasileiros

No Brasil, há uma crescente preocupação em disponibilizar o acesso aos periódicos de maneira aberta e gratuita. De maneira geral, os periódicos brasileiros são publicados majoritariamente em acesso aberto, garantindo alta visibilidade e acessibilidade ao conteúdo científico.

Quanto à representatividade da produção científica brasileira, os periódicos brasileiros são responsáveis por mais de um terço dos artigos indexados internacionalmente com afiliação brasileira (PACKER, 2011). Já quanto à origem da produção acadêmica dos periódicos brasileiros, os programas de pós-graduação brasileiros são os mais importantes fornecedores de trabalhos de alta qualidade para publicação, sendo a principal missão dos periódicos brasileiros facilitar a disseminação da pesquisa científica do país (MOREIRA, 2015). Entretanto, como parte da literatura científica universal, os periódicos brasileiros também publicam trabalhos de pesquisadores estrangeiros, mas em menor proporção. Sendo assim, os periódicos brasileiros são representantes de parte importante da ciência desenvolvida no país.

Uma vez que a ciência brasileira é fortemente publicada nas revistas nacionais, os periódicos brasileiros possuem um papel importante no desenvolvimento da mesma. Entretanto, a cobertura da ciência brasileira em bases de dados bibliográficas internacionais como a *Web of Science* e *Scopus* ainda é baixa, limitando a representatividade da pesquisa nacional (MELO; TRINCA; MARICATO, 2021). Embora haja uma dependência dos métodos de medição e avaliação da qualidade de pesquisa das métricas provindas destas duas bases bibliométricas, vistas como imparciais e de representatividade global, Tennant (2020) identifica que a maioria das produções indexadas nestas bases são de países da América do Norte e Europa, e que elas parecem “discriminar formas diferentes de conhecimento; e, em particular, o que não vem da parte do mundo ocidental que fala inglês” (TENNANT, 2020, p. 2, tradução nossa).

Ainda assim, bases de dados como *Web of Science* e *Scopus* são utilizadas como parâmetro de qualidade para seleção de periódicos importantes e, aqueles que não estiverem indexados nestas bases, sofrem com o risco de serem desqualificados (MUGNAINI *et al.*, 2019). Devido à baixa indexação da produção brasileira nas bases de dados internacionais, existe uma dificuldade de publicação dessa produção em revistas internacionais, justamente pela produção não ser considerada internacional, fator que contribui para que a ciência brasileira seja, em grande parte, publicada em periódicos brasileiros. Em algumas áreas de

conhecimento, como Ciências Agrárias e da Saúde, Ciências Sociais Aplicadas, Humanidades e Linguística (MUGNAINI *et al.*, 2019), a ciência brasileira está concentrada em revistas brasileiras, o que faz com que gerar indicadores de revistas brasileiras seja útil na avaliação da ciência brasileira para diversas áreas.

Desse modo, os estudos métricos brasileiros ainda carecem de pesquisas que tomem como referência fontes de dados nacionais que possuam uma maior cobertura da produção científica brasileira, representando com maior fidelidade o cenário da ciência nacional. Mugnaini *et al.* (2019) afirmam que

A realidade das recentes transformações deve ser considerada para avanços dos estudos quantitativos da ciência, significando que estudos que utilizem fontes brasileiras devem primar por conhecer especificidades da produção científica nacional, não se limitando ao que oferecem a bases tradicionais, ainda que isso implique enfrentar problemas e limitações típicas de fontes de dados de natureza diversa. (Ibid, p. 3).

As fontes de dados diversas das fontes internacionais e comerciais podem, por um lado, prover dados que demandem uma maior dificuldade na descarga e tratamento, mas a relevância do uso destas bases reside em poder acessar os dados por completo, com a segurança de que estes dados representam de maneira satisfatória a produção científica nacional, ou seja, que representam boa parte da produção existente.

Trabalhando em conjunto com os periódicos brasileiros em busca da divulgação e disseminação da produção científica, surgem diversas iniciativas de divulgação das publicações científicas brasileiras, sendo a principal delas a *SciELO*². A *SciELO* é um repositório multidisciplinar para depósito, preservação e disseminação de artigos submetidos e aprovados para publicação ou já publicados em periódicos da Rede *SciELO*, contando com mais de 1.800 periódicos. Destes, 391 são brasileiros, totalizando mais de 450 mil documentos, sendo 276 mil de afiliação brasileira (*SCIELO ANALYTICS*, 2021).

O objetivo da *SciELO* é desenvolver uma metodologia comum para a preparação, armazenamento, disseminação e avaliação da produção científica em formato eletrônico (OLIVEIRA, 2008). A *SciELO* publica e indexa artigos completos para todos os periódicos presentes na base, e é uma iniciativa de acesso aberto valorizada na América Latina por oferecer dados de citações para todos os artigos da plataforma.

² SCIELO BRASIL. Scientific Electronic Library Online. **Lista de periódicos**. São Paulo: SciELO, c2021. Disponível em: <https://www.scielo.br/>. Acesso em: 5 dez. 2021.

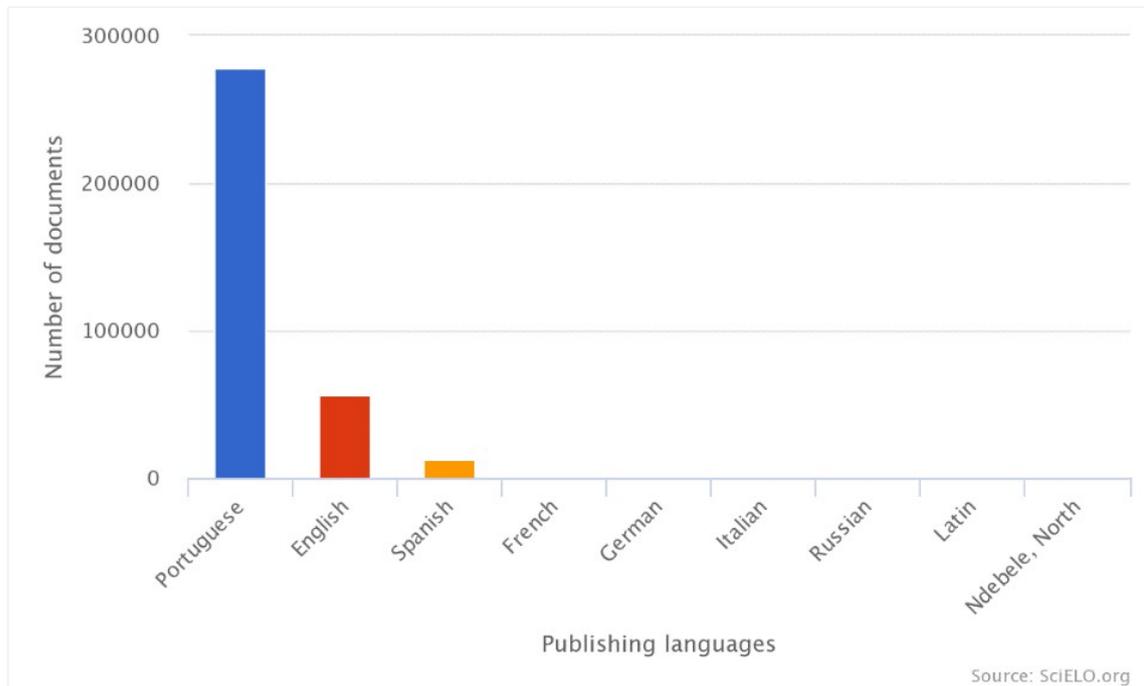
Além da *SciELO*, surge em 2000 o Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)³. Atualmente, o Portal da CAPES possui mais de 49 mil periódicos disponíveis com texto completo, e 455 bases de dados com conteúdo diversos, reunindo conteúdo científico de alta qualidade e se consolidando como “instrumento fundamental para as atividades de ensino e pesquisa no Brasil por disponibilizar acervo científico digital atualizado e com alto fator de impacto” (BRASIL, c2020).

A importância de se ter uma base de dados regional, como a *SciELO* e o Portal de Periódicos da CAPES, reside em capturar as citações de publicações não podem ser coletadas por outras bases de dados internacionais por conta do idioma usado na publicação. Isso porque a maioria da produção científica brasileira ainda é publicada no idioma nativo, o português, e as grandes bases de dados indexadoras de conteúdo internacional coletam dados de publicações em inglês.

Diversos periódicos brasileiros, com a preocupação de alcançar uma visibilidade internacional, têm como idioma oficial de publicação o inglês. Entretanto, as estatísticas da *SciELO* apontam que a grande maioria dos documentos científicos na base são publicados em português, como mostra a Figura 3. Isso impacta diretamente na representatividade da produção brasileira nas bases de dados internacionais, tanto em termos de produção quanto de citação.

³ BRASIL. Ministério da Educação. CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. **Periódicos**. Brasília: CAPES, c2020. Disponível em: <https://www-periodicos-capes-gov-br.ezl.periodicos.capes.gov.br/>. Acesso em: 5 dez. 2021.

Figura 3 - Distribuição dos documentos por idioma na base de dados *SciELO* (2021)



Fonte: *SciELO* Analytics (2021).

Além da sub-representatividade nas bases de dados internacionais, outra questão enfrentada por periódicos brasileiros é a aquisição de recursos. Com a forte política de publicação em acesso aberto e a gestão dos periódicos científicos brasileiros estar majoritariamente vinculada a instituições de ensino superior e seus programas de pós-graduação, os recursos são provindos de agências de fomento e recursos institucionais, que costumam ser escassos e distribuídos de acordo com critérios e seleções específicas baseados na qualidade das revistas. Dessa forma, periódicos brasileiros de maior qualidade recebem mais recursos para sua manutenção, e a disponibilização das métricas utilizadas na avaliação da qualidade do periódico contribui para o financiamento dele, além de contribuir para a disseminação do conteúdo do periódico em bases de dados indexadoras.

De acordo com Packer (2011), a presença dos periódicos brasileiros nas bases indexadoras bibliográficas internacionais é consistente com a produção científica indexada no Brasil, ainda que os números sejam inferiores quando comparados a produções científicas internacionais. Entretanto, o posicionamento brasileiro é superior à maioria dos países em desenvolvimento.

2.1.2 Periódicos de Acesso Aberto e sua indexação

Uma vez que a presente pesquisa analisa indicadores alométricos e bibliométricos para revistas de portais de periódicos de acesso aberto de universidades brasileiras, faz-se útil entender o que significa dizer que um periódico é de acesso aberto. Compreender o que é *Open Access* (OA) nos leva às três principais declarações que serviram de base para o surgimento dessa iniciativa: *Budapest Open Access Initiative* (BOAI, 2002), *Bethesda Statement on Open Access publishing* (2003) e *Berlin Declaration on Open Access to knowledge in the Sciences and Humanities* (2003).

A definição mais influente de *OA* vem da *BOAI* de 2002, que define *OA* como “tornar o conteúdo livre para leitura e reutilização, demandando a oportunidade para que os usuários do *OA* possam coletar artigos para indexação, passá-los em forma de dados para os softwares, ou usá-los para qualquer outro propósito legal” (BOAI, 2002). Outros pesquisadores preferem uma definição mais objetiva do que é *OA*, como simplesmente tornar a literatura acadêmica livre para ser lida online (WILLINSKY, 2003). Baseado no consenso entre as diversas conceituações para o *OA*, Piwowar *et al.* (2018) definem o acesso aberto como artigos livres para serem lidos online, seja no website da própria editora ou em um repositório de acesso aberto.

A principal razão para o interesse no acesso aberto é deixar o conteúdo aberto a todos sem restrições. Tendo em mente este conceito principal que a iniciativa de acesso aberto defende, seguimos para o entendimento dos tipos de acesso. Devido a diversas questões em debate sobre a clareza na identificação no tipo de acesso a um artigo, há um consenso limitado sobre a definição de todos os tipos de acesso aberto e até mesmo sobre a delimitação entre o acesso aberto e o acesso fechado (WHITE *et al.*, 2021). Entretanto, ao longo dos mais de quinze anos de pesquisas sobre o assunto, foram identificados seis principais tipos de acesso, descritos por White *et al.* (2021) conforme o Quadro 1:

Quadro 1 - Definição dos tipos de acesso aberto

OA Dourado	Versão publicada é de acesso aberto imediato. É cobrado a <i>APC</i> .
OA Híbrido	Publicação baseada em assinatura. A <i>APC</i> pode ser paga para tornar artigos individuais de acesso aberto.
OA Bronze	Livre para leitura no site da editora, mas o tipo de licença não é claro.
OA Verde	Somente acessível através de repositórios (redes sociais de universidades não estão inclusas). A versão da editora tem acesso pago.
Fechado	A versão publicada tem acesso pago.
OA Diamante	A versão publicada é de acesso aberto imediato. Recebe o status “Dourado” no <i>Unpaywall</i> ⁴ , mas o DOAJ ⁵ mostra que nenhuma <i>APC</i> é cobrada.

Fonte: White *et al.* (2021, p. 7, tradução nossa).

O acesso aberto busca remover barreiras de preço e de permissão ao conteúdo científico (SUBER, 2006) e, por meio de seus principais tipos, também chamados de vias de publicação, contribui para a tendência de retirar o controle da comunicação científica das editoras comerciais, que, visando o lucro, impõem políticas editoriais com restrição de acesso às publicações. Com o surgimento dos periódicos de acesso aberto, a disponibilização da produção científica passa a ser sem barreiras de taxas e assinaturas, forma de divulgação da ciência que pode se enquadrar nos tipos de acesso dourado ou diamante, a depender da política de disponibilização do conteúdo definido pelo periódico.

Os tipos de acesso afetam diretamente a publicação científica de determinados países, com o movimento de acesso aberto trazendo benefícios para países em desenvolvimento. De fato, o acesso aberto fortalece o sistema científico de países que não atingem os critérios internacionais de qualidade (AHMED, 2007), critérios estes que são impostos através de um sistema de avaliação que prioriza publicações indexadas em bases de dados como *Web of Science* e *Scopus*. Ahmed (2007) cita o exemplo do continente africano, afirmando que

[...] O acesso aberto fortalecerá as comunidades científicas da África, seus sistemas nacionais de ciência e, principalmente, irá expor essas fontes, sistemas e métodos de conhecimento local, que ainda precisam encontrar seu caminho no mercado internacional de ideias, muitas vezes porque se acredita que não atendem aos padrões internacionais. O acesso aberto também irá melhorar o sistema científico global, demonstrando estratégias e técnicas alternativas às que já existem, mas que não fazem parte dos sistemas científicos convencionais, mas que também produzem os mesmos resultados que aquelas pesquisas executadas com métodos padrão. (AHMED, 2007, p. 340)

⁴ O *Unpaywall* é uma base de dados aberta que disponibiliza o status de *OA* para mais de 30 milhões de artigos científicos (ELSE, 2018).

⁵ O DOAJ é um diretório de periódicos de acesso aberto com mais de 17 mil periódicos indexados, e que possibilita a consulta das diretrizes e licenças de acesso aberto de cada um deles. (DOAJ, c2020).

Desse modo, a disponibilização de publicações científicas em acesso aberto, em especial as de acesso diamante, permite que a ciência de países às margens do sistema de avaliação de qualidade internacional de publicações científicas e sem recursos para manterem assinaturas e pagamento de taxas alcancem visibilidade, com um fluxo de conhecimento de acesso livre sendo compartilhado a nível global, o que beneficia diretamente o avanço da própria ciência (AHMED, 2007).

Se tratando de artigos publicados em portais de periódicos brasileiros de acesso aberto, estes são classificados como de acesso Diamante, visto que os periódicos destes portais disponibilizam os artigos em acesso aberto no momento da publicação, não cobrando taxas de processamento do artigo (do inglês, *Article Processing Charges*, ou *APCs*) ao longo do processo. As *APCs* surgiram como uma forma de adequação dos modelos de negócio das editoras comerciais, redirecionando para o autor o pagamento dos custos no processamento dos artigos a serem disponibilizados em acesso aberto (ou seja, aqueles artigos que não geram ganhos para o periódico com o pagamento de assinatura para acessá-lo).

Alguns estudos (DEMETER; ISTRATII, 2020; SOLOMON; BJORK, 2016) identificaram que as *APCs* são uma questão preocupante a ser levada em consideração nos estudos de acesso aberto, uma vez que não há um consenso de como os custos que justifiquem o valor das *APCs* possam ser estimados (WHITE *et al.*, 2021). Istratii e Demeter (2020) identificaram que periódicos de alto fator de impacto tendem a ter *APCs* mais altas também. A preocupação com o valor das *APCs* se deve ao fato de que a migração para um modelo de acesso aberto em que as *APCs* sejam cobradas dos autores e das instituições de fomento e de apoio à pesquisa pode agravar diferenças econômicas entre pesquisadores de países desenvolvidos e periféricos (EJIKEME; EZEMA, 2019).

Pesquisas como a de Demeter (2020) já demonstraram que a academia global é caracterizada por um desbalanço significativo entre o Norte Global e o Sul Global⁶, principalmente em publicações em revistas de alto fator de impacto. Em sua pesquisa, Demeter (2020) demonstrou que revistas de alto fator de impacto da área de Estudos de Desenvolvimento são publicadas integralmente em países ocidentais considerados economicamente desenvolvidos, com a maior parte pertencendo ao Reino Unido (51%), aos EUA (37%), o restante à Europa Ocidental (12%). Essa tendência, que se fortalece com a

⁶ Demeter e Istratii (2021) descrevem que o Norte Global inclui os EUA, Canadá, Austrália e Nova Zelândia, Reino Unido, Europa Ocidental, Israel e países asiáticos como o Japão, Hong Kong, Cingapura, Coreia do Sul e Taiwan. O Sul Global inclui a América Latina (incluindo América do Sul e Central), Europa Oriental, Ásia (exceto os países incluídos no Norte Global), Oriente Médio e África (e partes da Oceania, exceto Austrália e Nova Zelândia).

conexão entre a publicação e o fomento à pesquisa, uma vez que os custos da publicação em acesso aberto são cobertos diretamente por agências de fomento à pesquisa (DEMETER; ISTRATII, 2020), reforça a dificuldade de publicação de pesquisas de países periféricos em revistas de qualidade, uma vez que os pesquisadores destes países enfrentam desafios, como dificuldade de conseguir financiamento, financiamento insuficiente e até mesmo falta de compromisso por parte das instituições de fomento (EJKEME; EZEMA, 2019).

Entretanto, os periódicos brasileiros se destacam “pela ampla visibilidade que têm na web e pelo uso dos seus artigos, medido pelo número de downloads dos arquivos HTML e PDF” (PACKER, 2014, p. 312). O incentivo e adoção do acesso aberto pelo SciELO contribuiu com o desenvolvimento da capacidade nacional de indexação de periódicos e suas publicações, maximizando a presença e a comunicação dos seus conteúdos na web (PACKER, 2014).

Os portais de periódicos, poderosos agregadores de conteúdo científico, quando vistos dentro do contexto brasileiro, são predominantemente de acesso aberto, sendo o Brasil o maior consumidor do sistema de periódicos de acesso aberto *OJS* na América Latina, e segundo maior do mundo, atrás somente da Indonésia (PKP, 2021a). Somente em 2018 o Brasil possuía 115 portais de periódicos de acesso aberto vinculados a universidades (SANTOS, 2019), reunindo periódicos de diferentes áreas de conhecimento e financiados por estas instituições. Desse modo, a adoção desta estrutura de agregação de periódicos deve ser estimulada com vistas ao suporte e expansão do acesso aberto e, conseqüentemente, da ciência brasileira publicada em periódicos nacionais.

Outra questão de interesse em estudos sobre acesso aberto é a *Open Access Citation Advantage* (*OACA*) ou, Vantagem de Citação de Acesso Aberto. Essa é uma hipótese que afirma que publicações disponibilizadas em acesso aberto tendem a ser mais citadas do que aquelas em acesso restrito (PIWOWAR *et al.*, 2018), que são todas aquelas acessíveis somente mediante o pagamento de uma assinatura ou taxa. Diversos estudos compararam citações de publicações de acesso aberto com as de acesso restrito, e quase sempre contam com correlações positivas entre o acesso aberto e o número de citações (WHITE *et al.*, 2021). Alguns identificavam um ligeiro aumento nas citações para acesso aberto (MCCABE; SNYDER, 2014), outros (ARCHAMBAULT *et al.*, 2014) chegam a descrever uma *OACA* de 40% para uma amostra de mais de um milhão de artigos.

Constatado que pesquisas publicadas em acesso aberto recebem, de maneira geral, um maior número de citações do que as de acesso restrito, e que integrado a este contexto, com a evolução da tecnologia e mídias sociais, as publicações científicas estão sendo cada vez mais

compartilhadas nas redes sociais, surge uma hipótese similar à *OACA*, porém do ponto de vista altmétrico. Estudos da recém proposta *Open Access Altmetrics Advantage (OAAA)* ou, Vantagem Altmétrica de Acesso Aberto, investigam se publicações em acesso aberto possuem maiores indicadores altmétricos do que as de acesso restrito (CINTRA; FURNIVAL; MILANEZ, 2017; HOLMBERG *et al.*, 2020).

Vários estudos comprovaram a *OAAA* a partir de perspectivas distintas. Adie (2014) analisou a atenção social de todas as publicações da revista *Nature* e identificou que artigos de acesso aberto parecem gerar significativamente mais *tweets* e atrair mais leitores no *Mendeley* que artigos pagos. Cintra, Furnival e Milanez (2017) identificaram uma vantagem no número potencial de citações e de menções na web social que as publicações recebem, e mostraram que essas vantagens se mantêm até o oitavo ano após a publicação do artigo.

Embora, por definição, periódicos de acesso aberto sejam mais acessíveis e, conseqüentemente, devam ser usados mais frequentemente do que periódicos de assinatura, sua visibilidade pode ser medida somente depois que eles são indexados em bases de dados bibliográficas e de citação, como *Scopus* e *Web of Science* (JOKIC; MERVAR; MATELJAN, 2018).

Uma parte significativa dos resultados de pesquisa publicados em periódicos de acesso aberto que não são indexados em bases de dados disponíveis publicamente, especialmente naquelas de maior prestígio, acabam não sendo percebidos pela audiência internacional, principalmente se eles não são escritos em uma das línguas mais utilizadas mundialmente, como o inglês. Porém, a visibilidade destes periódicos pode ser acessada através de indicadores altmétricos (WILSON, 2016), que mostram a usabilidade através de visitas e downloads, por exemplo, ou na correlação destes com o número de citações.

Sendo abordados aqui os principais tópicos para o entendimento do acesso aberto, o escopo de estudo de portais de periódicos de acesso aberto desta pesquisa é reforçado pelas hipóteses *OACA* e *OAAA*, em busca de que as análises dos indicadores altmétricos e bibliométricos apresentem a possibilidade de uma maior cobertura e maiores números possíveis para os periódicos dos portais de acesso aberto, visando um cenário mais completo e possibilitando uma análise fidedigna dos indicadores.

Os portais de periódicos de acesso aberto são uma estrutura de comunicação científica que têm à sua disponibilidade recursos informacionais gratuitos, como as plataformas de editoração em software livre, de código aberto. Porém, considerando que o acesso aberto significa que as publicações científicas se tornarão disponíveis abertas e gratuitamente

imediatamente após sua publicação, sem nenhum tipo de restrição ou cobrança de acesso, é de se esperar que esse sistema não se sustente sem fomento. Rodrigues e Fachin ressaltam que

A organização de vários periódicos, de uma mesma instituição em um Portal, requer diversas ações integradas como: estrutura organizacional para viabilizar as ações de migração para a plataforma adotada; alocação de recursos para custos associados; definição da responsabilidade institucional para com o grupo de periódicos, que tende a se configurar como uma “meta-editora” (RODRIGUES; FACHIN, 2010, p. 37).

Na estrutura de um portal de periódicos de acesso aberto, a garantia da confiabilidade das informações científicas publicadas segue sendo garantida pela avaliação pelos pares, da mesma forma que em periódicos científicos, impressos ou eletrônicos, publicados em plataformas independentes. Portanto, o reconhecimento que o pesquisador espera obter com sua descoberta permanece seguindo as mesmas premissas que em periódicos estruturados de maneira individual. Rodrigues e Fachin ressaltam que:

Esses portais intervêm em duas questões estratégicas: contribuem para o aumento da visibilidade e do valor público das instituições, servindo como indicador tangível da sua qualidade, e contribuem para o sistema de comunicação científica, expandindo o acesso aos resultados da investigação e assumindo o apoio aos editores e a responsabilidade da disseminação e preservação dos periódicos da instituição. (RODRIGUES E FACHIN, 2010, p. 38).

Uma vez que a maioria dos estudos que analisam as hipóteses *OAAA* e *OACA* demonstraram a vantagem do acesso aberto resultando em maiores indicadores altmétricos e bibliométricos para as publicações científicas, essa premissa, teoricamente, também se aplica a portais de periódicos que sejam de acesso aberto.

Estudos como o de Wang *et al.* (2014) e Liu *et al.* (2013) analisaram as coberturas e correlações entre indicadores altmétricos e bibliométricos para publicações das revistas do portal da editora PLOS⁷, que reúne em seu portal revistas de acesso aberto de diversas áreas de conhecimento e com foco especial nas áreas de Ciências e Medicina. Wang *et al.* (2014) justificam a análise de publicações do portal da PLOS com dois argumentos: devido a este prover métricas a nível de artigo detalhadas, e para cada publicação individualmente; e pelo fato de que os dados provindos de um portal de periódicos de acesso aberto eliminam o viés entre dados de acesso aberto e fechado, ou seja, previnem uma análise desproporcional entre os dados que seria causada pela diferença no tipo de acesso às publicações.

⁷ Disponível em: <https://plos.org/>. Acesso em: 15 dez. 2021.

Os portais de periódicos de acesso aberto são uma fonte interessante de indicadores. O *OJS*, por exemplo, provê indicadores altmétricos de acesso à página de resumo das publicações, acesso aos diversos tipos de arquivos da publicação e números de downloads efetuados. Através de *plugins* de contagem de citações, os portais de periódicos também podem fornecer métricas para a composição de indicadores bibliométricos. Outros dados detalhados sobre as publicações também estão disponíveis através de relatórios diversos, como contagem de usos no padrão COUNTER (c2021) para mensurar o acesso a cada periódico em determinados períodos, e metadados de cada publicação.

Além dos indicadores que podem ser extraídos a partir desses relatórios e *plugins*, o *OJS* permite a extração de metadados através do protocolo *Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)*⁸, sendo este um mecanismo facilitador da interoperabilidade entre repositórios digitais. Os metadados, ainda que possam ter problemas de preenchimento e estejam sujeitos a falhas humanas neste processo, são usualmente inseridos no *OJS* ao longo do fluxo editorial de uma publicação, o que permite uma indexação mais eficiente do periódico nas diversas bases de dados.

Marchiori *et al.* (2018) geraram, a partir de metadados extraídos pelo protocolo *OAI-PMH*, sete indicadores determinados apenas por metadados do *OJS* para todas as edições publicadas no periódico de acesso aberto brasileiro *AtoZ*, como: autores por artigo; colaboração internacional; autores mais referenciados; endogenia; e, palavras-chave. Através destes indicadores os autores puderam identificar diversos pontos de atenção a serem considerados pelos gestores do periódico estudado,

[...] tais como: a redução da endogenia; a captação de material que amplie a representatividade na distribuição institucional e geográfica dos autores [...]; e a verificação cuidadosa dos métodos de pesquisa utilizados pelos autores das submissões em relação à proposta editorial do periódico. Resultados obtidos na análise das palavras-chave revelam que o periódico está publicando alguns temas emergentes, tais como redes sociais, jogos educativos e inteligência artificial. (MARCHIORI *et al.*, 2018, p. 334)

Esse tipo de resultado demonstra que a análise dos indicadores das publicações beneficia autores e periódicos não apenas através da análise direta de citações, mas também possibilita uma análise de falhas e oportunidades às quais os gestores dos periódicos podem se adequar e utilizar para melhorar a qualidade do periódico, em busca de conceder maior visibilidade e prestígio aos trabalhos científicos publicados e à própria revista. O fato de as fontes dos dados utilizadas no estudo serem de acesso aberto “é um ponto positivo e

⁸ Disponível em: <https://www.openarchives.org/pmh/>. Acesso em: 20 nov. 2021.

estimulante para a continuidade dos estudos. Contudo, a qualidade dessas fontes de dados pode limitar o espectro dos indicadores sugeridos” (MARCHIORI *et al.*, 2018, p. 335).

Porém, ressaltamos a importância dos gestores de periódicos manterem os dados normalizados e organizados, com vistas a melhorar sua qualidade e, conseqüentemente, facilitar a análise e tratamento dos mesmos, especialmente quando a análise agrega um volume de dados maior e provindos de revistas com diferentes normas e padrões, como é o caso dos portais de periódicos. Neste caso, um trabalho conjunto entre editores, instituição gestora, biblioteca e demais agentes envolvidos no processo de publicação científica devem ser considerados para alcançar a qualidade dos dados que serão usados na construção dos indicadores, reduzindo a complexidade nas etapas de pré-tratamento dos dados e de verificação de inconsistências.

A aplicação dos indicadores aos portais de periódicos de acesso aberto permite mensurar o impacto científico e a qualidade dos periódicos. Entre os principais indicadores existentes para avaliação de qualidade utilizados pelas bases indexadoras de periódicos científicos, encontram-se os indicadores bibliométricos baseados em citação, que são essencialmente quantitativos (AVENA; BARBOSA, 2017). As métricas de citações por item de publicação, citações totais que uma publicação recebeu nos últimos anos, e a imediatez (velocidade que um artigo de um periódico é citado após ser publicado), são indicadores bibliométricos a nível de artigo ou revista que são aplicados na avaliação de periódicos de maneira individual ou agrupados.

A estrutura em si de um portal de periódicos não interfere na forma de avaliação da qualidade de publicações científicas, porém pode influenciar na cobertura dos indicadores extraídos de diferentes bases de dados indexadoras, a depender de fatores derivados das ações estratégicas das editoras gestoras de cada portal. Por exemplo, uma instituição que funcione como “meta-editora” (RODRIGUES; FACHIN, 2008) de um portal de periódicos pode ter, em seu planejamento estratégico de atividades, a responsabilidade de fornecer para as bases indexadoras os metadados dos periódicos hospedados no portal.

A institucionalização de um portal de periódicos também pode afetar a obtenção de DOI para as publicações, identificador que é utilizado por diversas bases indexadoras para rastrear as atividades de uma publicação, como acessos e citações, e gerar métricas para indicadores. Se nem todos os artigos científicos possuem DOI ou há problemas técnicos relacionados a este identificador, a cobertura de indicadores provenientes de bases que utilizam o DOI para indexação de publicações será mais baixa (RODRIGUES; FACHIN, 2008). Rodrigues e Fachin afirmam que, no processo de institucionalização de um portal,

as alternativas são: o portal se tornar uma “meta-editora” com política, estrutura e orçamento próprios; ou estar vinculado à biblioteca; ou estar vinculado à editora; ou ainda opções mistas. Qualquer das alternativas implica em ajustes nas políticas editoriais e de informação da instituição, com repercussão em todas as revistas, na biblioteca e na editora. (RODRIGUES; FACHIN, 2008, p. 5).

Outro ponto importante de se ressaltar na análise de indicadores aplicados a portais de periódicos é que o fato de periódicos estarem estruturados em um mesmo portal não garante que a cobertura dos indicadores seja a mesma para todos eles. Avena e Barbosa (2017) selecionaram sete periódicos brasileiros da área de Enfermagem indexados no portal de periódicos da *SciELO*, que provê acesso aberto aos conteúdos de periódicos científicos, com o objetivo de analisar os indicadores bibliométricos dos periódicos nacionais e internacionais na área de Enfermagem, sob a ótica das bases indexadoras. Os indicadores coletados são os utilizados pelas bases *Web of Science*, *Scopus*, *SciELO*, *CAPES*, *CUIDEN* e *Google Scholar* na avaliação da qualidade de periódicos, todos baseados em citação. As autoras constataram uma desigualdade entre estes indicadores bibliométricos analisados, e “mesmo os que utilizam a mesma metodologia, de citações ao longo de existência do periódico (índice H), se mostram incomparáveis justamente porque a cobertura dos periódicos nas diversas bases é diferente”. (AVENA; BARBOSA, 2017, p. 5).

Dito isso, o estudo das coberturas dos indicadores se mostra relevante para publicações de portais de periódicos da mesma maneira que para periódicos de editoras diferentes. No caso de portais de periódicos de acesso aberto, acredita-se que a cobertura de indicadores nas diferentes revistas reflitam os esforços dos editores, da instituição e da biblioteca responsáveis pela gestão do portal e das revistas hospedadas em ampliar a visibilidade e o impacto científico e social das publicações.

Como já mencionado anteriormente, o fomento necessário para que as revistas científicas brasileiras de acesso aberto mantenham suas atividades é concedido de acordo com uma série de critérios de avaliação da qualidade e do potencial destas revistas. São fatores como a qualidade das publicações, indexação nas bases de dados, e impacto científico que são levados em consideração no momento da seleção de revistas a serem beneficiadas com fomento. Inclusive, Avena e Barbosa (2017) identificam que são indicadores bibliométricos de citação como o Fator de Impacto (FI) e o *SCImago Journal Ranking (SJR)* que norteiam a avaliação de qualidade dos periódicos para concessão de fomento. No entanto, acredita-se que esses indicadores são insuficientes para se analisar a ciência publicada em periódicos brasileiros, que em sua maioria não estão indexados nestas bases de dados internacionais.

Os indicadores de acesso aos portais de periódicos também podem ser úteis para um conjunto de análises. Porém, de acordo com Aguillo *et al.* (2010), ainda não há um padrão para que tais indicadores sejam combinados e comparados entre si, uma vez que os conceitos podem ter significados diferentes. Por exemplo, portais de periódicos de acesso aberto comumente divulgam relatórios com estatísticas de visitantes e downloads, mas esses dados correspondem a ações diferentes, e devem ser interpretados de maneiras diferentes. Ainda assim, as possibilidades de aplicação destes indicadores de uso revelam novas formas de se analisar a relevância do portal de periódicos, a visibilidade dos seus periódicos e o impacto social de suas publicações científicas.

Levando em consideração que indicadores de acesso revelam quantitativos de visitas ao portal, número de visitantes, acessos à página de resumo, downloads de publicações, é necessário compreender o impacto que estes indicadores permitem medir. Este impacto pode estar alcançando a comunidade em geral, uma vez que os usuários contabilizados através destes indicadores não são, necessariamente, pesquisadores que virão a potencialmente citar as publicações acessadas em suas pesquisas.

Esforços no sentido de combinar indicadores baseados em citação e em uso vêm sendo feitos por diversos estudos desde antes do surgimento da altmetria. Em 2009, Watson já comparava citações e downloads para artigos individuais, concluindo que as estatísticas de download fornecem um indicador útil para prever eventuais citações com dois anos de antecedência, além de contribuírem no entendimento da relevância e do valor de artigos individualmente.

Aguillo *et al.* (2010) indicaram haver a necessidade de criação de medidas para quantificar repositórios de acesso aberto, o que se aplica aos portais de periódicos. Os autores justificam que essas medidas reforçariam as iniciativas do acesso aberto, promovendo o depósito de documentos por pesquisadores; e para prover dados empíricos sobre essa estrutura de comunicação científica que são os repositórios, para futuras análises comparativas. Com isso, Aguillo *et al.* (2010) propuseram o *Ranking Web of Repositories*, introduzindo os indicadores desenvolvidos para medir a atividade e visibilidade dos repositórios.

Para construir o *ranking*, Aguillo *et al.* (2010) atribuíram pesos aos indicadores utilizados, adotando um modelo baseado no FI, com pesos iguais para a atividade e impacto das publicações. De modo geral, no resultado obtido pelo *ranking*, a consistência entre os indicadores provindos da web (*Google Scholar*, *PDF*, links externos e número de páginas) é baixa, uma vez que os top repositórios rankeados de acordo com um indicador não são tão bem colocados quando rankeados por outros. Mais uma vez, isso indica uma diferença de

cobertura dos indicadores e demonstra o potencial de complementação entre eles na análise de publicações acadêmicas em portais de periódicos.

Um dos problemas mencionados por Aguillo *et al.* (2010) que pode justificar a baixa consistência entre os indicadores provenientes da web é o que os autores chamam de invisibilidade: a incapacidade dos rastreadores dos motores de busca para coletar dados, devido a barreiras criadas na construção das bases de dados. Este é um problema sério que deve ser levado em consideração pelos editores e instituições no momento da implantação de um portal de periódicos e na configuração das permissões de acesso e descoberta dos itens neles, uma vez que além de impactar na cobertura dos indicadores, pode restringir a descoberta das publicações pelos usuários a partir de outras fontes online, como as próprias ferramentas de busca e as bases indexadoras.

É importante, portanto, que os portais levem em consideração a indexação e posicionamento do portal nas ferramentas de busca como o *Google*, *Yahoo* e *Bing*. Além disso, Aguillo *et al.* (2010) sugerem, como forma de otimizar a coleta de dados pelos buscadores, o uso de domínios e subdomínios independentes como uma solução técnica para evitar barreiras aos robôs de coleta de dados, bem como a atribuição de valor semântico aos metadados das publicações.

No fim das contas, a análise da cobertura dos indicadores para revistas e publicações de um mesmo portal de periódicos pode revelar onde estão as possíveis falhas na divulgação dele, assim como uma possível discrepância nas coberturas de uma mesma fonte para revistas diferentes pode representar ações estratégicas isoladas dos editores que refletem positivamente nos indicadores.

Os indicadores de portais de periódicos demonstram também um grande potencial para analisar a correlação entre indicadores altmétricos e bibliométricos. Wang *et al.* (2014), ao estudarem as correlações dos indicadores para sete revistas do portal de periódicos de acesso aberto PLOS identificaram que, de modo geral, as visualizações de artigos se correlacionam bem com a citação. Ou seja, a correlação de Spearman varia de moderada a forte tanto na análise por faixa temporal quanto na análise por revista.

No entanto, diferentes tipos de visualização de artigos têm diferentes níveis de correlação com a citação, sendo que o download do arquivo do artigo em formato *PDF* (*Portable Document Format*) se correlaciona com a citação de forma mais significativa. Wang *et al.* (2014) também encontraram correlações moderadas entre a atenção social, medida através do indicador altmétrico extraído do *Mendeley*, e as citações. Os autores

também ressaltam a atenção social como uma métrica importante que permite identificar atividades em uma publicação que podem levar a uma citação subsequente.

Os indicadores de portais de periódicos de acesso aberto possuem, finalmente, a vantagem de transparência, quando comparados aos indicadores de portais de periódicos de assinatura e acesso restrito. Os indicadores extraídos de portais de periódicos de acesso aberto permitem aos editores, instituições e agentes científicos visualizar os dados e extraí-los diretamente do sistema, garantindo que os dados sejam precisos e seguros. Quando os indicadores são fornecidos por periódicos e portais de acesso restrito, a confiabilidade dos dados recai exclusivamente no prestígio e reputação da editora responsável por prover os dados, uma vez que não há transparência na forma com que são gerados. Indicadores extraídos de portais de acesso aberto podem trazer novas percepções sobre o impacto da ciência não indexada em grandes bases de dados comerciais.

2.2 INDICADORES E MÉTRICAS DE PRODUÇÃO CIENTÍFICA

A avaliação e monitoramento da atividade científica não é algo recente, e são feitas através de indicadores obtidos a partir das produções científicas, sobretudo publicações. O interesse na medição, monitoramento e avaliação dessas atividades científicas surge no século XX, quando começam a se sobressair estudos de pesquisadores como Lotka, Bradford, e Eugene Galfield, que trouxeram as primeiras contribuições para a edificação das áreas de bibliometria e cientometria (MARICATO; MARTINS, 2017).

Precisamente em 1923, surge então a área de estudos métricos da informação, uma área atualmente bastante consolidada dentro da Ciência da Informação e que possui variações nas metodologias de cada métrica deste grande campo. De acordo com Sutcliffe (1992), as três métricas clássicas da informação são:

- Bibliometria: estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada, desenvolvendo modelos e medidas matemáticas, com o objetivo de elaborar previsões e apoiar tomadas de decisão.
- Cientometria: estudo dos aspectos quantitativos da ciência e da pesquisa científica. Inclui atividades relacionadas à publicação e se sobrepõe à Bibliometria.
- Infometria: estudo da produção, disseminação e uso de todo tipo de informação registrada, independentemente do formato ou modo como é gerada. Considera as necessidades e usos da informação para qualquer atividade, seja proveniente de atividade intelectual ou não.

Para Maricato (2010), baseando-se no conjunto de métodos e técnicas abrangidas pelos estudos métricos da informação, a bibliometria possui caráter central e integrador, podendo ser considerada a “disciplina mãe”, apresentando relação interdisciplinar com as demais, que, por sua vez, possuem suas particularidades e aplicações próprias. O uso de indicadores e métricas de produção científica é uma prática comum na avaliação da qualidade e desempenho de pesquisa.

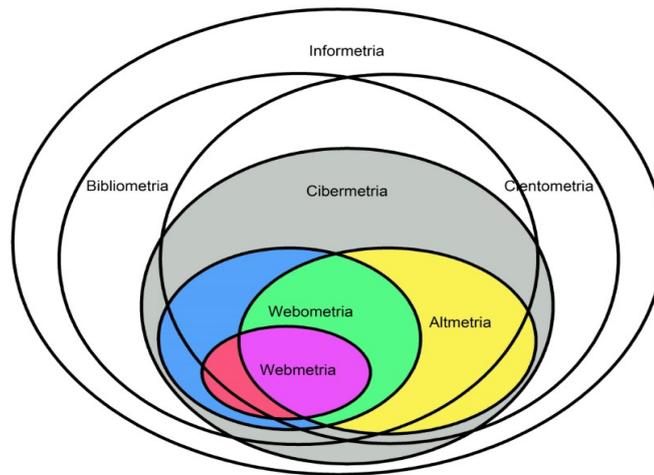
Hood e Wilson (2001, p. 309) afirmam que, como os interesses dos pesquisadores no campo das métricas da informação se voltam para a medição de páginas da Web ou sites, novos termos e métricas foram sendo criados. Sendo assim, com o avanço da tecnologia e da Web, diversos novos tipos de estudos métricos da informação emergiram, cabendo a definição de três deles:

- **Cibermetria:** o estudo de aspectos quantitativos da construção e uso de recursos informacionais, estruturas e tecnologias da Web a partir de abordagens bibliométricas e informétricas (BJÖRNEBORN; INGWERSEN, 2004, p. 12).
- **Webometria:** o estudo de aspectos quantitativos da construção e uso de recursos informacionais, estruturas e tecnologias da Web a partir de abordagens bibliométricas e informétricas (BJÖRNEBORN; INGWERSEN, 2004). Faz-se a medição de recursos, acesso, e utilidade da informação publicada online, bem como a ligação entre sites na web.
- **Altmétrie:** área que se ocupa do estudo e do uso de métricas de impacto acadêmico baseadas nas atividades realizadas em ambientes *online* (PRIEM, 2014). Faz uso de indicadores baseados em mídias sociais para quantificar o impacto social da informação acadêmica (PRIEM *et al.*, 2011).

Observe que, enquanto a webometria se restringe a estudos que possuem a web como suporte, como sites, links e outras citações textuais, a cibermetria é mais ampla, abrangendo outras métricas da internet e do ciberespaço, como ambientes virtuais de interação social e fóruns de discussão online, e desse modo, a cibermetria englobaria também a webometria. Já a altmetria engloba estudos focados nas atividades online que acontecem nas mídias sociais, um nicho mais específico que os campos de cibermetria e webometria, porém não contemplado totalmente pelas abordagens anteriores.

Gouveia (2013) elaborou um esquema de relações entre diversos campos de estudos métricos da informação, de modo a facilitar a compreensão e visualização das suas intersecções, conforme ilustra a Figura 4.

Figura 4 - Esquema de relações entre os campos da Informetria, Bibliometria, Cientometria, Cibermetria, Webometria, Webmetria e Altmtria



Fonte: Gouveia (2013).

Gouveia (2013) afirma que o campo da altmetria se situa dentro da interseção da cientometria com a cibermetria e a webometria, com sobreposições também com a bibliometria. Dessa forma, observamos que os diversos campos de estudos métricos da informação possuem suas similaridades e diferenças, de acordo com o objeto estudado. Porém, apesar da variedade de métricas utilizadas e das diversas terminologias e discussões ao redor da definição de cada um deles, alguns se destacaram e se tornaram métodos tradicionais para quantificar a informação, em especial, a informação registrada em forma de conhecimento científico.

As métricas tradicionais são todas as aquelas consolidadas e amplamente utilizadas para a avaliação da ciência. A bibliometria se tornou a métrica tradicional para medição do impacto acadêmico, e com ela surgiram diversos indicadores utilizados mundialmente de maneira bem difundida na avaliação da qualidade de revistas e artigos científicos, especialmente através das citações.

As fontes destes indicadores bibliométricos tradicionais são em sua maioria as bases de dados que indexam revistas e artigos científicos, abrangendo uma gama variada de bases consolidadas, como *Scopus*, *SciELO*, *Medical Literature Analysis and Retrieval System Online (MEDLINE)*, *Library & Information Science Abstracts (LISA)*, e *Web of Science (WoS)*. Essas bases de dados são construídas de modo a prover o usuário com filtros e mecanismos de busca poderosos, permitindo a exportação dos dados de publicações indexadas. É exatamente através destas funcionalidades presentes na interface ou *Application Programming Interface (API)* das bases que é possível coletar os dados que permitirão medir

alguma faceta da ciência, como por exemplo o número de citações de um artigo, base para a criação de indicadores bibliométricos como *Science Citation Index (SCI)*, índice-h e *Journal Impact Factor (JIF)*.

2.3 INDICADORES DE QUALIDADE DE PERIÓDICOS CIENTÍFICOS

Comumente pesquisadores buscam publicar seus trabalhos em revistas de maior qualidade, de maneira a agregar prestígio ao seu trabalho. Diversos estudos abordam os critérios de avaliação da qualidade de revistas científicas, tanto brasileiras quanto internacionais.

No âmbito das pesquisas internacionais, Passos *et al.* (2018), afirmam que o intuito principal da avaliação da qualidade é estabelecer uma classificação que aponte quais as principais revistas de cada área. Para identificar as revistas de maior credibilidade, os autores trabalham com diferentes metodologias, mas, em geral, as pesquisas versam sobre levantamento de reputação entre os pares, cálculos estatísticos de impacto ou *rankings* definidos com base em critérios estabelecidos pelas principais bases de dados internacionais (PASSOS *et al.*, 2018).

Como já abordado anteriormente, determinadas bases de dados internacionais possuem o poder de definir quais publicações científicas são significantes, sendo estas as publicações de maior qualidade, sendo as duas bases de maior prestígio internacional a *Web of Science* e a *Scopus*. Estas são as bases mais citadas e que recebem as melhores pontuações em muitas avaliações de desempenho (CÉSPEDES, 2021), sendo utilizadas na maioria dos estudos bibliométricos e cientométricos.

No entanto, essas bases vêm sendo alvo de críticas aos seus critérios de indexação (CÉSPEDES, 2021), que refletem uma sub-representatividade de produção científica de países fora do eixo América do Norte – Europa, bem como de publicações cuja língua não é o inglês, considerada a língua neutra e oficial para publicações acadêmicas. Para Céspedes (2021),

[...] a ênfase na neutralidade do inglês como língua acadêmica franca muitas vezes implica uma valorização acriticamente positiva do papel desta língua, ao mesmo tempo que mascara condições diferentes e desiguais sob as quais os textos científicos são escritos, publicados e avaliados em países centrais e periféricos. (CÉSPEDES, 2021, p. 148)

Ainda assim, são estas e outras bases de dados internacionais de renome que detêm, através dos critérios para seleção de revistas a serem indexadas, o poder de atribuir prestígio as publicações contidas nestas revistas, como uma determinação dos atributos de qualidade da revista à pesquisa publicada e, conseqüentemente, atribuindo prestígio ao autor. Mas como se define e quantifica a qualidade de um periódico?

Existem diversas formas de avaliação da qualidade de periódicos científicos, com critérios baseados em características como normalização, periodicidade, indexação, colaboração entre autores, número de artigos publicados e número de citações. Um dos primeiros modelos de avaliação de periódicos foi proposto por Braga e Oberhofer em 1982 (PASSOS, 2018), com atribuição de pontuações para as revistas de acordo com as variáveis de cada critério de avaliação, sendo: normalização; duração; periodicidade; indexação; difusão; colaboração e divisão do conteúdo; e autoridade.

Este modelo não leva em consideração o número de citações da publicação, possui um foco direcionado para a estrutura do periódico, e o critério de difusão faz referência ao suporte impresso. Desse modo, é natural que o modelo tenha sofrido adaptações ao longo do tempo, se adequando a novos critérios.

Outros métodos de avaliação de qualidade que surgiram ao longo do tempo eram direcionados priorizando o que é considerado mais relevante em cada cenário. Por exemplo, periódicos brasileiros podem ser classificados por mérito, com vistas a priorizar apoio financeiro de agências de fomento. Em 1998, a CAPES propõe a base Qualis (BRASIL, c2016a) para construir indicadores de qualidade para as revistas científicas brasileiras vinculadas a programas de pós-graduação, classificando periódicos cadastrados em sua base de dados, a plataforma Sucupira (BRASIL, c2016b). Souza (2018) detalha a classificação Qualis, que

possui um indicativo de qualidade, sendo os conceitos de A1 como a maior nota, perpassando por A2; B1; B2; B3; B4; B5 E C para a nota de entrada. Essa classificação é realizada pelos comitês de consultores de cada área de atuação e seguem os critérios aprovados pelo Conselho Técnico-Científico da Educação Superior (CTC-ES), que disponibilizam uma lista de classificação dos periódicos vinculados aos programas de pós-graduação. (SOUZA, 2018, p. 49).

Para Mugnaini (2006), é também importante haver qualidade nas análises das informações contidas na produção científica, sendo necessário que a avaliação pelos pares seja rigorosa para validar a pesquisa avaliada. Porém, na avaliação da qualidade a nível de periódico, os critérios para a construção de indicadores de avaliação não costumam analisar

características do conteúdo das publicações, uma vez que se subentende que esta análise criteriosa foi cumprida justamente na etapa de revisão pelos pares.

É de extrema importância que os indicadores usados para avaliação da qualidade de periódicos sejam robustos, transparentes e precisos, não gerando dúvidas quanto a sua composição. Por muitos anos, o *Journal Impact Factor (JIF)* foi a métrica predominante na medição de qualidade das revistas. Desenvolvido por Eugene Garfield e Irving Sher em 1963, o *JIF* era a métrica usada para selecionar revistas a serem inseridas no recém-criado *Science Citation Index (SCI)*. A composição do *JIF* é baseada na média entre dois elementos: “o numerador, que é o número de citações no atual ano a quaisquer itens publicados em um jornal nos 2 anos anteriores, e o denominador, que é o número de artigos publicados nos mesmos 2 anos” (GARFIELD, 1999). Segundo o autor, a importância de se considerar estes dois indicadores na composição do *JIF*, e não apenas o número de citações, é que periódicos de menor porte e menor impacto poderiam ser prejudicados se dependessem somente do número de citações, e o *JIF* acabaria beneficiando somente grandes revistas, uma vez que o fator determinante para o impacto de periódicos é a média de citações por artigo (chamada de densidade) e o imediatismo das citações (GARFIELD, 1999).

Porém, a maneira que o *JIF* é composto permite que os periódicos manipulem e distorçam os resultados, podendo, por exemplo, publicar muitos artigos de revisão (que são mais altamente citados), e publicar resenhas e relatórios que não são incluídos na contagem de artigos, aumentando o *JIF* da revista (MINGERS; YANG, 2016).

Ao longo do tempo, e com as críticas ao *JIF*, outros indicadores utilizados na avaliação da qualidade de periódicos surgiram, como o índice-h; número médio de citações por artigo (do inglês, *citation per publication*, ou *CPP*); Eigenfactor e *SCImago Journal Rank (SJR)*. Todos estes indicadores possuem uma característica em comum: o cálculo sempre envolve uma variável baseada no número de citações. Maricato e Santos (2021) relacionam os principais indicadores utilizados para avaliar a produção científica e os pesquisadores, descritos no Quadro 2.

Quadro 2 - Indicadores bibliométricos mais usados para avaliar pesquisadores individuais e produção científica

Índice	Característica
índice-h	É um indicador numérico único que avalia a quantidade e o impacto da produção científica de um pesquisador. Apesar de ser um indicador de fácil compreensão, o índice-h apresenta inconsistência nos resultados, pois uma de suas limitações é não levar em conta artigos altamente citados.
índice-g	Desenvolvido para corrigir o índice-h em relação aos artigos mais citados, dá maior peso para eles. Por mais que o índice-g corrija essa limitação, ele pode não ser o mais adequado para comparar as produções científicas de diferentes idades, devido ter uma restrição quanto ao tempo de publicação.
índice-i10	Esta medida é utilizada pelo Google Acadêmico, e seu valor corresponde ao número de publicações, nos últimos 5 anos, com pelo menos 10 citações. Embora seja útil para identificar pesquisadores com certo nível de impacto em um determinado campo, o índice-i10 é limitante porque não contabiliza publicações com menos de 10 citações, bem como desconsidera publicações altamente citadas em seu cálculo.
RG Score	Um indicador bibliométrico/altmétrico da ResearchGate, uma plataforma proprietária de rede social acadêmica. A base para o cálculo do RG Score é qualquer contribuição que o pesquisador faz em seu perfil (artigos publicados, projetos, perguntas e respostas, entre outras interações). Um dos principais problemas de seu uso está relacionado à dificuldade de aplicação em outros contextos e a impossibilidade de auditar os dados.

Fonte: Adaptado de Maricato e Santos (2021, p. 4, tradução nossa).

Como é possível observar, as diversas metodologias de avaliação da produção científica através destes indicadores possuem restrições, seja ao não captar todo o escopo de produções científicas de um pesquisador devido a antiguidade da publicação, ou devido a falta de transparência dos dados utilizados nos cálculos dos indicadores. O vasto número de índices e indicadores desenvolvidos mostram um interesse em corrigir as limitações dos anteriores (MARICATO; SANTOS, 2021), e ainda que diversos destes indicadores sejam considerados bem difundidos e utilizados em larga escala na avaliação da produção científica, nenhum deles é completo, captando integralmente o impacto científico das publicações.

Alguns pesquisadores consideram que a altmetria é uma métrica melhor para avaliar o impacto científico por ser capaz de diminuir algumas limitações desses indicadores bibliométricos (GALLIGAN; DYAS-CORREIA, 2013). Entretanto, estudos desenvolvidos nos últimos anos (MARICATO; MARTINS, 2017; HUANG; WANG; WU, 2018; MOHAMMADI *et al.*, 2015; SUN *et al.*, 2018; OUCHI *et al.*, 2019;) apontam para um potencial de uso conjunto de indicadores baseados em citações e indicadores altmétricos, visando a complementação entre estes indicadores em busca de formas mais completas e imparciais de se avaliar o impacto científico.

Nesta pesquisa, objetivando a compreensão das dinâmicas entre indicadores bibliométricos e indicadores altmétricos no âmbito dos portais de periódicos brasileiros de

acesso aberto, são apresentados a seguir os principais conceitos e características sobre a bibliometria e a altmetria, de modo a explicitar a diferença entre estas métricas e as possibilidades que cada uma permite para quantificar o impacto científico online, e sua importância para o escopo deste trabalho.

2.4 BIBLIOMETRIA E ESTUDOS DE CITAÇÃO

A bibliometria é considerada a principal técnica dedicada a estudar indicadores científicos. Clássica e bem difundida, seu uso remonta ao final do século XIX. Ela surgiu primeiramente sob o nome de bibliografia estatística, e de acordo com Sengupta (1992), o primeiro estudo bibliométrico foi realizado por Campbell em 1896, cujo trabalho foi supostamente a primeira tentativa envolvendo métodos estatísticos para estudar a dispersão de assuntos. Na época, os métodos utilizados por Campbell sequer levavam o nome de bibliografia estatística. O termo foi originalmente utilizado por E. Wyndham Hulme em 1922 (PRITCHARD, 1969), e demorou algumas décadas para ser difundido e largamente utilizado. Hulme usou o termo bibliografia estatística como significado dos processos da ciência e da tecnologia por meio da contagem de documentos. Ao longo dos anos, a definição conceitual do que é bibliografia estatística teve suas variações. Segundo Raisig (1962), bibliografia estatística

[...] é a montagem e interpretação de estatísticas relativas a livros e periódicos para demonstrar movimentos históricos, podendo ser usada em uma variedade de situações [...] com o propósito principal de determinar o uso de pesquisa nacional ou universal de livros e periódicos, e para verificar em muitas situações o uso geral de livros e periódicos. (RAISIG, 1962, p. 450).

Pritchard (1969) não considerava o termo bibliografia estatística satisfatório, e propõe então o termo bibliometria (termo original em inglês: *bibliometrics*). Pritchard esperava que o termo bibliometria fosse explicitamente usado em todos os estudos futuros que buscassem quantificar os processos de comunicação escrita e que rapidamente ganhasse aceitação no campo da Ciência da Informação.

Em 1978, Nicholas e Ritchie foram os primeiros a categorizar a área de bibliometria (SENGUPTA, 1992), baseado no escopo da bibliometria: o provimento de informação sobre a estrutura do conhecimento e como ele é comunicado. Eles classificaram a bibliometria em dois grupos: bibliometria descritiva; e bibliometria comportamental. A descritiva descreve

características ou funcionalidades da literatura, enquanto a comportamental é aquela que examina as relações formadas entre componentes da literatura.

Indo além dessa categorização inicial, a bibliometria pode ser considerada uma ciência quantitativa, objetiva, e seu escopo envolve o estudo das relações dentro de uma literatura, como é o caso de estudos de citações. Com esse enfoque, Stevens (1953, apud SENGUPTA, 1992, p. 78) divide a bibliometria em duas categorias básicas: bibliometria descritiva, usada para produtividade; e bibliometria avaliativa, para contar o uso da literatura de um tópico, assunto ou disciplina específica. Dentro da categoria bibliometria avaliativa de Stevens (1953) estão os estudos de contagem de referências e de citações (SENGUPTA, 1992).

Tendo em vista o grande potencial das análises bibliográficas para avaliação da literatura acadêmica, a bibliometria faz uso do número de citações como indicador. Os documentos científicos, aqueles publicados nos canais de comunicação formal da ciência, adquiriram uma configuração estrutural estável ao longo do tempo. O tipo de documento científico por excelência é o artigo científico (BARBA, 2003, p. 97), e sua estrutura é composta pelos seguintes elementos: título; lista de autores e afiliação institucional dos autores; resumo; palavras-chave; texto principal, em suas subdivisões; e lista de referências bibliográficas. Várias dessas partes que constituem um artigo científico recebem atenção das análises bibliométricas, mas a lista de referências reúne as informações imprescindíveis que refletem com maior clareza a inserção do trabalho científico individual em uma obra coletiva. Barba (2003, p. 105) afirma que o duplo caráter da lista de referências, cognitivo e social, deu origem a uma enxurrada de estudos sobre suas funções e sua utilidade para o estudo da ciência através das conexões que ela estabelece entre trabalhos e autores.

A lista de referências se mostra, portanto, uma parte essencial da publicação científica que aponta para os estudos anteriores a ela, e que foram utilizados como fonte de informação. Afinal, um artigo científico não é desconectado dos demais estudos dentro da mesma área de conhecimento, ele faz parte de uma rede de conhecimento científico que atesta, através de suas referências bibliográficas, o reconhecimento a outra pesquisa. A pesquisa que recebe o reconhecimento é, portanto, a pesquisa citada. A citação implica em uma relação entre uma parte ou todo o documento citado e uma parte ou todo o documento que a cita (MALIN, 1968). De acordo com a ABNT, citação é a “menção de uma informação extraída de outra fonte” (ABNT, 2002, p. 1).

Ao longo dos anos, o significado de citação permaneceu imutável, o que muda e evolui constantemente são as formas de coletar a citação através de diferentes ferramentas de coleta de dados, em especial as proporcionadas pelo avanço da tecnologia, bem como as

técnicas e possibilidades de análise dos dados se desenvolvem graças ao aumento da capacidade de processamento computacional. Entretanto, apesar das citações serem consideradas uma métrica norteadora na medição de impacto científico, estas possuem limitações. Primeiramente, porque de acordo com Maricato e Martins (2017) as citações medem apenas um aspecto limitado da qualidade.

[...] tanto as citações, quanto as revisões por pares, são consideradas indicadores parciais de "impacto científico", não sendo capazes de revelar suficientemente o "impacto" total da pesquisa. Dadas estas limitações, a combinação de revisão por pares com "abordagem multi-métrica" é proposta necessária para a avaliação da pesquisa (MARICATO; MARTINS, 2017, p. 58).

Em segundo lugar, o real significado das citações vem sendo amplamente debatido. Entende-se que a análise de citação é baseada na suposição de que um documento referenciado em uma publicação subsequente marca a influência intelectual do documento citado sobre o documento citante (HAUSTEIN; LARIVIÈRE, 2015). Porém, de acordo com Small (1987 apud HAUSTEIN; LARIVIÈRE, 2015, p. 125), uma vez que o número de itens citados é geralmente restrito pelo comprimento da publicação, uma lista de referência não deve ser considerada uma lista completa, mas uma seleção das fontes mais influentes relacionadas a uma obra.

Desse modo, ao se assumir que o número de citações reflete o impacto científico das publicações acadêmicas, uma parcela deste impacto não é mensurado por ir além do espectro que as citações conseguem captar e medir. Assim, métricas que consigam captar outras formas de se mensurar a influência de publicações científicas devem ser consideradas na avaliação do impacto acadêmico.

Diversos estudos que analisaram o impacto da ciência consideram exclusivamente a análise de indicadores bibliométricos como abordagem metodológica. Bornmann e Mutz (2015) analisaram o crescimento da ciência baseados em dois conjuntos de dados: no número de publicações consideradas itens de origem na *Web of Science* divididas por ano de publicação; e no número de referências citadas nas publicações dos itens de origem por ano de citação. Os autores reconhecem uma limitação do uso de análises baseadas em citação, de que, por usarem o número de citação, parte da literatura que ainda não foi citada não é considerada no estudo (BORNMANN; MUTZ, 2015).

Ellegaard e Wallin (2015) buscam mensurar o impacto científico através da análise bibliométrica da produção científica extraída da *Web of Science*, julgando o impacto dos artigos analisados utilizando a média de citações dos dados coletados. Os autores consideram

que o surgimento de ferramentas alternativas de medição do impacto científico como contagem de downloads, serviços de gestão de referências como o *Mendeley*, e o compartilhamento de conteúdo acadêmico em redes sociais possibilitam a análise e interpretação do impacto científico da literatura de várias formas. Porém, os autores consideram que essa tendência de análise se torna complicada com o crescimento do número de fontes de dados alternativas (ELLEGAARD; WALLIN, 2015).

Purkayastha *et al.* (2019) compararam dois indicadores baseados em citação, tradicionalmente utilizados para publicações da área de Biomedicina, e os aplicaram a todos os artigos coletados da base *Scopus*, divididos em seis grandes áreas de conhecimento, e correlacionaram estes indicadores. Os autores identificaram que a força das correlações de Pearson e Spearman varia de forte a muito forte, sendo 0.79 a menor correlação de Spearman para a área de Engenharia, e a maior de 0.98 para a área de Biomedicina. Estes resultados apontam para uma performance similar entre os indicadores analisados, uma vez que ambos são baseados em citações (PURKAYASTHA *et al.*, 2019), bem como para a validade do uso destes para todas as áreas de conhecimento. Os autores reconhecem que existem limitações no uso exclusivo de métricas tradicionais, como a contagem pura de citações e o *JIF*, e ressaltam que a tendência de estudos bibliométricos recentes é se concentrarem no desenvolvimento de métricas alternativas a nível de artigo para medir o impacto de citação.

Apesar das bases de dados bibliométricos *Web of Science* e *Scopus* serem as mais utilizadas nos estudos bibliométricos e cientométricos tradicionais, outras bases de dados que coletam citações também foram consideradas a medida que foram surgindo e se consolidando, como é o caso do Google Scholar. Bar-Ilan (2010) analisou comparativamente a cobertura da *Web of Science*, *Scopus* e Google Scholar para um único livro, identificando que estas bases de dados são suplementares, não havendo uma base de dados única capaz de substituir todas as outras. Sendo assim, faz-se necessário conhecer as principais fontes de dados bibliométricos disponíveis na web.

2.4.1 Fonte de dados bibliométricos

Ao longo dos capítulos anteriores foram mencionadas diversas fontes de dados bibliométricos utilizadas em diferentes contextos e estudos. Nesta seção, é apresentado um panorama sobre as principais delas, sendo: *Web of Science*, *Scopus*, *Google Scholar* e *Crossref*.

A *Web of Science* é uma base de dados por assinatura da *Clarivate Analytics*. A *Web of Science* provê serviços de indexação de citação para produções acadêmicas, com uma coleção multidisciplinar e global que indexa milhares de periódicos, livros, conferências e outros materiais acadêmicos, disponibilizados para o desenvolvimento de estudos quantitativos que têm a ciência como campo de estudo. A principal coleção da *Web of Science* é composta de dez índices de citação, sendo que os seguintes índices indexam citações de artigos, cobrindo mais de 12 mil periódicos de alta qualidade, inclusive de acesso aberto: *Science Citation Index* (SCI), *Social Sciences Citation Index* (SSCI), *Arts & Humanities Citation Index* (A&HCI) (CLARIVATE, 2020).

A *Scopus* é uma base de dados de resumos e citações, com uma curadoria feita por especialistas em cada disciplina. Possui mais de 25 mil títulos indexados, provindos de mais de cinco mil editoras internacionais. A *Scopus* oferece uma visão abrangente das produções científicas mundiais nas áreas de Ciência, Tecnologia, Medicina, Ciências Sociais e Artes e Humanidades (ELSEVIER, 2020). Os tipos de fontes de dados cobertas pela *Scopus* são publicações periódicas que possuam um *International Standard Serial Number* (ISSN) como periódicos, séries de livros e de conferências, ou publicações não-periódicas que possuam um *International Standard Book Number* (ISBN). Para garantir a cobertura, a descoberta e as medidas de impacto para pesquisa em todas as áreas de conhecimento que a *Scopus* abrange, esta base cobre diferentes tipos de fontes de dados (ELSEVIER, 2020).

De acordo com Tennant (2020), as bases *Web of Science* e *Scopus* são “sinônimo de dados de pesquisa internacional. Elas são amplamente consideradas pela comunidade acadêmica como as duas fontes mais confiáveis e competentes de dados bibliométricos (TENNANT, 2020, p. 1, tradução nossa). Entretanto, apesar de serem consideradas globais, a representatividade de ambas as plataformas vem sendo questionada por diversas pesquisas (MELO; TRINCA; MARICATO, 2021; TENNANT, 2020; VERA-BACETA *et al.*, 2019) que mostram que a maior parte dos artigos indexados por essas duas bases são provenientes de alguns poucos países ocidentais de língua inglesa.

Diferente da *Web of Science* e da *Scopus*, o *Google Scholar* segue uma abordagem inclusiva e automatizada, indexando qualquer documento acadêmico que seus rastreadores conseguem encontrar e acessar, incluindo aqueles por trás de assinaturas. Além disso, o acesso do *Google Scholar* é gratuito, permitindo que os usuários acessem um índice de citação abrangente e multidisciplinar sem o pagamento de taxas (MARTÍN-MARTÍN *et al.*, 2021). Considerando que as plataformas disponíveis há mais tempo são a *Web of Science*, ativa desde 1997, *Scopus* e *Google Scholar*, ambas criadas em 2004, a maioria dos estudos que envolvem

indicadores bibliométricos utiliza estas bases como fonte de dados (MARTÍN-MARTÍN *et al.*, 2021).

A *Crossref* é uma agência oficial de registro de identificadores para objetos digitais, tais como artigos, livros e outros produtos de pesquisa, com o objetivo de torná-los fáceis de serem encontrados, citados, referenciados acessados e reutilizados. Assim, a *Crossref* é uma organização sem fins lucrativos que existe para melhorar a comunicação científica, marcando e compartilhando metadados de pesquisa e disponibilizando uma gama de serviços para instituições e editoras para ajudar a colocar a pesquisa em contexto, incentivando sua visibilidade e conectividade.

Entre os serviços oferecidos para membros que registram conteúdo junto à *Crossref*, estão o fornecimento de *DOI*, links de referência, e recuperação de metadados. O *DOI* é um identificador persistente e único, criado ao mesmo tempo ou antes que o próprio objeto, a fim de proporcionar uma etiqueta que identifique o objeto por toda sua vida útil, independentemente de sua localização (BUENO; AGUIAR; SILVA, 2018). Quando os membros da *Crossref* registram *DOIs* para seu conteúdo, a *Crossref* exige que um conjunto relativamente pequeno de metadados bibliográficos seja incluído junto com o *DOI* e sua *URL*. No entanto, o esquema de metadados permite uma gama muito maior de metadados que também podem ser incluídos opcionalmente. Entre esses dados opcionais está uma lista de referência que pode ser incluída no registro de metadados de cada publicação. O sistema da *Crossref* tentará então combinar os dados fornecidos em uma referência depositada com o *DOI* da obra citada, com base nos metadados fornecidos ao *Crossref* pelo editor da obra citada quando o *DOI* foi registrado.

Isso significa que a *Crossref* pode encontrar e "contar" referências quando tanto a obra citante quanto a citada têm *DOI*'s registrados por meio da *Crossref*. E, da mesma forma, pode contabilizar as referências quando o editor da obra citada opta por fornecer sua lista de referências. Podemos perceber que a correspondência (ou resolução) de referências bibliográficas para encontrar registros em uma coleção é um algoritmo crucial no ecossistema da *Crossref*. A correspondência automática de referências permite descobrir relações entre citações em grandes coleções de documentos, possibilitando à *Crossref* calcular indicadores como contagem de citações, índices-h e fatores de impacto (TKACZYK, 2018).

A *Crossref* disponibiliza, através de sua API, acesso automatizado aos indicadores disponíveis publicamente, possibilitando uma coleta de dados bibliométricos a partir de metadados de qualidade e confiáveis, ao contrário de outras fontes de dados bibliométricos, como o Google Scholar. No caso do Google Scholar, por mais que seja uma fonte com ampla

cobertura, com estimativa de mais de 300 milhões de registros acadêmicos indexados (DELGADO LÓPEZ-CÓZAR *et al.*, 2019), a relativa baixa qualidade dos metadados disponíveis e a dificuldade de extração tornam o uso do Google Scholar um desafio nas análises bibliométricas (MARTÍN-MARTÍN *et al.*, 2018).

Sendo assim, a *Crossref* se mostra uma fonte de dados bibliométricos transparente, de alta cobertura e consistente, tendo sido utilizada em diversas pesquisas que abordaram indicadores bibliométricos (CHUDLARSKÝ; DVOŘÁK, 2020; YAN; GERSTEIN, 2011; LIU *et al.*, 2013, YAN).

2.5 ALTMETRIA E A MENSURAÇÃO NA WEB 2.0

Antes de tratar sobre o impacto da Web 2.0 na forma de mensurar a ciência, deve-se deixar claro o salto tecnológico que representa a evolução da Web 1.0 para a Web 2.0. A Web 1.0 representou a transposição do meio físico para o digital, de modo estático, sem interatividade, com o leitor como um agente passivo em relação ao conteúdo digital. Com o avanço das tecnologias de comunicação e informação, surge a Web 2.0, que passa a possibilitar uma maior interatividade entre o usuário e o conteúdo. Observa-se que aqui o leitor passa a ser considerado usuário, pela forma como este passa a se relacionar com o conteúdo online. A Web 2.0 possibilita a criação de blogs e inserção de conteúdos que podem ser manipulados. A organização deste conteúdo passa de uma forma taxonômica para uma infinidade de possibilidades de categorização.

A possibilidade de interação entre o usuário e o conteúdo de blogs, que agora passa a ser chamado de serviço, foi só o começo. De acordo com Santos (2011), esse processo “desencadeou a formação de comunidades não só entre os autores dos blogs, mas entre os visitantes dos mesmos. E depois, comunidades dentro das comunidades” (SANTOS, 2011, p. 27). Face ao desenvolvimento tecnológico presenciado nos últimos anos e com a criação da internet e constante aderência de novos usuários que buscam nela um meio de comunicação e de conectividade com o mundo, as redes sociais digitais surgem sustentadas por softwares com funcionalidades que permitem o relacionamento e interação entre os usuários da rede.

Assim, a forma de se comunicar online também evolui com o surgimento da Web 2.0, influência que também impactou a comunicação científica, que se baseia nos suportes tecnológicos disponíveis e incorporou diversas facilidades trazidas pelos meios digitais. Graças aos meios eletrônicos e à internet, a comunicação científica pôde romper barreiras geográficas, e as publicações científicas alcançam públicos distribuídos mundialmente, bem

como se tornam mais acessíveis à comunidade de forma geral, e não apenas aos membros da academia.

Não apenas as formas de comunicação científica oficial foram beneficiadas com a Web 2.0, mas também as correspondências entre pesquisadores. Através de diversos canais como e-mail, blogs, redes sociais, os membros da comunidade acadêmica se comunicam e trocam informações valiosas para o desenvolvimento de pesquisas e produção de conhecimento científico de forma muito mais rápida e eficiente do que antes da Web 2.0. As coautorias de artigos científicos também passam a transpor barreiras geográficas, permitindo parceria entre pesquisadores de um mesmo campo do conhecimento, mas atuantes em continentes diferentes, ampliando as redes de colaboração científica. A velocidade na comunicação e facilidade de acesso ao conhecimento científico influenciam no aumento do número de pesquisas publicadas e surgimento de novos periódicos científicos, agora em formato eletrônico.

O surgimento das mídias e redes sociais digitais também beneficiou a comunicação científica. O próprio conceito de mídias sociais objetiva reunir pessoas através de grupos de interesse, formando comunidades específicas. No meio digital, os dados das comunicações entre usuários destas comunidades são registrados pelas plataformas online. Isso quer dizer que, quando um pesquisador usa o *Twitter* para comunicar à sua rede de contatos que publicou um artigo em determinado periódico científico, esta informação fica registrada.

As diversas ações que um usuário executa nas redes sociais são dados passíveis de serem quantificados e transformados em informação, assim como o conteúdo de publicações científicas online podem ser coletadas e armazenadas por bases indexadoras. Surgem então novas ferramentas online que permitem utilizar filtros para quantificar a literatura acadêmica. Neste cenário de quantificação da literatura acadêmica, das menções de publicações científicas e outros tipos de documentos acadêmicos em redes sociais digitais, surgem os indicadores alométricos.

Após uma vasta discussão sobre o uso da contagem de citações para medir o impacto acadêmico, fica claro que existem críticas sobre os seus usos, uma vez que as citações refletem um espectro limitado dos usos da literatura acadêmica (PRIEM *et al.*, 2011), não captando determinados tipos de uso que acabam passando invisíveis aos olhos da bibliometria.

A altmetria não surge objetivando substituir a bibliometria, mas sim captar e quantificar os usos da literatura acadêmica que esta não abrange. Isso não quer dizer que métricas que utilizam a contagem de citações não são úteis, mas elas não são mais suficientes (PRIEM *et al.*, 2011). Afinal, com todas as possibilidades de interações nas mídias sociais no

ciberespaço, é de se esperar que as publicações científicas sejam compartilhadas por diversos canais digitais e usadas por diferentes públicos que não apenas o acadêmico. Portanto, tais interações entre pessoas e publicações que circulam nas mídias sociais deixam rastros que podem ser mensurados.

O termo *altmetrics*, original em inglês para altmetria, surgiu em 2010, e foi definido por Priem (2010) como “a criação e estudo de novas métricas baseadas na Web Social para análise de informações acadêmicas”. Antes, os primeiros estudos se referiam a este campo como *article level metrics* (métricas a nível de artigo) ou, *alternative metrics* (métricas alternativas). O termo *altmetrics* foi usado pela primeira vez em um *tweet* de Jason Priem, em seu perfil pessoal na rede social digital *Twitter*. Para Priem, o problema com o termo “métricas a nível de artigo” é que este não implica a diversidade de medidas, nem o potencial da altmetria. As fontes de dados são inúmeras, bem como as ações que a altmetria é capaz de medir. São compartilhamentos, menções, acessos, downloads, leituras, capturas, “curtidas” e diversas outras formas de interação com o conteúdo acadêmico possíveis de serem mensuradas. Estas refletem parte do impacto das publicações científicas para além do público acadêmico, ampliando o espectro de fenômenos passíveis de serem observados.

Analisando as origens da altmetria, Maricato e Martins (2017) percebem a associação do termo com métricas em nível de artigo e métricas alternativas, este segundo remetendo à altmetria como alternativa às métricas tradicionais (bibliometria e cientometria). O potencial da Web para medir usos da pesquisa científica que as métricas tradicionais não são capazes de mensurar, e que podem ser utilizados na avaliação da pesquisa, foi reconhecido em estudos anteriores (ZAHEDI; COSTAS; WOUTERS, 2014). Mas, devido ao formato “não convencional” dos indicadores altmétricos para avaliação da pesquisa, se tornou um consenso para pesquisadores chamá-los de métricas alternativas. O próprio nome *altmetrics* é uma contração das palavras *alternative* e *metrics*.

Porém, o posicionamento de utilizar o termo altmetria como uma alternativa às bibliométricas e cientométricas tradicionais pode ter dificultado seu desenvolvimento, uma vez que a ideia de alternativa está relacionada a superioridade, o que não se aplica neste caso. O conceito de altmetria como métricas alternativas às tradicionais acaba sendo equivocado devido a essa ideia, porém o termo permanece sendo utilizado (MARICATO; MARTINS, 2017) e, a esta altura, já é bem difundido entre os pesquisadores da área.

Para Roemer e Borchardt (2015), as altmétricas se mostram dependentes das tecnologias da informação e comunicação e do ambiente online. Os autores ressaltam que a maior parte da conceituação da altmetria a compara com outras métricas já existentes, e a

relaciona com a análise de citações de artigos em mídias e redes sociais digitais. Mas a grande controvérsia que a área da altmetria vem enfrentando desde o seu surgimento é justamente por conta das tentativas dos pesquisadores em compará-la com a bibliometria. Os autores afirmam que

[...] a bibliometria e a altmetria compartilham das mesmas intenções ao buscar analisar quantitativamente o conhecimento científico, apesar de suas definições do que é conhecimento científico e seus métodos de análise divergirem significativamente. No entanto, com a altmetria oferecendo uma imagem muito mais imediata do impacto acadêmico do que as bibliométricas baseadas em citações, os pesquisadores têm ficado naturalmente curiosos para saber se a altmetria pode ser usada como um preditor de citações futuras, o que é desejável como uma métrica de longo prazo do sucesso acadêmico. (ROEMER; BORCHARDT, 2015, p. 22).

A altmetria pode ser considerada uma área fortemente ligada à Comunicação Científica e ao campo de Ciência da Informação, com grande proximidade (e ao mesmo tempo diferenças) das áreas de bibliometria e cientometria (MARICATO; VILAN FILHO, 2018). A altmetria tem como vantagem poder medir o impacto da produção científica em diversas fontes de dados, incluindo as mais recentes e não abordadas por outras métricas, como as redes sociais. Porém, essa variedade de fontes e ferramentas torna a altmetria uma atividade complexa e delicada, o que é reforçado pela falta de clareza do significado das ações e reações nas mídias e redes sociais, devendo a altmetria estar relacionada aos estudos de sociabilidade na rede para auxiliar no entendimento do impacto destas ações e reações (MARICATO; MARTINS, 2017, p. 62).

Nesse sentido, Rasmussen e Andersen (2013) levam em consideração que o próprio conceito de altmetria envolve múltiplas variáveis para poder medir o impacto acadêmico, com o propósito de prover uma imagem mais completa das publicações científicas do que os indicadores bibliométricos tradicionais. Essas variáveis, ou ações nas mídias sociais, tais como contagem de visualizações, acessos, marcações e outras já previamente citadas, variando de acordo com as funcionalidades e objetivos de cada plataforma de mídia social, podem nos dar a impressão do quanto uma publicação é interessante e o quanto ela está sendo efetivamente usada online. Porém, é importante entender que essas ações não possuem significado equivalente ao número de citações, ou seja, não é porque um artigo científico possui um alto número de downloads que ele será necessariamente altamente citado.

Mesmo havendo estudos demonstrando que há uma correlação moderada entre indicadores de downloads e citações (PERNEGER, 2004; WANG *et al.*, 2014), outras variáveis influenciam nessa correlação de uma maneira que não pode ser mensurada. Por

exemplo, um artigo científico de um estudo em estágio inicial sobre um assunto que é tendência na comunidade em geral pode ter seu número de downloads elevado, mas não ser altamente citado com o tempo, devido não apresentar resultados suficientes para que seja um dos principais artigos consumidos pelos pesquisadores da mesma área de conhecimento a qual pertence. Assim, os números de citações, downloads e leitores podem ser similares para muitos artigos, mas para tantos outros, as variáveis representam algo diferente (RASMUSSEN; ANDERSEN, 2013).

Adquirir dados para avaliação altmétrica também tem seus desafios. Por mais que as plataformas de fontes de dados busquem facilitar a busca e recuperação dos dados ao colocar a disponibilidade dos usuários ferramentas na própria interface ou na API como filtros de busca avançada para refinamento das informações, algumas vulnerabilidades na coleta devem ser consideradas. Priem, Piwovar e Hemminger (2011) citam a possibilidade de inconsistência nas bases de dados como uma vulnerabilidade que pode afetar a confiabilidade dos dados altmétricos provenientes das bases.

Wouter e Costas (2012, p. 40) afirmam que, no modelo de avaliação da pesquisa, a transparência e consistência dos dados e indicadores pode ser até mesmo mais importante que a disponibilidade desses dados, se referindo ao acesso aberto. Porém, as próprias plataformas fontes de dados altmétricos buscam ser transparentes quanto ao método de coleta e indexação dos indicadores altmétricos que reúnem, deixando claro em suas políticas os métodos utilizados, buscando sanar questionamentos a esse respeito. A transparência é, de fato, uma das qualidades da altmetria, onde as próprias inconsistências, se encontradas, são documentadas e disponibilizadas para o público (RASMUSSEN; ANDERSEN, 2013).

A transparência das altmétricas também busca minimizar a manipulação dos indicadores, um dos grandes problemas do *Journal Impact Factor*, baseado em citações. Algumas plataformas que são fontes de dados altmétricos tornam particularmente difícil a manipulação das variáveis, como é o caso do *Mendeley*, site gerenciador de referências onde os usuários podem adicionar publicações à sua biblioteca pessoal, e que ao fazê-lo, o site considera o usuário como leitor da publicação.

Uma vez que determinado leitor é contabilizado para uma publicação, ele não será contabilizado duplamente se remover e adicionar novamente esta mesma publicação em sua biblioteca e, portanto, o indicador altmétrico de leitores no *Mendeley* não é manipulável. O *Twitter*, rede social que permitem postagem (*tweet*) e compartilhamento (*re-tweet*) de textos curtos (que podem ou não conter outros tipos de mídias não-textuais), permite a fácil diferenciação das fontes de *tweets* e *re-tweets*, dificultando também a manipulação. Já o

número de downloads pode ser manipulado, porém seu significado não é tão claro quanto outras ações em outras redes e mídias sociais, visto que o download de uma única publicação não necessariamente representa uma única leitura desta publicação, já que o rastreamento do uso do documento digital da publicação se encerra na contagem do download, mas o documento pode ser compartilhado por outros meios com outros usuários, que também o lerão.

As alométricas também podem revelar como uma maior diversidade de indivíduos são afetados pelos conteúdos científicos. Para Sun *et al.* (2018), é necessário investigar as pessoas que mencionam artigos científicos e o conteúdo desses *tweets* para compreender a extensão com que o *Twitter* captura o impacto social. Os autores propõem um novo método para identificação dos usuários influenciadores de opinião, ou seja, aqueles que *tweetam* artigos diretamente da fonte, diferenciando-os dos usuários que seguem o perfil no *Twitter* dos influenciadores e apenas *re-tweetam* seu conteúdo. Este tipo de estudo leva em consideração análises qualitativas do conteúdo dos *tweets*, analisando fatores emocionais através da atribuição de valores para determinados termos léxicos para conseguir converter fatores qualitativos em quantitativos.

Nesse tipo de análise do impacto social, o contexto em que os artigos científicos são mencionados e discutidos é mais valioso do que apenas números, e ajuda a compreender melhor tanto os usuários, quanto o significado dos *tweets* como uma medida alométrica (SUN *et al.*, 2018), porém as análises são muito mais complexas e englobam mais variáveis qualitativas para dar sentido a cada ação executada na rede social que envolva a produção acadêmica.

A velocidade com que os agregadores coletam os dados alométricos também é um ponto a favor da almetria. Indicadores alométricos começam a ser produzidos imediatamente após uma publicação científica estar disponível online, com a coleta sendo diária ou semanal, a depender do agregador. A disponibilidade imediata desses dados, a depender das ações executadas a partir da publicação, permitem aos autores, editores e demais agentes envolvidos no processo de produção de conhecimento científico terem um feedback com tempo de resposta mais rápido do que as citações, refletindo a influência que a publicação científica está exercendo em um campo particular. Além disso, os indicadores alométricos acabam contribuindo para a tomada de decisão de Instituições de Ciência e Tecnologia, agências de fomento, que usam esse tipo de informação de maneira estratégica para direcionar os investimentos ou recursos públicos para determinados departamentos e projetos de pesquisa de maior impacto na comunidade.

Desde o surgimento dos indicadores altmétricos há discussões sobre sua credibilidade e possíveis ameaças para a avaliação das atividades científicas. Em 2013, Jeffrey Beall, autor de um blog que publica o nome de revistas predatórias de acesso aberto, acusou a altmetria de possuir um alto potencial de ser manipulável, além de não possuir nenhum tipo de regulação ou autoridade que regule a coleta de dados. Brigham (2014) afirma que não é possível utilizar altmétricas para comparações entre disciplinas.

Quanto à possibilidade de manipulação, como dito anteriormente, há plataformas que dificultam a manipulação dos dados altmétricos, já outras possuem brechas que podem ser uma oportunidade de manipulação como, por exemplo, o *Facebook* permite que empresas criem perfis automatizados feitos exclusivamente para curtir determinadas postagens. Porém, conforme demonstrado em estudos recentes (PRIEM; PIWOWAR; HEMMINGER, 2011, THELWALL *et al.*, 2013, OUCHI *et al.*, 2019), o *Facebook* não chega a ser uma fonte de dados altmétricos relevante na composição de indicadores devido sua cobertura extremamente baixa, o que pode em parte ser ocasionado pela quantidade de perfis fechados, sem possibilidade de coleta de dados. Conforme Stacy Konkiel, autora no blog do agregador Impactstory.org, em resposta às críticas de Jeffrey Beall, “não há como negar que a manipulação acontece, e não está limitada apenas a altmetria” (KONKIEL, 2014, tradução nossa).

Quanto à hipótese de não ser possível utilizar indicadores altmétricos para comparar disciplinas, diversos estudos mostram que a altmetria permite essa comparação (HAMMARFELT, 2014; HAUSTEIN; LARIVIÈRE, 2014, MOHAMMADI *et al.*, 2014, WANG *et al.*, 2014; ZAHEDI; COSTAS; WOUTERS, 2017). Roemer e Bordhardt (2015) abordam as diversas discussões, controvérsias e críticas que cercaram a altmetria desde o seu surgimento, de maneira a deixar claro que as ameaças à confiabilidade da altmetria são menores do que os críticos se fazem pensar, e as possibilidades de uso são a defesa ideal do movimento altmétrico.

É interessante observar as críticas e defesas à altmetria através dos estudos e perceber que, à medida que o campo da altmetria se torna cada vez mais bem estruturado com estudos e resultados promissores, a direção dos estudos passam da análise das novas possibilidades e das ameaças às métricas tradicionais, para a complementação e uso conjunto entre as métricas. Os estudos mais recentes em altmetria concentram-se mais na avaliação do impacto, na própria importância da pesquisa altmétrica, e na relação entre indicadores altmétricos e indicadores de citação tradicionais (SUN *et al.*, 2018).

Estas considerações sobre as possibilidades e desafios que a altmetria enfrenta nos ajudam a entender melhor o cenário em que as análises altmétricas se inserem, e as possibilidades de interpretação dos dados. Nenhuma métrica se mostrou perfeita até o momento, e dificilmente alguma será. Porém, o uso de diferentes técnicas de análise do impacto acadêmico possibilita uma visão mais ampla do alcance das publicações, dentro e fora da comunidade científica, considerando, principalmente, o uso em conjunto de indicadores altmétricos e indicadores bibliométricos.

Para facilitar a compreensão do desenvolvimento do campo da altmetria, apesar de ainda ser considerada uma métrica nova, a altmetria já possui mais de 842 artigos científicos neste tema indexados na base de dados *Web of Science*, em uma busca simples através da palavra-chave “*altmetrics*” feita em outubro de 2021, quando este trabalho é escrito. Ao restringir a busca para publicações anteriores a 2016, abrangendo os anos de 2012 a 2015, o total de artigos encontrados caiu para 118, mostrando o recente interesse de pesquisadores pelo tema e o crescimento da área. Vale ressaltar que a altmetria possibilita indicadores para medição do impacto acadêmico não apenas de artigos científicos, mas também de outros tipos de informação acadêmica fruto de pesquisas e publicadas online, como dados de pesquisa e softwares. Essa é mais uma vantagem da altmetria, visto que esse tipo de produção acadêmica acaba ficando esquecida pelas métricas tradicionais baseadas em citação (BRIGHAM, 2014).

Antes de demonstrar a evolução e variedade dos estudos altmétricos, é interessante listar as principais ferramentas agregadoras de dados altmétricos. Já ficou claro que os dados altmétricos podem ser extraídos de diferentes redes e mídias sociais, sites e blogs. Porém, na prática, seria extremamente oneroso para o pesquisador extrair individualmente dados de cada plataforma online, uma vez que cada uma delas possui sua própria *API* e funções para coleta desses dados, o que também exigiria do pesquisador conhecimento de cada uma das documentações de *APIs* disponíveis.

Para otimizar a coleta e reunir dados de diversas plataformas diferentes, existem as ferramentas agregadoras altmétricas. Embora essas ferramentas usadas para coletar altmétricas estejam em constante evolução, atualmente existem três principais agregadores altmétricos que qualquer pessoa pode usar para explorá-los, e que são fontes de dados altmétricos para diversos estudos, a saber:

- Altmetric.com: monitora mais de 20 fontes para menções de resultados de pesquisa para trazer a imagem mais relevante e atualizada da atividade e discussão online, entre elas: *Facebook*, *Twitter*, *Youtube*, *Wikipedia*, *blogs*, e gestores de referências online (ALTMETRIC, [202-]).

- PlumX: categoriza as métricas disponíveis em cinco categorias distintas: citações, uso, capturas, menções e mídia social. Define a produção de pesquisa em 67 tipos diferentes sobre os quais coletam métricas sendo que 56,5% das publicações pertencem ao tipo artigo científico (PLUM ANALYTICS, c2021).
- Impactstory.org: usa o *Open Researcher and Contributor ID* (ORCID) para encontrar e importar os trabalhos acadêmicos, e suas fontes de dados são o *Altmetric.com*, *Bielefeld Academic Search Engine* (BASE), *Mendeley*, *Crossref*, *ORCID* e *Twitter* (IMPACTSTORY, [202-]).

Logo nos primeiros anos após o surgimento da altmetria, os estudos altmétricos eram direcionados para reunir um corpo emergente de pesquisa para o estudo e uso de altmetria, demonstrando sua importância e construindo argumentos para fundamentar seu uso. Priem, Groth e Taraborelli (2012) reuniram uma coleção de estudos com este objetivo, que abordam áreas como análise estatística de fontes de dados altmétricos; validação de métricas e identificação de possíveis vieses nas medições; validação de modelos de descoberta científica baseados em altmetria; guias teóricos com base empírica para o uso da altmetria; e outras pesquisas relacionadas ao impacto acadêmico, baseadas em ferramentas e ambientes online. Os autores afirmaram que essas áreas abordadas na coleção

irão publicar novas pesquisas continuamente, e esperamos ver contribuições adicionais aparecerem nos próximos meses. Nós estamos ansiosos para construir uma base de pesquisa inicial para apoiar este novo campo. (PRIEM; GROTH; TARABORELLI, 2012, p. 2)

Os estudos defensores do campo da altmetria também citam várias razões que exigem (e justificam) a criação de novas métricas (GALLIGAN; DYAS-CORREIA, 2013): revisão por pares, contagem de citações e fatores de impacto de periódicos têm sido tradicionalmente usados como um meio de verificar o valor do trabalho acadêmico e como uma forma de filtrar apenas o material mais significativo e relevante do enorme volume de literatura acadêmica produzida. Afinal, as métricas tradicionais lidam apenas com as revistas e artigos científicos, e não avaliam outras formas de comunicação científica online importantes para o diálogo acadêmico. Inicia-se então uma tendência de estudos altmétricos baseados nos debates anteriores sobre a análise de citações, e discutindo as estratégias de avaliação altmétrica, como testes de correlação e análises de conteúdo (SUD; THELWALL, 2014).

Com o crescente uso da altmetria na avaliação do impacto acadêmico online, começa a surgir a necessidade de selecionar os indicadores altmétricos que deverão ser adotados para

que os editores possam escolher qual usar em seus sistemas, uma vez que os resultados obtidos a partir desses indicadores também serão relevantes para outros participantes do processo de produção de conhecimento, como para as agências de fomento ao direcionar recursos para pesquisas. Busca-se então comparar os indicadores altmétricos entre si (BORNMANN, 2015; PRIEM; GROTH; TARABORELLI, 2012; YAN; GERSTEIN, 2011; YU *et al.*, 2021) e analisar suas coberturas (COSTAS; ZAHEDI; WOUTERS, 2014; HAUSTEIN; COSTAS; LARIVIÉRE, 2015; ORTEGA, 2018; ROBINSON-GARCÍA *et al.* 2014; THELWALL *et al.*, 2013).

Entre os estudos altmétricos produzidos recentemente, as análises variam tanto em busca de demonstrar a importância da altmetria para a mensuração do impacto acadêmico, quanto para compará-la com a bibliometria (COSTAS; ZAHEDI; WOUTERS, 2014; HAUSTEIN; COSTAS; LARIVIÉRE, 2015; HUANG; WANG; WU, 2018; MOHAMMADI *et al.*, 2015; OUCHI *et al.*, 2019; ZAHEDI; COSTAS; WOUTERS, 2014). Quanto à comparação da altmetria com a bibliometria, é necessário primeiramente compreendermos como essas duas métricas podem se complementar e os benefícios para a avaliação da informação trazido por essa combinação.

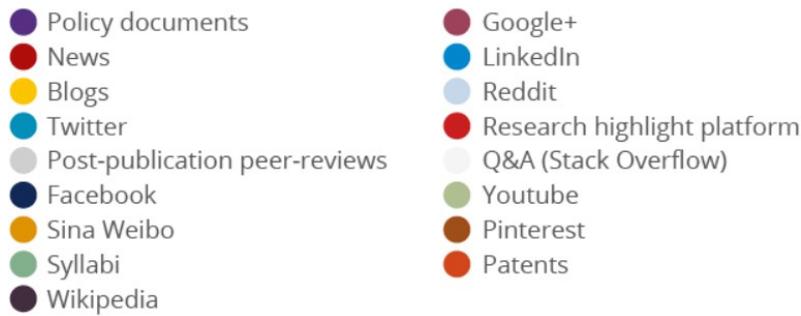
2.5.1 Fontes de dados altmétricos

Entre as diversas fontes de dados altmétricos existentes, serão apresentadas a seguir as mais relevantes dentro do escopo deste estudo, sendo elas: *Altmetric*, *Twitter*, *Mendeley* e *Open Journal Systems*. A plataforma agregadora de dados altmétricos *Altmetric* consegue quantificar o impacto de qualquer publicação científica com um DOI ou outro identificador padrão. Ele pode encontrar menções de livros, redes sociais como *Twitter*, *Facebook*, *Pinterest*, *blogs* de ciência, vários veículos de mídia convencionais, como jornais, documentos de política governamental e de gerenciadores de referência online para menções de trabalhos acadêmicos.

Um recurso que se tornou a marca registrada do *Altmetric* é o “*donut*”, que é a forma de visualização do *Altmetric Attention Score (AAS)*, um indicador altmétrico calculado com base na atribuição de pesos para as ocorrências de uma publicação, a depender da fonte onde foi detectada a menção. As cores que compõem o *donut* refletem a mistura de fontes que mencionam a publicação, conforme legenda disponível no próprio site da plataforma e ilustrada na Figura 5.

Figura 5 - Legenda de cores de acordo com a fonte que compõem o donut do *Altmetric Attention Score*

The Colors of the Donut



Fonte: Altmetric (2021).

Uma questão que diversos estudos de indicadores altmétricos encaram é a possibilidade de utilizar ou não o *AAS* como indicador altmétrico. A maioria dos estudos optam por utilizar os dados altmétricos coletados do *Altmetric* separados por fonte de dados (BORNMANN, 2014; THELWALL *et al.*, 2013), ou o somatório desses dados (COSTAS; ZAHEDI; WOUTERS, 2014). Outros, utilizam o *AAS* calculado pelo *Altmetric* (EZEMA; UGWU, 2019; HUANG; WANG; WU, 2018;) ou ambas as abordagens (OUCHI *et al.*, 2019). Entretanto, a maioria dos estudos de indicadores altmétricos até então utilizam os dados coletados do *Altmetric* separando-os por fonte de dados ou somando-os.

O *Altmetric* é uma ferramenta de assinatura com base em taxas. É voltado para editores e instituições com vários planos e preços. A *Nature Publishing Group*⁹ é um exemplo de editora que incorporou as informações do *Altmetric.com* em seu site. Desde o início de 2012, o *Altmetric* monitora uma série de fontes de dados não tradicionais, em busca de links e referências para pesquisas publicadas. Hoje, o banco de dados do *Altmetric* contém mais de 191 milhões de menções e mais de 35 milhões de resultados de pesquisa, incluindo artigos de periódicos, conjuntos de dados, imagens, relatórios, e outros tipos de dados, e está em constante crescimento.

Para identificar e cruzar os dados de publicações com as menções nas diversas mídias sociais, o *Altmetric* utiliza identificadores, que também ajudam a plataforma a reconhecer diferentes versões de uma mesma pesquisa. Assim, as informações do *Altmetric* são limitadas a documentos científicos que possuam um *DOI*, um registro *PubMed ID*, um *ArXiv ID*¹⁰, ou outro identificador padrão (ERFANMANESH, 2017). O número de *DOI* que uma publicação científica recebe ao ser publicada é um identificador que possibilita, portanto, o rastreamento

⁹ Disponível em: <https://www.nature.com/>. Acesso em: 12 dez. 2021.

¹⁰ Disponível em: https://arxiv.org/help/arxiv_identifier. Acesso em: 12 dez. 2021

da pesquisa e o cruzamento de dados entre as fontes de dados para que o *Altmetric* contabilize as menções e as associe à publicação.

Sun et al. (2018) afirmam que os *tweets* são considerados, para a altmetria, potenciais indicadores importantes do impacto social imediato de artigos científicos. Devido aos resultados de pesquisas anteriores que demonstraram, em todos os casos, que depois do *Mendeley*, a rede social com maior cobertura altmétrica é o *Twitter* (COSTAS, ZAHEDI, WOUTERS, 2014; HAUSTEIN, COSTAS, LARIVIÈRE, 2015; ROBINSON-GARCÍA *et al.*, 2014), é interessante considerar estas duas fontes de dados altmétricos como as principais em estudos futuros.

O *Mendeley* é um gerenciador de referências online gratuito que permite aos usuários armazenar, organizar, anotar, compartilhar e citar referências e dados de pesquisa (MENDELEY, 2021). Possui funcionalidades como gerar bibliografias automaticamente, permitir a colaboração com outros pesquisadores online, encontrar artigos relevantes com base no que o usuário está lendo, e permite ao usuário acessar seus documentos de qualquer lugar online. Sua base de dados possui mais de 29,5 milhões de publicações científicas indexadas e mais de oito milhões de usuários cadastrados (ATKISON-BONASIO, 2018).

Estudos anteriores mostraram correlações moderadas entre leitores do *Mendeley* e o número de citações para publicações científicas (ZAHEDI; COSTAS; WOUTERS, 2017). As correlações entre leitores do *Mendeley* e citações são sempre maiores e mais fortes do que entre citações e outros indicadores altmétricos (THELWALL *et al.*, 2013), nos permitindo assumir que há uma similaridade maior entre estes dois indicadores do que entre citações e indicadores provindos de outras fontes de dados altmétricos.

Além disso, publicações com mais pontuações de leitores de *Mendeley* tendem a ter maior número de citações e são publicadas em periódicos de maior impacto em comparação com aqueles com menos ou sem leitores (ZAHEDI; COSTAS; WOUTERS, 2014). Uma característica importante do *Mendeley* é que suas estatísticas de leitores incluem dados sobre o status acadêmico, disciplinas e países dos usuários. Esse tipo de informação possibilita uma melhor compreensão dos padrões de salvamento de publicações científicas por diferentes grupos de usuários.

Além das vantagens e potencialidades do *Mendeley* supracitadas, os baixos números de cobertura altmétrica identificados nas mídias sociais como *Twitter*, *Facebook* e blogs contrastam com os resultados observados em estudos que analisam os leitores do *Mendeley*, onde as coberturas identificadas foram consideravelmente maiores, acima de 60% (HAMMARFELT, 2014; HAUSTEIN; LARIVIÈRE, 2014; ZAHEDI; COSTAS; WOUTERS,

2014), indicando o *Mendeley* como a maior fonte de altmétricas e uma ferramenta promissora de avaliação de pesquisa (LI; THELWALL; GIUSTINI, 2012).

O *Open Journal Systems (OJS)* é um software livre para gestão e configuração de revistas científicas. Originalmente desenvolvido e lançado pela *PKP* em 2001 para melhorar o acesso à pesquisa, o *OJS* é a plataforma de código aberto de publicação de periódicos mais amplamente usada, com mais de 25.000 periódicos usando-a em todo o mundo. De acordo com o projeto desenvolvedor do *OJS*, o *Public Knowledge Project (PKP)*, o *OJS* é disponibilizado gratuitamente para periódicos em todo o mundo com o propósito de tornar a publicação em acesso aberto uma opção viável para mais periódicos, uma vez que o acesso aberto traz benefícios para o periódico podendo, por exemplo, aumentar o número de leitores, bem como a contribuição do periódico para o bem público em escala global, em consonância com as diretrizes da iniciativa *OA*. O próprio *PKP* é uma iniciativa multi universitária que desenvolve softwares de código aberto e conduz pesquisas para melhorar a qualidade e o alcance da publicação acadêmica (*PKP*, c2021b).

Seguindo, portanto, a tendência mundial de utilização do *OJS* para revistas de acesso aberto, a recomendação de software do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) para periódicos científicos é a plataforma *OJS*. O IBICT, como unidade de pesquisa do Ministério da Ciência, Tecnologia e Inovação e órgão nacional de informação, realiza estudos no campo da ciência da informação e temas relacionados, dando apoio a diversas instituições de ensino superior na gestão e inovação de seus acervos, pesquisas e softwares. Nesse sentido, o IBICT disponibiliza guias, capacitação e suporte operacional ao *OJS* para universidades brasileiras, contribuindo para a adesão de revistas e portais de periódicos de universidades brasileiras de acesso aberto a utilizarem o *OJS* como plataforma online de publicação acadêmica e, por consequência, ajudando na disseminação da produção científica brasileira na web.

O *OJS* é uma ferramenta abrangente para gerenciar toda a submissão, fluxo de trabalho editorial e publicar artigos científicos e edições de revistas online. Alguns dos recursos interessantes que ele oferece são:

- fluxo de trabalho editorial flexível e configurável;
- submissão online e gerenciamento de todo o conteúdo;
- módulo de assinatura com opções de acesso aberto retardado;
- integrado com serviços de publicação acadêmica, como *Crossref*, *ORCID*.
- geração de relatórios para monitoramento de ações no sistema, como usos, acessos a publicações, *logs*, avaliação, submissões de documentos, dentre outros.

Entre as diversas funcionalidades e ferramentas do *OJS* disponíveis para auxiliar os agentes científicos no processo de comunicação científica, o sistema *OJS* armazena informações sobre os documentos submetidos nos periódicos hospedados na plataforma, e disponibiliza estas informações na forma de relatórios, com diversas possibilidades de filtros e seleção de metadados a serem exibidos nos mesmos.

Um dos relatórios relevantes para estudos de indicadores altmétricos disponível no *OJS* é o relatório de acesso, que possui informações de cada publicação tais como identificador (ID) do artigo, título, edição, data de publicação, acessos ao resumo, e número de downloads do arquivo *PDF* do documento. Nesse sentido, diversos estudos (LIU *et al.*, 2013; MOED, 2005; PERNEGER, 2004) analisaram correlações entre indicadores altmétricos de acesso e downloads e o indicador bibliométrico de citação, identificando correlações moderadas e concluindo que indicadores de acesso e downloads são importantes para entender estas relações.

2.6 ESTUDOS SOBRE COBERTURAS DE FONTES DE INFORMAÇÃO BIBLIOMÉTRICAS E ALTMÉTRICAS

A cobertura é uma medida que permite avaliar a proporção com que determinada amostra de dados está presente em uma determinada fonte. A cobertura de citação permite a identificação de quais fonte de dados bibliométricos abrangem a maior parte da produção acadêmica, de acordo com a delimitação da amostra pré-definida na análise. Por exemplo, Noris e Oppenheim (2007) analisaram comparativamente as coberturas das bases *Web of Science*, *Scopus*, Google Scholar, e CSA Illumina para a literatura da área de Ciências Sociais. Neste estudo, os autores identificaram que a base de dados com maior cobertura a nível de artigo não é, necessariamente, a de maior cobertura a nível de revista. Esse tipo de resultado demonstra que, devido a questões inerentes aos critérios utilizados por cada base de dados para indexação do conteúdo, a cobertura para uma determina amostra pode variar, dependendo do nível de análise utilizado na abordagem.

Desse modo, com o objetivo de analisar a cobertura das fontes de dados para construção de indicadores altmétricos e bibliométricos, faz-se necessário identificar na literatura científica as potencialmente melhores fontes do ponto de vista de maior cobertura a nível de artigo. Nesse sentido, Bornmann (2014) selecionou 104.633 artigos da base F1000, de literatura biomédica, que possuísem número de *DOI*. Destes, 91% (n=94.641) possuíam

algum dado bibliométrico disponível, e destes, apenas 69% (n=65.535) possuíam algum dado altmétrico coletado pela plataforma *Altmetric*.

Li, Thelwall e Giustini (2012) identificaram, para artigos da revista *Nature* publicados em 2007, uma média de 90,88 citações no *Google Scholar*, 78,06 citações na *Web of Science* e 10,71 de leitores no *Mendeley*, com médias um pouco menores para artigos da revista *Science* publicados no mesmo ano. Em termos de cobertura, apenas 8% dos artigos não foram lidos por algum usuário no *Mendeley*, um indício da amplitude de uso desta base.

Para mais de 1,3 milhões de artigos indexados na *Web of Science* em 2012 de cinco áreas de conhecimento distintas, Haustein, Costas e Larivière (2015) encontraram coberturas de citações entre 45,8% e 74,3%, sendo a cobertura mais baixa para a área de Ciências Sociais e Humanidades, e a mais alta para a área de Ciências da Vida e da Terra, e uma média geral de 66,8%. Entre as mídias sociais analisadas, as maiores coberturas encontradas foram do *Twitter*, com a menor cobertura para a área de Ciências da Computação e Matemática, com 7,5%, e a maior cobertura para Ciências da Saúde e Biomédicas, com 31,7%.

As médias encontradas por Haustein, Costas e Larivière (2015) em todas as áreas foram inferiores às encontradas anteriormente por Li, Thelwall e Giustini (2012), com a maior média de citação (3,83) para a área de Ciências Naturais e Engenharia. As diferenças entre as coberturas e médias entre estes dois estudos pode estar relacionada à reputação das revistas *Nature* e *Science*, bem como ao fato do intervalo de tempo entre a coleta dos dados e a publicação dos artigos ser menor no estudo de Haustein, Costas e Larivière (2015), fator cujo impacto será abordado mais à frente nesta pesquisa.

Ao analisar 389.795 publicações do portal de periódicos latino-americano *SciELO* e coletar os indicadores altmétricos da plataforma *Altmetric*, Alperin (2015) identificou que a fonte altmétrica com mais menções é o *Mendeley* (80%), seguido distantemente pelo *Twitter* (8%). Comparativamente com resultados de estudos focados em fontes como *Web of Science*, *Scopus* e periódicos e portais consolidados internacionalmente, Alperin (2015) identifica que a diferença encontrada nas coleções da América Latina é que a cobertura dos indicadores tende a ser um pouco mais baixa, podendo essa baixa cobertura ter três explicações plausíveis:

- 1) os artigos publicados na SciELO simplesmente têm menor uso de modo geral, e isso se reflete em um uso menor das mídias sociais; 2) os níveis de uso de mídias sociais, pelo menos entre acadêmicos, é menor na América Latina do que em contextos estudados anteriormente; ou 3) existe uma cultura diferente na prática de compartilhar pesquisas nas mídias sociais entre os pesquisadores latino-americanos. (ALPERIN, 2015, p. 15, tradução nossa).

Mohammadi *et al.* (2014) analisaram a cobertura no *Mendeley* para 480.979 publicações indexadas na *Web of Science* em 2008. Como resultado, obtiveram uma cobertura geral de 45,6% e variações na cobertura por área, com artigos da área de Medicina Clínica alcançando a maior cobertura, de 71,6%, enquanto para Física e Química as coberturas encontradas foram as mais baixas, de 31,4% e 33,7%, respectivamente. Em um estudo similar, voltado para publicações das áreas de Ciências Sociais e Humanidades, Mohammadi e Thelwall (2014) encontraram uma cobertura no *Mendeley* de 58% para Ciências Sociais e 28% para Humanidades para publicações também indexadas na *Web of Science* em 2008.

Haustein *et al.* (2014) identificaram, para uma amostra de mais de 1,4 milhões de publicações de diferentes áreas indexadas no PubMed, que 66,2% destes foram lidos no *Mendeley*, e apenas 9,1% foram tweetados. Em uma análise por área a maior cobertura no *Mendeley*, de 81%, foi para artigos de Psicologia, e a menor, de 40,7%, para Humanidades. A média de leitores por artigo foi de 9,7, e a média de *tweets* foi de 2,5. A baixa cobertura para a área de Humanidades pode ser justificada pela escolha da *PubMed*, uma base voltada para artigos com foco em Biomedicina, indexando apenas aqueles que são representativos para a área de alguma forma, como fonte de dados. Entretanto, na análise geral das coberturas, os resultados refletem números similares aos de estudos anteriores.

É possível observar que a maioria dos estudos indica o *Mendeley* como uma fonte de dados altmétricos de cobertura satisfatória, superior a mídias sociais como o *Twitter*. Entretanto, no estudo conduzido por Ortega (2018), 67.147 artigos científicos coletados da base *PlumX* tiveram as coberturas dos indicadores altmétricos nas bases *PlumX* e *Altmetric* analisadas. Nesse estudo a cobertura do *Twitter* identificada pela plataforma *Altmetric* se mostrou elevada, de 91,38%, e muito próxima da cobertura de 95,76% do *Mendeley*. Essas altas coberturas se devem ao fato da coleta de dados ter sido feita diretamente de uma fonte de dados altmétricos que possui, por pressuposto, apenas informações de publicações que tiveram alguma menção altmétrica rastreável e coletada por esta fonte.

Robinson-García *et al.* (2014) analisaram publicações indexadas na *Web of Science* com DOI durante o período de 2011 a 2013 e coletaram seus dados da plataforma *Altmetric*. 19% do conjunto original de publicações possuía algum dado altmétrico recuperado pelo *Altmetric*. Dentre as 16 mídias sociais identificadas como fontes de dados utilizadas pelo *Altmetric*, apenas cinco cobriam 95,5% do total de menções altmétricas recuperadas. Destas, o *Twitter* (87,1%) e o *Mendeley* (64,8%) tiveram as maiores coberturas. Deve-se ressaltar que estas coberturas identificadas por Robinson-García *et al.* (2014) dizem respeito às publicações da amostra reduzida de 19% que tiveram alguma menção altmétrica coletada pelo *Altmetric*, e

não à amostra geral. Na amostra geral, a cobertura do *Twitter* cai para 16%, e a do *Mendeley* para 12%. A cobertura baixa de leitores no *Mendeley* identificada por Robinson-García *et al.* (2014) se deve ao fato da coleta das métricas ter sido feita através da plataforma *Altmetric*, e não diretamente pela *API* do *Mendeley*.

Zahedi, Costas e Wouters (2014) também analisaram as coberturas para publicações indexadas na *Web of Science*. A maior cobertura encontrada foi para o *Mendeley*, com ao menos um leitor para 62,6% das publicações. Já o *Twitter* obteve uma cobertura de apenas 1,6%. Em outro estudo mais recente, Zahedi e Costas (2018) analisaram a cobertura de publicações do portal de periódicos da PLOS em diferentes fontes altmétricas, como *PlumX*, Lagotto ([202-]), *Mendeley* e *Altmetric*. Os principais resultados apontam para uma cobertura de 99,9% na *PlumX*, seguida pela Lagotto (99,8%), *Mendeley* (95,9%) e *Altmetric* (61%). Comparando exclusivamente o indicador altmétrico de leitores no *Mendeley* para publicações que tiveram ao menos um leitor registrado em alguma das fontes, a cobertura foi de 95,8% para o próprio *Mendeley* e 60,6% para leitores do *Mendeley* coletados pelo *Altmetric*. Isso demonstra que a política de coleta de dados do *Altmetric* influencia na cobertura do indicador, sendo recomendável a coleta de números de leitores no *Mendeley* diretamente nessa plataforma, através de sua *API*. Quanto a cobertura no *Twitter*, Zahedi e Costas (2018) identificaram que o *Altmetric* exibe a maior cobertura (57%) de publicações com ao menos um *tweet* entre todas as bases analisadas.

Numa análise de cobertura das plataformas para publicações indexadas no *Altmetric*, Costas, Zahedi e Wouters (2014) identificaram que o *Twitter* possui a maior cobertura entre as fontes analisadas, com 13,3% de todas as publicações. As demais coberturas por plataforma podem ser analisadas na Figura 6. Os resultados obtidos por Costas, Zahedi e Wouters (2014) concordam com os resultados obtidos por Thelwall *et al.* (2013), que também obtiveram baixas coberturas de indicadores altmétricos para publicações da PLOS em mídias sociais.

Figura 6 - Distribuição geral das publicações com altmétricas

	Docs.	% within altmetrics	%pubs
Total altmetrics	75,569	100%	15.1%
Facebook walls	12,386	16.4%	2.5%
Blogs	9,444	12.5%	1.9%
Twitter	66,591	88.1%	13.3%
Google+	3,021	4.0%	0.6%
News outlets	2,331	3.1%	0.5%

Fonte: Costas, Zahedi e Wouters (2014, p. 5)

Em uma análise por ano de publicação, para 20.000 publicações coletadas do *ImpactStory* ([202-]) publicadas entre 2005 e 2011, Zahedi, Costas e Wouters (2014) identificaram que a principal fonte de dados altmétricos foi o *Mendeley*, com uma cobertura de 80% para publicações de periódicos multidisciplinares, e apenas 7% de cobertura no *Twitter*. Estes resultados reforçam os obtidos por estudos anteriores (ALPERIN, 2015; HAUSTEIN *et al.*, 2014; ZAHEDI; COSTAS; WOUTERS, 2014), reafirmando o *Mendeley* como a fonte de dados altmétricos de maior cobertura, seguido pelo *Twitter*.

Quanto aos indicadores bibliométricos, numa análise detalhada e comparativa de fontes de dados bibliométricas, Martín-Martín *et al.* (2021) identificaram o Google Scholar como a base de maior cobertura para publicações de oito grandes áreas de conhecimento, com coberturas variando entre 85% e 90%. As menores coberturas encontradas foram para as citações abertas de DOI para DOI da *Crossref* (COCI). As demais fontes de dados e suas coberturas podem ser visualizadas na Figura 7.

Figura 7 - Porcentagens de citações encontradas por área de conhecimento

	N	Google Scholar	Microsoft Academic	Scopus	Dimensions	Web of Science	COCI
Humanities, Literature & Arts	89,337	87%	39%	31%	29%	25%	18%
Social Sciences	406,661	88%	47%	40%	36%	33%	20%
Business, Economics & Management	235,338	88%	47%	34%	32%	29%	19%
Engineering & Computer Science	691,164	88%	63%	61%	54%	48%	30%
Physics & Mathematics	317,320	90%	57%	64%	59%	59%	36%
Health & Medical Sciences	1,001,507	85%	63%	59%	58%	51%	27%
Life Sciences & Earth Sciences	571,817	89%	68%	64%	63%	60%	32%
Chemical & Material Sciences	253,990	90%	69%	75%	72%	72%	32%

Fonte: Martín-Martín *et al.* (2021, p. 885).

A baixa cobertura de citações na fonte *Crossref* pode estar relacionada a forma de coleta dos dados, feita através da lista de referências abertas disponibilizada pela *Crossref*. Martín-Martín *et al.* (2021) ressaltam que, no momento da escrita de sua pesquisa, apenas 59% dos artigos com referências depositados na *Crossref* tinham suas referências abertas.

2.7 ESTUDOS SOBRE CORRELAÇÕES ENTRE INDICADORES BIBLIOMÉTRICOS E ALTMÉTRICOS

Os indicadores altmétricos podem ser utilizados em conjunto com os bibliométricos, trazendo benefícios na avaliação do impacto das publicações científicas. De acordo com Watson (2009) e Wang *et al.* (2014) os indicadores altmétricos, de acesso e de download podem ajudar a prever citações. Sendo assim, é de interesse de toda comunidade acadêmica (e não apenas da área de estudos métricos da informação) a análise das correlações entre estes indicadores e o indicador bibliométrico de citação, objetivando obter resultados que reafirmem os resultados desses estudos e identificar quais as melhores e mais consistentes fontes de indicadores altmétricos para prever citações.

Além da expectativa dessa previsão, mensurada na força da correlação entre os indicadores altmétricos e bibliométricos, conhecer estes indicadores e suas correlações é de interesse dos portais de periódicos de acesso aberto, para direcionar suas estratégias na tomada de decisões que favoreçam a ampla divulgação de suas publicações científicas com vistas a aumentar a atenção social recebida.

Para entender melhor as possibilidades e as limitações de uso dos indicadores altmétricos e de citação, é preciso realizarmos análises comparativas entre eles, e a partir destas análises é que será possível identificar se há influência ou relação entre os indicadores altmétricos e os bibliométricos de citação. De acordo com Costas, Zahedi e Wouters (2014):

O desenvolvimento do conceito de altmetria foi acompanhado por um crescimento na diversidade de ferramentas baseadas na web com o objetivo de capturar e rastrear uma ampla gama de resultados de um pesquisador, agregando dados altmétricos em uma ampla variedade de fontes. O uso predominante da web social por estudiosos também levou a análises de altmetria e sua relação com métricas de impacto já estabelecidas, como a análise de citações. A maioria desses estudos encontrou algum grau de correlação entre a altmetria e o número de citações, sugerindo que essas duas abordagens estão de alguma forma relacionadas, mas não são a mesma coisa. (COSTAS; ZAHEDI; WOUTERS, 2014, p. 2004, tradução nossa)

Pesquisadores e acadêmicos costumam ter interesse em medir o impacto que suas descobertas tiveram (BRIGHAM, 2014). Não apenas no seu próprio campo, mas também em um contexto mais amplo da comunidade científica e nas próprias mídias. Devido, portanto, à necessidade de mensurar o impacto das publicações científicas online nas diversas redes e mídias sociais existentes (não apenas nos canais formais de comunicação científica) e a possibilidade dos indicadores altmétricos preverem citações futuras, o uso de indicadores altmétricos se apresenta como uma solução para quantificar o impacto da produção acadêmica, indo além das potencialidades dos indicadores bibliométricos tradicionais e suas métricas baseadas em citação.

Diversos estudos buscaram explorar a possibilidade de predição de citações da altmetria, mas seus resultados foram inconclusivos, contraditórios ou pouco promissores (ROEMER; BORCHARDT, 2015). Porém, cada vez mais, estudos recentes encontram correlações moderadas entre indicadores altmétricos e bibliométricos, nas mais variadas fontes de dados, e alguns resultados passam a ser mais recorrentes do que outros, mostrando quais são as possíveis fontes de dados altmétricos mais promissoras ao serem correlacionadas com citações.

A cada estudo de correlações entre indicadores altmétricos e bibliométricos concluído, os pesquisadores apontavam a necessidade de desenvolvimento de outros estudos em larga escala que complementassem os anteriores e que combinassem abordagens qualitativas e quantitativas. Ao analisar qualitativamente os estudos altmétricos desenvolvidos no Brasil, Gouveia (2019) identifica que uma mudança no perfil dos artigos publicados sobre altmetria vem ocorrendo ao longo do tempo: inicialmente, eram feitos estudos teóricos propondo a altmetria como alternativa às citações.

Mais recentemente, as análises dos estudos são específicas por plataformas e fontes de dados, e envolvendo o uso de indicadores. Dessa forma, o enfoque nas correlações entre os indicadores e a análise deles por plataforma são uma tendência dos estudos da área da altmetria que se justifica pela necessidade de um melhor entendimento da complementaridade entre os indicadores altmétricos e bibliométricos e o potencial desta para a avaliação da ciência em geral e, mais especificamente, a publicada em portais de periódicos brasileiros de acesso aberto.

O uso dos indicadores altmétricos para medir o impacto social deve ser avaliado antes de colocado em prática. Para isso, um conjunto de estratégias foram adotadas para avaliar os indicadores altmétricos (SUD; THELWALL, 2014) e, entre elas, a mais popularmente utilizada é a correlação entre os indicadores altmétricos e a contagem de citações (THELWALL, 2020), sendo este um indicador bibliométrico tradicional. De acordo com Thelwall (2020), essa estratégia consiste em avaliar os indicadores altmétricos e bibliométricos de citação com um teste de correlação,

[...] para saber se eles fornecem informações sobrepostas. Este teste é justificado com base em que: (a) quase qualquer indicador de impacto genuíno deve se correlacionar com as contagens de citações, uma vez que, outros fatores sendo iguais, as pesquisas mais impactantes têm uma maior probabilidade de atrair citações de estudos subsequentes; e (b) correlações estatisticamente significativas são evidências de que os indicadores altmétricos, pelo menos, não são aleatórios. (THELWALL, 2020, p. 3, tradução nossa).

Os estudos abordando correlações entre os indicadores altmétricos e bibliométricos de citação encontraram resultados interessantes, como a identificação das melhores fontes de dados de indicadores altmétricos que se correlacionam mais fortemente com as citações, resultados estes que guiaram a escolha de algumas fontes utilizadas na presente pesquisa. Mas antes de falarmos das fontes de dados, alguns estudos de correlações entre os indicadores altmétricos e bibliométricos de citação merecem destaque.

Um dos primeiros estudos a correlacionar indicadores altmétricos e bibliométricos de citação foi o de Pernerger (2004), antes mesmo de Jason Priem nomear o campo da altmetria como tal. Pernerger (2004), ao analisar a relação entre número de acessos e citações subsequentes para 153 publicações da revista BMJ, identificou que os artigos que atraíram mais acessos na primeira semana após a publicação eram subsequentemente mais citados do que artigos acessados com menor frequência. A correlação de Pearson obtida pelo autor entre os acessos e citações para essa pequena amostra foi de 0,54, indicando uma correlação moderada entre os indicadores. O autor propõe a hipótese de que:

[...] o ‘valor científico’ explica a associação entre os acessos e as citações, uma vez que os leitores online julgam o valor científico de um artigo baseado no título e resumo no *website* da revista, e se estes forem satisfatórios, então eles acessam o artigo completo. O valor científico de um artigo também leva à citação por outros pesquisadores. Essa hipótese é suportada pela grande frequência de acessos e citações para artigos que usam métodos de estudo cientificamente rigorosos. (PERNEGER, 2004, p. 2, tradução nossa).

Pernerger (2004), desde então, mesmo com um estudo de uma pequena amostra de dados providos de uma mesma revista, já identificava haver uma relação (pelo menos dentro da revista BMJ), entre acessos e citações. Inclusive, o autor encoraja a divulgação dos indicadores de acesso a publicações científicas pelas editoras, prática que se mostra estratégica para aumentar a atenção social recebida.

Um estudo semelhante foi conduzido por Moed (2005), porém analisando comparativamente o indicador de downloads com as citações extraídas do *Science Citation Index (SCI)*, para 1.190 documentos publicados na revista *Tetrahedron Letters*, da editora *Elsevier*. O índice de correlação de *Spearman* encontrado entre os indicadores foi de 0,22 para a amostra total de publicações, porém, quando downloads considerados recentes (aqueles executados durante os três meses iniciais após a data de publicação do artigo) foram descartados da amostra, a correlação subiu para 0,35, mas mantendo-se, ainda assim, fraca. O processo inverso ocorreu quando o autor correlacionou estes downloads recentes com as citações mais antigas, com o índice de correlação caindo para 0,11.

Moed (2005) interpretou estes resultados como um indicativo de que downloads recentes e citações são relativos a diferentes fases do processo de coletar e processar informação científica relevante, que eventualmente irá levar à publicação de um novo artigo científico (MOED, 2005). Porém, ao analisar os downloads de uma publicação nos três meses que se seguem após esta ser citada, o autor identificou um aumento de 25% no número de downloads, comparativamente ao esperado para publicações que não foram citadas recentemente, o que aponta para a hipótese de que foi a citação que ocasionou o aumento nos downloads da publicação. Entretanto, essa influência da citação nos downloads não ocorre para artigos de publicação recente. Esse comportamento observado nas correlações e nos indicadores nos mostra que as fases temporais pelas quais uma publicação passa, são importantes para entender as relações entre os indicadores de download e de citação.

Priem, Piwowar e Hemminger (2011) buscaram responder se as altmétricas poderiam prever citações. Para isso, a partir de 24.334 artigos publicados até o final de 2010 no portal de periódicos da *Public Library Of Science (PLOS)*, de sete revistas de áreas de conhecimento distintas, foram coletados indicadores bibliométricos de citação de diversas bases de dados, como *Web of Science*, *Scopus*, *Crossref*, e *PubMed Central*. Indicadores altmétricos também foram coletados de diferentes fontes, como *Wikipedia*, *blogs*, comentários no próprio portal da *PLOS* e redes sociais como *Mendeley*, *CiteULike*, *Twitter* e *Facebook*. Os autores obtiveram correlações muito variadas entre os indicadores, com destaque para a correlação entre os indicadores provenientes do *Mendeley* e as citações ($r = 0,26$), levando-os a discutir a importância de pesquisas futuras utilizarem indicadores provindos de mídias sociais de uso acadêmico, como o *Mendeley* e *Twitter*, e que as amostras de publicações sejam ainda mais diversas (PRIEM; PIWOWAR; HEMMINGER, 2011).

Com o foco dos estudos altmétricos se voltando para indicadores provenientes de redes sociais de uso acadêmico, Eysenbach (2011) analisou todos os *tweets* contendo links para artigos da revista *Journal of Medical Internet Research* publicados em edições dos anos de 2009 e 2010 e comparou-os com os números de citações das bases Google Scholar e *Scopus*. Essa foi a primeira pesquisa sistemática, prospectiva e longitudinal executada a nível de artigo e a nível de revista investigando como citações e *tweets* de artigos científicos se acumulam ao longo do tempo.

Eysenbach dividiu as amostras agrupando os *tweets* de acordo com o número da edição ao qual o artigo tweetado pertence, para então correlacionar com as citações. As correlações de Pearson se mostraram moderadas na maioria dos casos, variando entre 0,42 e 0,72. As correlações se mostraram fortes o suficiente para que o autor fizesse previsões

baseadas em modelos multivariados e testes binários, chegando à conclusão de que artigos altamente tweetados possuem onze vezes mais chances de serem mais citados do que artigos menos tweetados, e que os *tweets* podem prever os artigos que serão mais citados nos primeiros três dias após a publicação do artigo, sendo recomendável utilizar medidas de impacto social baseadas em *tweets* para complementar as métricas tradicionais baseadas em citação (EYSENBACH, 2011).

Li, Thelwall e Giustini (2012) buscaram contribuir com as investigações referentes ao uso de indicadores altmétricos para medir o impacto acadêmico, coletando indicadores do *Mendeley* e do *CiteULike* para 1.613 publicações das revistas *Nature* e *Science*, e com números de citações extraídos da base *Web of Science*. As correlações de *Spearman* obtidas entre o indicador de leitores do *Mendeley* e as citações foram moderadas para ambas as revistas, variando entre 0,540 e 0,603, enquanto as correlações entre o indicador do *CiteULike* e as citações se mantiveram em torno de 0,36, sendo correlações significativamente mais fracas. Os autores são cautelosos em tirar conclusões, afirmando que, apesar dos resultados obtidos,

correlações estatisticamente significativas entre duas fontes de dados nunca provam um relação causal, então não podemos concluir que o alto impacto da pesquisa, medido por citações, causa alta contagem de usuários do gerenciador de referências online para artigos (*Mendeley*). Assumir que o fato dos usuários salvarem artigos em suas contas do *Mendeley* pode indicar que o artigo tem algum valor é consistente com a revisão da literatura, mas permanece uma suposição não totalmente comprovada. (LI; THELWALL; GIUSTINI, 2012, p. 468, tradução nossa)

Os resultados de Li, Thelwall e Giustini (2012) começam a apontar o gerenciador de referências online *Mendeley* como uma fonte de dados relevante nas análises futuras nas correlações, inclusive com a recomendação dos autores de que análises semelhantes ampliem o alcance das amostras para outras áreas do conhecimento e períodos de tempo. Esse tipo de estudo contribui para confirmar a suposição de Li, Thelwall e Giustini (2012) de que a ação executada por leitores no *Mendeley* de salvar um artigo na sua biblioteca de referências pode significar que este artigo possua relevância para futuros trabalhos científicos do leitor, possibilitando que esse indicador venha a fazer parte dos métodos de avaliação de pesquisa.

Bar-Ilan *et al.* (2012) analisaram publicações de 57 pesquisadores da Ciência da Informação e Biblioteconomia, com indicadores altmétricos provindos do *Mendeley* e *CiteULike*, e indicadores bibliométricos de citação extraídos do *Scopus*. A correlação de *Spearman* encontrada entre leitores do *Mendeley* e as citações foi moderada, de 0,448, enquanto a correlação entre as marcações no *CiteULike* e as citações foi fraca, de apenas

0,232, reforçando os resultados encontrados por Li, Thelwall e Giustini (2012). Assim como dito por Li, Thelwall e Giustini (2012), os autores reforçam que é necessário ter cautela na interpretação dos resultados das correlações, uma vez que é preciso primeiramente compreender melhor como e por que os usuários de plataformas como o *Mendeley* usam estas ferramentas (BAR-ILAN *et al.*, 2012).

Liu *et al.* (2013) executaram uma pesquisa com uma amostra consideravelmente maior que de estudos anteriores, correlacionando indicadores altmétricos e bibliométricos para 33.128 artigos científicos do portal de periódicos da editora *PLOS*, com indicadores altmétricos extraídos do próprio portal, como downloads, acesso à página do artigo, e citações nas bases *Crossref*, *Scopus* e *PubMed Central*. Os autores utilizaram a correlação de Spearman por esta ser mais recomendável do que a correlação de Pearson no caso de distribuições anormais de variáveis. As correlações entre o indicador altmétrico de acesso ao artigo e os indicadores de citação variaram entre 0,226 e 0,402, sendo a correlação mais fraca relativa às citações do *PubMed Central*. Entre o indicador altmétrico de downloads e as citações, as correlações se mantiveram entre 0,268 e 0,444, mantendo uma similaridade na força quando comparadas com as correlações entre acessos e citações para as três bases bibliométricas analisadas.

Os resultados de Liu *et al.* (2013) reforçam os encontrados por Moed (2005) e chamam atenção para a importância das editoras explorarem o valor de interações sociais online após a publicação de pesquisas para rastrear a disseminação dos artigos publicados, usando “dados altmétricos atualizados para determinar a qualidade do artigo ou o contexto de impacto para decisões de estabilidade e promoção” (LIU *et al.*, 2013, p. 15, tradução nossa).

Thelwall *et al.* (2013), levando em consideração as correlações moderadas entre indicadores altmétricos e indicadores bibliométricos de citação encontradas nos estudos anteriores, buscam contribuir para preencher a necessidade de pesquisas mais abrangentes ao analisarem 11 indicadores altmétricos para 208.739 publicações das áreas de conhecimento de Medicina e Ciências Biológicas indexadas na PubMed e correlacioná-los com as citações. Neste estudo, o *Mendeley* não foi incluído como fonte de dados altmétricos, e a única fonte de indicador altmétrico que apresentou uma correlação relevante com citações foi o *Twitter*, como uma correlação de Spearman negativa de -0,190.

Os autores Thelwall *et al.* (2013) ressaltam que esse resultado é justificado pelo fato do *Twitter* ser uma rede social cujo uso vem crescendo rapidamente, e artigos mais recentes acabam sendo muito tweetados imediatamente após serem disponibilizados online, porém são pouco citados devido à demora no processo até que uma citação ocorra. Esse comportamento

oposto entre a velocidade e imediatez de menções de uma publicação online e a demora que uma citação leva para ocorrer implica que editoras e bibliotecas devem levar o tempo desde a publicação em consideração ao usar alométricas para classificar resultados de pesquisa (THELWALL, *et al.*, 2013). Os autores reforçam que a correlação negativa

Não deve ser interpretada como evidência de que maiores números de tweets não se associam com artigos de alta qualidade. Muito pelo contrário, a evidência (...) é de que os tweets são úteis para indicar artigos altamente citados; a correlação negativa (...) é devido aos tweets para artigos não citados que, se a tendência continuar, serão mais citados com o passar do tempo (THELWALL *et al.*, 2013, p. 6, tradução nossa).

Podemos observar que muitos dos estudos obtiveram correlações de força moderada entre indicadores alométricos e bibliométricos, o que dificulta a interpretação dos resultados, já que as métricas parecem medir coisas similares, mas não idênticas (BORNMANN, 2014). Porém, ainda que não haja um consenso ou modelo oficial de como usar as alométricas na avaliação da pesquisa, devemos levar em consideração que, com o crescente entendimento do significado dos indicadores alométricos, é necessário utilizá-los para que os editores possam escolher quais destes indicadores aplicar nos seus sistemas.

As instituições, bibliotecas e agências de fomento responsáveis pelo apoio e publicação de periódicos também devem tomar conhecimento das informações alométricas para tomar decisões relativas a financiamento, ao firmar compromissos ou decidir promoções e benefícios (SUD; THELWALL, 2014). Desse modo, em se tratando de portais de periódicos institucionais, cabe às bibliotecas coletar e exibir essas métricas de publicações dos repositórios e periódicos de sua competência e, assim, agregar valor aos seus serviços e aumentar a percepção e relevância das próprias bibliotecas (KONKIEL; SCHERER, 2013).

É possível perceber que, nos primeiros anos de estudos exploratórios e empíricos sobre as correlações entre indicadores alométricos e bibliométricos, os pesquisadores analisaram publicações de amostragem com características específicas, como a partir de revistas de qualidade reconhecida, como *Nature* e *Science* (LI; THELWALL; GIUSTINI, 2012), ou de outras revistas das áreas de Medicina e Ciências, como as indexadas na base *PubMed* (LIU *et al.*, 2013; THELWALL *et al.*, 2013).

A cada estudo finalizado, é feita a recomendação de que novos estudos com um número maior de publicações e diversificação das fontes de dados sejam desenvolvidos. Nesse sentido e, sendo um dos primeiros a analisar uma amostra considerável de publicações de diferentes áreas do conhecimento, tipos e anos de publicação, Zahedi, Costas e Wouters (2014) analisaram coberturas e correlações dos indicadores alométricos e de citação para

20.000 publicações diversas com número de *DOI* indexadas na *Web of Science*, coletando os indicadores altmétricos da plataforma *Impact Story*. O principal resultado, tanto em cobertura altmétrica quanto na força da correlação com a citação, foi para o indicador altmétrico de leitores do *Mendeley*, com uma correlação de *Spearman* de 0,49, com pelo menos um leitor no *Mendeley* para 62,6% das publicações. Para efeito comparativo, a segunda maior fonte de dados altmétricos para a amostra foi o *Twitter*, com apenas 1,6% das publicações tendo alguma menção na plataforma, resultados que demonstram que o *Mendeley* é a maior e mais útil fonte de dados altmétricos, confirmando resultados anteriores (BAR-ILAN *et al.*, 2012; PRIEM; PIWOWAR; HEMMINGER, 2011).

Mohammadi e Thelwall (2014) analisaram publicações indexadas na *Web of Science* para dez disciplinas dentro das grandes áreas de Humanidades e Ciências Sociais, e obtiveram correlações de *Spearman* entre os leitores do *Mendeley* e as citações de 0,428 e 0,516, respectivamente. Na análise por disciplina, as correlações também se mantiveram moderadas, concordando com os resultados de estudos anteriores, não sendo elas fortes o suficiente para que se possa afirmar que os indicadores de leitores do *Mendeley* e de número de citações medem o mesmo aspecto do impacto acadêmico, mas sendo suficientes para afirmar que os dados altmétricos provindos do *Mendeley* são úteis como medida suplementar as citações.

Outro estudo que confirma a importância do *Mendeley* foi o conduzido por Haustein e Larivière (2014), que selecionaram uma amostra similar a de estudos como Li, Thelwall e Giustini (2012), Thelwall, *et al.* (2013), e Liu *et al.* (2013), com publicações indexadas na *Web of Science* e *PubMed*, porém em maior escala, totalizando mais de 1,3 milhões de documentos publicados entre 2010 e 2012. A correlação geral entre o indicador altmétrico de leitores do *Mendeley* com o indicador de número de citações da *Web of Science* foi de 0,456, se mantendo próxima aos resultados anteriores. Numa análise qualitativa do tipo de leitores no *Mendeley* para as publicações, os autores concluem que enquanto as citações refletem o impacto das publicações na comunidade acadêmica, o indicador provindo do *Mendeley* espelha o uso dessas publicações por uma audiência mais ampla, mas ainda assim acadêmica, que consiste em uma maior extensão de público composto por estudantes e pós-graduandos.

Dessa forma, Haustein e Larivière (2014) confirmam o que Bornmann (2014) afirmou: as altmétricas e as citações parecem medir coisas similares, mas não idênticas. Outros estudos como os de Thelwall e Sud (2016) e Yang, Xing e Wolfram (2018) identificaram o *Mendeley* como a melhor fonte de indicadores altmétricos, tanto em cobertura quanto pela força das correlações deste indicador com as citações. Para Mohammadi *et al.* (2015), as estatísticas do *Mendeley* podem também revelar um impacto oculto de algumas publicações científicas,

como o valor educacional para usuários que não são autores de pesquisas, mas fazem parte da academia, ou o impacto das publicações na prática de leitores de fora da comunidade acadêmica.

Estando clara a distinção entre os aspectos captados pelos indicadores alométricos de downloads, acessos, leitores e menções na web social e o indicador bibliométrico de citação, é necessário ressaltar a capacidade dos indicadores alométricos em captar informações que os indicadores bibliométricos não conseguem enxergar, principalmente impactos mais próximos à data de publicação do artigo (KEELE, 2016). Nesse sentido, Costas, Zahedi e Wouters (2014) analisaram indicadores alométricos para 718.315 publicações com *DOI*, extraídas da plataforma Altmetric em 2013, e correlacionaram-nos com o indicador bibliométrico de citação da *Web of Science*. Ao observarem que esses dados alométricos são mais frequentes para publicações dos anos mais recentes (de 2011 a 2013), os autores destacam que “os dados alométricos são válidos apenas para os anos de publicação mais recentes e não têm nenhum valor real quando aplicados a publicações mais antigas, uma vez que sua presença é insignificante” (COSTAS; ZAHEDI; WOUTERS, 2015, p. 3, tradução nossa).

Considerando as características de imediatez e velocidade (ROEMER; BORCHARDT, 2015; SUN et al., 2018; THELWALL *et al.*, 2013) que os indicadores alométricos têm em refletir o impacto das publicações científicas logo após a estarem disponíveis online, Costas, Zahedi e Wouters (2015) focaram suas análises nas publicações publicadas após julho de 2011. A correlação de Spearman entre o indicador alométrico de número de *tweets* e o indicador bibliométrico de citação obtido foi de 0,167. Ao investigarem a média de citações de acordo com o número de alométricas por publicação, os autores identificam que publicações com mais alométricas tendem também a serem mais citadas.

Nessa combinação de análises é possível perceber que a correlação por si só não foi suficiente para transparecer a tendência da relação entre alométricas e citações, reafirmando a necessidade de se combinar metodologias de análises diferentes. Além disso, os resultados de Costas, Zahedi e Wouters (2014) apontam as alométricas como potencial ferramenta de filtragem de publicações relevantes.

Um dos estudos que obtiveram resultados para as correlações diferentes da maioria foi o de Ouchi *et al.* (2019). Eles analisaram uma amostra dos mil artigos da revista *Nature* mais citados no *Scopus*, fazendo deste um estudo dirigido por número de citação. A correlação de Spearman obtida entre o indicador alométrico de leitores do *Mendeley* foi forte, de 0,724, acima dos resultados de estudos anteriores. A cobertura alométrica obtida foi de 98,9%,

excepcionalmente alta quando comparada com outras pesquisas, podendo ser justificada pela reputação da revista *Nature*.

Todos esses estudos mediram as correlações entre indicadores altmétricos e indicadores bibliométricos de citação, porém os resultados normalmente apontam para correlações moderadas entre as métricas, o que dificulta um consenso interpretativo. Assim, seguem sendo necessários mais estudos avaliando diferentes cenários e com as mais variadas bases de dados, em busca de uma resposta satisfatória para a questão de se as altmétricas são apropriadas para medir o impacto das publicações científicas.

Portanto, este estudo propõe contribuir para esta investigação, analisando as correlações entre os indicadores altmétricos e bibliométricos de citação para as publicações de portais de periódicos de acesso aberto de universidades brasileiras, de modo que as altmétricas, com seu maior alcance e velocidade de medição do impacto acadêmico e social (PRIEM; PIWOWAR; HEMMINGER, 2011), possam ser consideradas para complementar métricas tradicionais na avaliação da pesquisa provendo novos insights neste processo.

2.8 ESTUDOS SOBRE INDICADORES ALTMÉTRICOS E BIBLIOMÉTRICOS EM DIFERENTES ÁREAS DE CONHECIMENTO

Meadows (1999) afirma que as áreas de conhecimento têm preferências diferentes quanto aos canais escolhidos para comunicar seus resultados e menciona a quantidade de artigos publicados em periódicos científicos como uma medida normalmente usada para avaliar a quantidade de informações que um pesquisador comunica. Porém, o padrão de comportamento das diferentes áreas de conhecimento se reflete nas diferentes práticas e preferências nas formas de comunicar a ciência.

No estudo conduzido por Mueller (2005), que analisou oito anos de produção científica de 226 pesquisadores bolsistas de pós-doutorado da CAPES, os canais preferenciais de comunicação científica para oito grandes áreas de conhecimento foram identificados: pesquisadores das Ciências Exatas e da Terra, das Biológicas e da Saúde preferem publicar suas pesquisas em periódicos estrangeiros; as Ciências Agrárias preferem ou costumam publicar em periódicos nacionais.

Já para as Engenharias, os anais de congresso são canais prestigiosos e frequentemente usados para a comunicação científica da área; e a publicação de livros e capítulos de livros é significativa para as áreas de Ciências Humanas, Ciências Sociais Aplicadas, e Linguística,

Letras e Artes. Pesquisadores das áreas de Ciências Exatas e da Terra, Biológicas e Engenharias não costumam recorrer a publicação em livros (MUELLER, 2005).

Tendo em vista que existem diferentes preferências entre as grandes áreas do conhecimento ao escolher o canal de comunicação científica, é interessante analisar os indicadores altmétricos e bibliométricos de citação de publicações em periódicos científicos destas diferentes áreas de maneira separada, levando em consideração as práticas da comunidade científica. Os próprios periódicos científicos, ao estabelecerem sua estrutura, normas e temáticas, levam em consideração essas práticas.

Assim como as revistas podem atuar como agentes no processo científico, determinando restrições e determinando forma e conteúdo do que se publica, também podem buscar a especialização de seus conteúdos e focar em um tema determinado, de audiência mais restrita. Essa forma de organização por especialidade explica a existência de diferenças no comportamento de comunicação científica das diferentes áreas, como critérios de publicações diferentes. Para avaliar adequadamente as áreas, os indicadores devem ser normalizados de modo que eliminem os efeitos dessas diferenças por conta das especialidades de cada área.

Tudo isso nos mostra que as revistas científicas têm o poder de estruturar a comunicação científica, delimitando documentos, proporcionando critérios e definindo marcos de referência adequados para estes e permitindo estimar a qualidade dos mesmos. De acordo com Barba (2003):

Na prática, parece que os cientistas, unidos em um princípio de compartilhar objetivos e problemas de investigação, se relacionam para trocar informações e experiências sobre temas determinados. Destas relações podem derivar grupos estruturados, ou comunidades disciplinares, cujos membros o são de fato, pois para pertencer a uma destas comunidades é necessário trabalhar no campo de interesse desta comunidade e respeitar certas normas básicas de relação científica (BARBA, 2003, p. 78).

Apesar das diferenças entre práticas de comunicação científica nas áreas de conhecimento, dentre as diversas formas de literatura científica, como anais de congressos, livros e artigos em periódicos, já evidenciamos anteriormente a publicação em periódicos científicos como a mais importante para a maioria das áreas de conhecimento (BARBA, 2003, VILÁN FILHO, 2010) e utilizada por todas elas, independente das preferências por esta ou aquela forma de publicação. Desse modo, a análise dos indicadores altmétricos e de citação para publicações em periódicos científicos, permanece relevante para compreender coberturas

e correlações entre estes, especialmente dentro do escopo desta pesquisa, para publicações em portais de periódicos brasileiros de acesso aberto.

Com o objetivo de analisar a cobertura das fontes de dados altmétricos e bibliométricos de artigos dos portais nas diferentes áreas do conhecimento, é preciso realizarmos análises comparativas entre eles para então identificarmos se há influência ou relação entre os indicadores altmétricos e os bibliométricos de citação nas diferentes áreas do conhecimento. Nesse sentido, Eysenbach (2011), ao analisar artigos da *Journal of Medical Internet Research (JMIR)*, revista da área de Medicina, dentro da grande área da Saúde, encontrou correlações de Pearson entre o indicador altmétrico de número de *tweets* e os indicadores de citação da *Scopus* e do Google Scholar variando entre 0,42 e 0,72.

Para Eysenbach (2011), a revista *JMIR* é ideal para análise devido ao seu fator de impacto alto, o que gera muitos eventos de citação, e devido seu tema ser justamente sobre internet e mídias sociais, com um público leitor mais engajado com ferramentas da web 2.0, o que pode impactar na cobertura dos indicadores altmétricos. Ainda que este seja um estudo inicial destas correlações entre indicadores altmétricos e bibliométricos, Eysenbach (2011) deixa a possibilidade de que seus resultados sejam um pivô para criação de uma nova métrica de avaliação baseada em impacto em mídias sociais, mesmo que ainda somente para a área de Medicina.

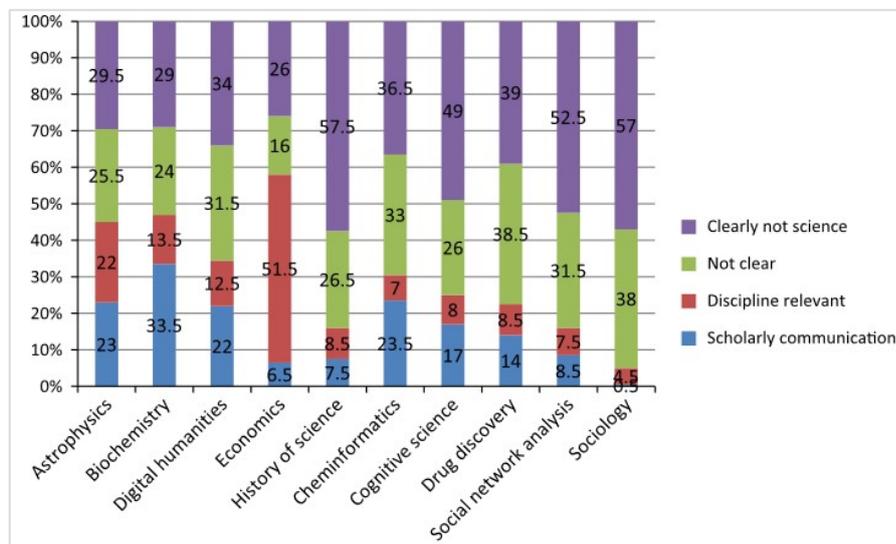
Como citado anteriormente, Bar-Ilan *et al.* (2012) analisaram publicações das áreas de Ciência da Informação e Biblioteconomia, dentro da grande área de conhecimento de Ciências Sociais Aplicadas, obtendo uma correlação entre o indicador altmétrico de leitores do *Mendeley* e o indicador de número de citações de 0,448. Já Mohammadi e Thelwall (2013), em sua análise das correlações entre indicadores altmétricos de leitores no *Mendeley* e indicadores bibliométricos de citação na *Web of Science* para disciplinas da Humanidades e Ciências Sociais, a correlação foi sempre positiva e moderada, no entanto, a correlação foi mais forte em disciplinas mais próximas das chamadas ciências duras, relacionadas a questões quantificáveis, como Negócios e Economia, com correlação de 0,573, e mais fracas em disciplinas orientadas para as Humanidades, como Religião, com correlação de 0,363 (MOHAMMADI; THELWALL, 2013).

Evidenciando a importância das fontes de dados selecionadas como fontes de indicadores altmétricos também nos estudos por área de conhecimento, Holmberg e Thelwall (2013) analisaram, em um estudo de cobertura o uso do *Twitter* como fonte de indicadores altmétricos, *tweets* feitos por 447 pesquisadores de dez disciplinas de áreas de conhecimento distintas. Os autores descobriram que apenas 2,2% de todos os *tweets* feitos possuíam link

para os artigos científicos mencionados, chegando à conclusão de que a plataforma *Twitter* é usada por muitos pesquisadores como ferramenta de comunicação científica, mas que não é frequentemente utilizada no compartilhamento de publicações científicas (HOLMBERG; THELWALL, 2013).

A principal diferença identificada entre as disciplinas analisadas por Holmberg e Thelwall (2013) foi que pesquisadores da área de História da Ciência eram menos propensos a usar o *Twitter* para fins acadêmicos, em comparação com pesquisadores de outras disciplinas. Pesquisadores da disciplina de Humanidades Digitais usam o *Twitter* mais para facilitar a conversação com outros pesquisadores, enquanto pesquisadores da disciplina de Economia são mais propensos a compartilhar links em seus *tweets* (HOLMBERG; THELWALL, 2013). O percentual de *tweets* identificados como parte da comunicação científica em cada disciplina analisada é exibido na Figura 8.

Figura 8 - Relevância dos *tweets* de comunicação científica por disciplina



Fonte: Holmberg e Thelwall (2013), p. 578.

A disciplina com maior percentual de *tweets* de comunicação científica é a Bioquímica (33,5%), e a com menor percentual é a História da Ciência (7,5%), seguida pela Análise de Redes Sociais (8,5%). Essas observações evidenciam que, ainda que a cobertura geral do *Twitter* para a amostra analisada seja baixa, em conformidade com a cobertura encontrada por estudos anteriores (COSTAS; ZAHEDI; WOUTERS, 2015; THELWALL *et al.*, 2013), existem diferenças nas práticas entre áreas, com as coberturas seguindo inclusive um comportamento similar ao encontrado por Mohammadi e Thelwall (2013) nas correlações entre indicadores altmétricos e de citação, com correlações mais fortes em disciplinas de ciências duras (como a Bioquímica) e mais fracas nas Humanidades, área que abrange as

disciplinas de História da Ciência e de Análise de Redes Sociais na amostra de Holmberg e Thelwall (2013).

Hammarfelt (2014) coletou dados de todas as publicações para o ano de 2012 da base *SwePub* (c2018) da área de Humanidades, separando-as por tipo de publicação e por disciplina. Capítulos de livro corresponderam a 31,5% das publicações, e 22,7% eram artigos científicos publicados de periódicos, corroborando as práticas de comunicação científica por área de conhecimento mencionadas por Mueller (2005). Quanto à cobertura altmétrica dos artigos científicos, 61,3% estavam presentes no *Mendeley*, 20,6% no *Twitter* e menos de 5% em outras plataformas de mídias sociais, como *Facebook* e *blogs*. Em termos de cobertura, o *Mendeley* se mostra a fonte de dados altmétricos mais promissora para analisar artigos de periódicos científicos. O autor afirma que, no geral, “parece que abordagens altmétricas funcionam melhor em campos onde os periódicos são o canal de publicação preferido, e são menos eficazes em áreas de pesquisa onde os acadêmicos publicam em uma variedade de canais” (HAMMARFELT, 2014, p. 1428).

Haustein e Larivière (2014), ao explorarem o número de leitores no *Mendeley* para 1,2 milhões de publicações de periódicos das áreas de Pesquisa Biomédica, Medicina Clínica, e Saúde e Psicologia, encontraram uma correlação geral de *Spearman* de 0,512 entre o indicador altmétrico extraído do *Mendeley* e o indicador de citações da *Web of Science*. A área com a correlação mais forte foi a Pesquisa Biomédica, com correlação de 0,575, enquanto a área com correlação mais forte foi a de Saúde, com 0,434.

De modo geral, considerando o tamanho da amostra e que todas as áreas analisadas são pertencentes à grande área de Ciências da Saúde, o estudo de Haustein e Larivière (2014) aponta para correlações moderadas nesta área. Em termos de cobertura, o *Mendeley* cobre 65,9% das publicações estudadas, similar à cobertura encontrada por Hammarfelt (2014) para publicações da área de Humanidades.

Haustein, Costas e Larivière (2015) analisaram a cobertura de citações e menções em diversas plataformas de mídias sociais, como *Twitter*, *Facebook* e *Google+*, para mais de 1,3 milhões de publicações científicas de cinco disciplinas de diversas áreas. A plataforma de maior cobertura altmétrica foi sempre o *Twitter*, com a maior cobertura para a disciplina de Ciências da Saúde e Biomédicas (31,7%), e a menor para a disciplina de Matemática e Ciências da Computação (7,5%). Para os autores, esses resultados podem evidenciar que

[...] disciplinas que têm laços mais fortes com a sociedade [...] ou que lidam com preocupações específicas da vida cotidiana das pessoas (problemas de saúde ou ambientais; ou seja, Ciências Biomédicas e da Saúde e Ciências da Vida e da Terra)

têm uma maior probabilidade de aparecer em plataformas de mídia social do que publicações de disciplinas mais técnicas e aplicadas (por exemplo, Ciências Naturais e Engenharia) ou com um nível técnico superior e componentes de complexidade (por exemplo, Matemática e Ciências da Computação) (HAUSTEIN; COSTAS; LARIVIÈRE, 2015, p. 6, tradução nossa).

Ao verificarmos a correlação entre os indicadores da fonte de dados altmétricos de maior cobertura, o *Twitter*, e as citações, a correlação de Spearman encontrada é de 0,194 para o universo de publicações, próxima a de -0,190 encontrada anteriormente por Thelwall *et al.* (2013).

Mohammadi *et al.* (2015) coletaram indicadores altmétricos de leitores no *Mendeley* e indicadores de citação na *Web of Science* para 480.979 artigos de cinco disciplinas de áreas diversas, sendo: Medicina Clínica, Engenharia e Tecnologia, Ciências Sociais, Física, e Química. As correlações encontradas entre os indicadores foram próximas entre as disciplinas, mas as correlações mais fortes foram para a Medicina Clínica e Ciências Sociais, de 0,561 para ambas.

Em uma análise do perfil dos leitores do *Mendeley*, Mohammadi *et al.* (2015) identificaram que uma porcentagem considerável das publicações de Medicina Clínica eram lidas por pessoas que não são acadêmicas. No caso de artigos de Ciências Sociais, foi identificado que uma fração notável das publicações foram lidas por bibliotecários, o que indica que estes artigos podem ser das sub-áreas de Biblioteconomia e Ciência da Informação. Esse tipo de análise é interessante para identificar as influências das práticas de leitura das diferentes áreas de conhecimento e seu reflexo na cobertura dos indicadores altmétricos, reforçando a importância da análise dos mesmos por área.

Huang, Wang e Wu (2018) consideraram 2.406 artigos para seis revistas do portal de periódicos PLOS em sua análise comparativa das correlações entre o indicador altmétrico *Altmetric Attention Score (AAS)*, calculado pela plataforma Altmetric, e o indicador bibliométrico de citação da base *Web of Science*, com a coleta dos indicadores feita em 2015. As correlações de Spearman obtidas variaram entre 0,156, para a revista Doenças Tropicais Negligenciadas, e 0,341, para a revista Biologia, com as demais revistas com correlações neste intervalo, encontrando variação de força das correlações entre as áreas de cada revista. Wang *et al.* (2014) fizeram uma análise de amostra similar, considerando 63.906 publicações revistas da PLOS, porém considerando a revista multidisciplinar *PLOS One*, além das demais revistas da área da Saúde, e dividindo também os artigos para serem analisados por ano de publicação (2004 a 2012). Os indicadores altmétricos considerados incluíram leitores do *Mendeley*, acessos à página *HTML* e ao *PDF* do artigo, e o *AAS*.

Na análise por revistas, Wang *et al.* (2014) dividiram as amostras em dois períodos de publicação: de 2004 a 2008, e de 2009 a 2012, devido a plataforma Altmetric prover dados apenas de menções a partir de 2011. As correlações mais fortes encontradas são sempre entre o acesso ao *PDF* do artigo e as citações, para todas as revistas, sendo a maior para a revista multidisciplinar *PLOS One*, com correlação de 0,715, seguida pela revista da área de Genética, com 0,714.

Em contrapartida, as correlações mais fracas são sempre entre o indicador altmétrico AAS e as citações, com pouca variação e se mantendo entre -0,122 e -0,013, sendo estes resultados similares ao encontrado por Thelwall *et al.* (2013). Para Wang *et al.* (2014), não há muitas possibilidades de que a atenção social recebida por uma publicação leve à citação. Pesquisas populares ou interessantes para a comunidade fora da academia tendem a ter uma atenção altmétrica alta, o que de acordo com os autores, pode justificar as correlações muito fracas encontradas.

Em busca de identificar a relação entre citações e indicadores altmétricos para revistas das áreas de Biblioteconomia e Ciência da Informação, Ezema e Ugwu (2019) selecionaram publicações de 85 periódicos destas áreas e coletaram o número de citações das bases *Scopus*, *Google Scholar* e *Web of Science*, e os dados altmétricos da plataforma *Altmetric*. Considerando como único indicador altmétrico o *Altmetric Attention Score (AAS)*, calculado pela plataforma *Altmetric*, os autores encontraram correlações entre moderadas e fortes, com destaque para a correlação de 0,733 entre o *AAS* e as citações do *Scopus*, e de 0,541 entre o *AAS* e as citações totais das três bases bibliométricas. As correlações notavelmente diferem de resultados anteriores, como os de Haustein, Costas e Larivière (2015), Huang, Wang e Wu (2018), Thelwall *et al.* (2013), e Wang *et al.* (2014), porém se assemelha aos resultados de Mohammadi *et al.* (2015), ainda que este tenha considerado apenas o *Mendeley* como fonte de dados altmétricos. Considerando que o *AAS* é calculado com base na atribuição de pesos para as ocorrências de menção a depender da fonte altmétrica, este pode ter influenciado para que a correlação obtida por Ezema e Ugwu (2019) seja discrepante de estudos anteriores.

A cautela no uso do AAS como indicador altmétrico em estudos de correlações se estende a outras áreas de conhecimento, e não apenas para a Biblioteconomia e Ciência da Informação. Uma das razões é a forma de contabilização dos leitores do *Mendeley* pela plataforma *Altmetric*, pois apesar dela coletar as informações de leitores do *Mendeley*, só são contabilizados leitores de publicações mencionadas em pelo menos uma das demais plataformas de mídias sociais cobertas pelo *Altmetric* (HAUSTEIN; COSTAS; LARIVIÈRE,

2015). Em outras palavras, não serão contados leitores de artigos que o Altmetric não encontrou menções em outras mídias sociais.

Com base nos diversos estudos aqui apresentados demonstrando que o *Mendeley* é a fonte de dados altmétricos mais promissora, apresentando as maiores coberturas e correlações dentre os indicadores altmétricos, entende-se que o ideal seja recorrer diretamente aos dados do *Mendeley* para realizar análises, garantindo que todos os leitores existentes para as publicações sejam contabilizados.

Outro estudo que utilizou o *Mendeley* e o *Twitter*, que até então se mostram as melhores fontes de dados de indicadores altmétricos tanto em cobertura quanto em correlação com citações, Cho (2017) analisou o impacto de 383 artigos coreanos publicados em revistas internacionais nas áreas de Ciências Médicas, Engenharia, Ciências Sociais, e Artes e Humanidades, e constatou que os artigos mais discutidos em redes sociais como o *Twitter* são da área de Ciências Médicas. Já os artigos mais lidos no *Mendeley* são, em sua maioria, das Ciências Sociais e Artes e Humanidades.

Como resultado da análise de correlação entre os indicadores altmétricos e de citação, Cho (2017) descobriu uma relação positiva de entre o número de leitores do *Mendeley* e as citações para todas as áreas de conhecimento, com a correlação mais forte ($r = 0,705$) para a área de Engenharia, seguida pelas Ciências Médicas ($r = 0,504$). Cho (2017) também observou que a correlação entre os *tweets* e os leitores do *Mendeley* é moderada nas Ciências Médicas, de 0,525, maior do que em outras áreas. Isso mostra que os resultados de estudos médicos que são amplamente utilizados por profissionais da área também são discutidos frequentemente em redes sociais, mostrando um alto impacto social (CHO, 2017).

Chen *et al.* (2020) analisaram as correlações entre indicadores altmétricos do Altmetric, downloads e o indicador bibliométrico de citação para artigos da área de Reumatologia coletados da base da editora Oxford entre 2010 e 2015. Os autores identificaram que, quando o número de downloads aumentava, também aumentavam as citações para as publicações e vice-versa. De modo contrário, encontraram correlação fraca entre os downloads e o indicador do Altmetric ($r = 0,028$) e entre o indicador do Altmetric e as citações ($r = 0,004$).

Outro estudo que obteve resultados voltados para publicações da área da Saúde foi o de Bornmann (2014), que analisou publicações da base F-1000, de literatura científica da área de Biomedicina, coletando dados altmétricos de mídias sociais como o *Twitter*, *Facebook*, *blogs*, *Youtube* e *LinkedIn* através da plataforma Altmetric, e números de citação da *Web of Science*. Diferente de estudos anteriores que obtiveram baixas coberturas (COSTAS;

ZAHEDI; WOUTERS, 2014; THELWALL *et al.*, 2013), Bornmann (2014) obteve uma cobertura altmétrica de 69% para a amostra, com a maior cobertura sendo no *Twitter*. A diferença na metodologia de Bornmann que pode ter elevado a cobertura altmétrica está na seleção de artigos publicados após 2011, uma vez que o Altmetric só provê dados confiáveis para publicações após este período (BORNMANN, 2014).

Como podemos observar ao longo da exposição da literatura, a maioria dos trabalhos sobre correlações entre indicadores altmétricos e indicadores bibliométricos de citação obtiveram correlações moderadas, dificultando a interpretação dos resultados (BORNMANN, 2014) e demandando mais estudos com diferentes amostras e em larga escala, tanto para compreender as diferenças nas relações de indicadores entre as áreas do conhecimento quanto de forma geral na literatura científica publicada em periódicos.

2.9 ESTUDOS SOBRE A INFLUÊNCIA DO FATOR TEMPORAL ENTRE INDICADORES ALTMÉTRICOS E BIBLIOMÉTRICOS

Com base nos estudos vistos até aqui, podemos perceber que a seleção de pesquisas mais recentes como amostra para as análises de correlações e coberturas dos indicadores altmétricos e bibliométricos pode influenciar nos resultados obtidos. Uma questão de pesquisa comumente enfrentada por estudos anteriores, conforme mencionado por Costas, Zahedi e Wouters (2014), é que a plataforma *Altmetric* possui informações de menções em redes sociais coletadas a partir de julho de 2011.

Ezema e Ugwu (2019) utilizaram esta limitação para restringir o período de busca nas redes sociais em sua pesquisa, e no próprio site do *Altmetric* é declarada a possibilidade de que menções feitas a artigos antes de 2011 podem não terem sido coletadas pela ferramenta. Considerando esta restrição da plataforma *Altmetric*, estudos que utilizem dados altmétricos provenientes desta plataforma devem considerar analisar, em um primeiro momento, as menções a artigos publicados a partir de 2012 e, posteriormente, as menções à totalidade de artigos publicados independentemente da data de publicação.

Diversos estudos que analisaram publicações menos recentes (publicadas antes de 2011) optaram por dividir a amostra em fatias temporais para observar a influência do tempo na força das correlações (COSTAS; ZAHEDI; WOUTERS, 2014; WANG *et al.*, 2014; ZAHEDI; HAUSTEIN, 2018). É importante ressaltar que um artigo, independentemente de ter sido publicado há mais de uma década ou em 2021, pode ter seus indicadores altmétricos coletados pela plataforma *Altmetric*, porém o que o *Altmetric* não faz adequadamente é coletar

as menções em mídias sociais feitas antes de 2011, independente da data de publicação do artigo.

Com a análise de publicações divididas por fatias temporais é possível comparar, em algum nível, se a antiguidade do artigo influencia nos indicadores altmétricos e nas respectivas correlações com outros indicadores. Um artigo científico publicado há mais tempo (por exemplo, antes do advento da internet) pode ser mencionado a qualquer momento nas mídias e redes sociais, sendo possível analisar se este continua recebendo atenção online. Mesmo que um artigo tenha sido publicado em um momento em que já existisse as mídias e redes sociais, os indicadores altmétricos, conforme limitações apontadas, sofrem influência das metodologias de coleta de dados. Assim, é importante reconhecer que existe certo viés na análise da produção em períodos anteriores à 2012 (no que se refere a indicadores oriundos da plataforma *Altmetric*), mas, também, das dinâmicas e datas de coleta de dados de download, acessos e de qualquer outra fonte de dados, inclusive de citação.

Eysenbach (2011) identificou em sua pesquisa que a correlação entre *tweets* e citações para as publicações divididas por edição publicada foi maior para a edição nº 4 de 2009 do que para a edição mais recente analisada, nº 2 de 2010. Já Shema, Bar-Ilan e Thelwall (2014) coletaram dados de artigos de dez periódicos de diferentes áreas de conhecimento, analisando as publicações de 2009 e 2010 separadamente, calculando as medianas e correlacionando as menções no blog ResearchBlogging.org (POST, c2015) com as citações. Como resultado, os autores obtiveram evidências estatísticas de que as citações atraídas por artigos mencionados no blog tendem a estar acima da mediana da revista na qual estão publicados.

De modo geral, a nível de revista, 58% das revistas na lista de 2009 e 68% de 2010 tiveram resultados estatisticamente significativos, ou seja, indicaram que menções no blog atraíram mais citações para os artigos destas revistas. O maior número de resultados significativos foi para o ano de 2010, o que pode estar relacionado ao número crescente de postagens no blog. Esses resultados mostram que alguns dos artigos mencionados no blog tendem a ser posteriormente mais citados do que outros artigos da mesma revista (SHEMA; BAR-ILAN; THELWALL, 2014).

Wang *et al.* (2014), em sua análise de publicações do portal de periódicos da *PLOS*, dividiu as estatísticas por periódico por ano, de 2004 a 2012, identificando que artigos mais antigos possuem maiores médias e medianas de leitores no *Mendeley*, porém somente são identificadas menções em mídias sociais a partir da plataforma *Altmetric* para publicações a partir de 2009, com a média do *AAS* se elevando com o passar dos anos.

É interessante observar, nas estatísticas de Wang *et al.* (2014), que o número de publicações por ano não é um fator que influencia nas médias e medianas dos indicadores de leitores no *Mendeley* e visualizações dos artigos. De modo geral, a tendência observada para as médias de citação, visualização e leitores no *Mendeley* é de diminuir para anos mais recentes, enquanto a média do *AAS* aumenta de 2009 até 2012, conforme mostra a Figura 9.

Figura 9 - Estatísticas descritivas de citação e altmétricas dos artigos da PLOS divididas por ano

Year	Number	Scopus citation		Article views		Mendeley readership		Altmetrics score	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
2004	189	135.09	97	16,667	12,253	88	54		
2005	419	73.79	43	13,199	9,944	72	50		
2006	899	51.36	33	10,836	8,401	56	37		
2007	2,199	40.18	26	8,097	6,074	44	29		
2008	4,403	37.03	24	7,131	5,249	39	25		
2009	6,038	27.55	19	5,873	4,436	33	22	1.95	0
2010	8,653	18.40	13	4,932	3,756	30	20	2.31	0
2011	15,871	11.59	8	3,837	2,970	22	15	2.68	0
2012	25,634	5.31	4	3,025	2,212	14	10	3.89	1
2004-2012									
2004-2008	7,709	44	27	8,402	6,127	45	29		
2009-2012	56,196	11	7	3,854	2,837	21	13	3	0

Fonte: Wang *et al.* (2014, p. 4).

Quanto à análise das correlações entre os indicadores altmétricos e de citação, que podem ser observadas na Figura 10, Wang *et al.* (2014) identificaram que a maior correlação, de 0,772, é entre as citações e downloads do *PDF* dos artigos publicados em 2004, com as demais correlações entre estes indicadores se mantendo acima de 0,517 para todos os períodos analisados, superando inclusive as correlações entre o indicador altmétrico de leitores no *Mendeley* e as citações para todos os anos. Para os autores, a atenção social “não tem influência direta sobre a citação acadêmica, no entanto, pode causar a visualização do artigo e levar à citação posteriormente, o que significa que uma parte das citações pode ser causada por atenção social, incluindo *Twitter*, *Facebook*, blogs, notícias, etc” (WANG *et al.*, 2014, p. 9, tradução nossa).

Figura 10 - Correlações entre citações e outras métricas dos artigos divididas por ano

Dateset	N	Total	HTML	PDF	Mendeley	Altmetrics	T H P M A
2004	189	0.705	0.683	0.772	0.500	NA	■ ■ ■ ■
2005	419	0.515	0.503	0.517	0.363	NA	■ ■ ■ ■
2006	899	0.492	0.441	0.629	0.469	NA	■ ■ ■ ■
2007	2199	0.581	0.549	0.672	0.525	NA	■ ■ ■ ■
2008	4003	0.676	0.648	0.740	0.550	NA	■ ■ ■ ■
2009	6038	0.664	0.637	0.719	0.547	0.221	■ ■ ■ ■ _
2010	8653	0.612	0.587	0.658	0.484	0.198	■ ■ ■ ■ _
2011	15871	0.627	0.606	0.651	0.490	0.143	■ ■ ■ ■ _
2012	25634	0.569	0.544	0.608	0.420	0.149	■ ■ ■ ■ _
2004-2008	7709	0.631	0.603	0.699	0.538	NA	■ ■ ■ ■
2009-2012	56196	0.673	0.647	0.715	0.546	-0.031	■ ■ ■ ■

Fonte: Wang *et al.* (2014, p. 6).

Costas, Zahedi e Wouters (2014) confirmam os resultados de Wang *et al.* (2014) ao observar que, para a amostra de mais de 1,3 milhões de publicações com DOI analisadas, os dados altmétricos são mais frequentes entre as publicações mais recentes, particularmente dos anos de 2011, 2012 e 2013, considerando que a coleta dos dados foi feita em outubro de 2013. De acordo com Costas, Zahedi e Wouters (2014),

[...] em 2011, cerca de 10,8% de todas as publicações com DOI [...] recebeu alguma atenção altmétrica. Essa participação sobe para 23,8% nas publicações de 2012, e acima de 25% em 2013. Considerando isso, faz sentido destacar que as altmétricas são válidas apenas para os anos de publicação mais recentes, e não têm nenhum valor real quando aplicadas a publicações mais antigas, uma vez que sua presença é insignificante (Costas; Zahedi; Wouters, 2014, p. 3, tradução nossa).

Isso indica o valor das altmétricas quando aplicadas a publicações recentes, mas também é preciso considerar que a plataforma *Altmetric* só passou a efetivamente coletar dados altmétricos a partir de julho de 2011.

Em se tratando de análise de correlações entre os indicadores, Thelwall e Sud (2016) encontraram correlações positivas entre o indicador altmétrico de leitores no *Mendeley* e as citações no *Scopus* para todos os anos de publicação examinados em sua amostra. Para as cinco áreas do conhecimento analisadas, as correlações de Spearman diminuem lentamente entre os anos de 2004 e 2009 e, em seguida, diminuem mais rapidamente até o ano de 2014.

Na análise das médias de citações, e sendo os anos de 2013 e 2014 os mais recentes da amostra deste estudo, a média para esses dois anos são próximas de zero, o que explicaria as baixas correlações com leitores do *Mendeley* para publicações destes anos, confirmando a necessidade de se procurar alternativas que substituam o uso do indicador de citação para artigos publicados recentemente (THELWALL; SUD, 2016).

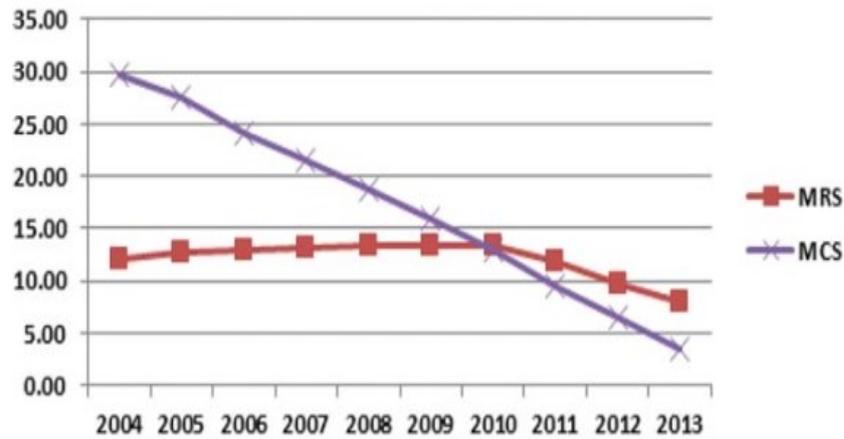
Além disso, os resultados obtidos das análises das correlações por ano também confirmam um padrão encontrado em um estudo anterior feito por Maflahi e Thelwall (2015) para quatro periódicos das áreas de Biblioteconomia e Ciência da Informação, que indica que as correlações entre os leitores no *Mendeley* e as citações para artigos tendem a aumentar ao longo dos cinco anos que se seguem após a data de publicação, e então se estabilizar. Para Thelwall e Sud (2016),

[...] a causa das baixas correlações nos anos iniciais (da publicação) parece ser o número muito baixo de citações, porque há uma tendência de haver substancialmente mais leitores do que citações nos primeiros anos após a publicação. Consequentemente, os resultados sugerem que o número de leitores no Mendeley é um bom substituto para o número de citações para análises cientométricas envolvendo artigos com menos de 5 anos, especialmente quando os artigos estão em seus primeiros anos, e em campos com níveis mais elevados de uso do Mendeley (THELWALL; SUD, 2016, p. 3048, tradução nossa).

Dessa forma, o indicador altmétrico de leitores do *Mendeley* mostra potencial na medição do impacto acadêmico de publicações recentes, que ainda não são quantificáveis através de indicadores bibliométricos. O comportamento das correlações observado por Thelwall e Sud (2016) possui diferenças mínimas entre as áreas de conhecimento, o que também sugere que o fenômeno é generalizado, sem restrição por áreas ou disciplinas.

Zahedi, Costas e Wouters (2017) também estudaram a distribuição e presença dos leitores do *Mendeley* ao longo do tempo e em diferentes disciplinas para mais de nove milhões de publicações com DOI indexadas na *Web of Science* e publicadas entre os anos de 2004 e 2013. Os resultados de Zahedi, Costas e Wouters (2017) confirmam a tendência encontrada nos resultados de Thelwall e Sud (2016), mostrando que a cobertura das publicações com ao menos um leitor no *Mendeley* aumentou de 2004 para 2009, sofrendo uma queda nos anos mais recentes, de 2010 a 2013. A distribuição nas médias de citação e de leitores no *Mendeley* obtida por Zahedi, Costas e Wouters (2017) e ilustrada na Figura 11 sugere que publicações mais recentes possuem, em média, mais leitores do que citações.

Figura 11 - Distribuição das médias de citação (MCS) e leitores no *Mendeley* (MRS) ao longo do tempo



Fonte: Zahedi, Costas e Wouters (2017, p. 2514).

Em defesa do uso do *Mendeley* como ferramenta cientométrica, estão os diversos resultados de estudos anteriores aqui previamente elucidados que obtiveram correlações positivas e moderadas entre o indicador altmétrico de leitores no *Mendeley* e o indicador bibliométrico de citação, sendo estas correlações sempre muito mais fortes do que entre qualquer outro indicador altmétrico provindo de mídias sociais e as citações (COSTAS; ZAHEDI; WOUTERS, 2014), indicando uma semelhança entre essas duas métricas.

Entretanto, estudos anteriores obtiveram correlações ainda mais fortes entre o indicador altmétrico de downloads do artigo e o indicador de citação, assim como entre o indicador de acessos ou visualizações do artigo e o indicador de citação (CHEN *et al.*, 2020; LIU *et al.*, 2013; PERNEGER, 2004; WANG *et al.*, 2014). Entre as fontes de dados que são mídias sociais, o *Twitter* se mostrou a mais relevante nos estudos altmétricos correlacionais e em cobertura (BORNMANN, 2014; CHO, 2017; COSTAS; ZAHEDI; WOUTERS, 2014; THELWALL *et al.*, 2013). Dessa forma, a seguir serão apresentadas em detalhes cada uma destas fontes de dados altmétricos, bem como a fonte de dados bibliométricos, consideradas nesta pesquisa.

3 METODOLOGIA E PROCEDIMENTOS METODOLÓGICOS

3.1 CARACTERIZAÇÃO DA PESQUISA

De acordo com Creswell (2010), um projeto de pesquisa envolve a intersecção de filosofia, estratégias de investigação e métodos específicos. A filosofia, ou concepções filosóficas do estudo, é uma orientação geral sobre a natureza da pesquisa desenvolvida (CRESWELL, 2010), influenciando nas práticas metodológicas adotadas. Portanto, faz-se necessário explicitar a concepção filosófica do presente estudo.

A presente pesquisa possui uma concepção filosófica pós-positivista, caracterizada pela observação e mensuração a realidade objetiva, buscando desenvolver declarações relevantes que expliquem as relações causais de interesse do estudo. Devido a suas características de objetividade, reducionismo e observação e mensuração empíricas (CRESWELL, 2010), o pós-positivismo está intimamente relacionado ao desenvolvimento de uma abordagem de pesquisa quantitativa, pois suas suposições são mais válidas para esta abordagem, com a investigação da relação entre variáveis e a apresentação em termos de questões. Estas variáveis são medidas para que os dados numéricos possam ser analisados por procedimentos estatísticos.

A abordagem metodológica utilizada é quantitativa, marcada por aspectos essenciais como a objetividade e a confiabilidade (CRESWELL, 2010), sendo necessário examinar os métodos utilizados para evitar vieses interpretativos. Este relatório possui a estrutura fixa da metodologia quantitativa, bem como as estratégias de pesquisa utilizadas (análise de dados numéricos) são caracterizadas pela abordagem quantitativa.

Quanto aos fins, esta pesquisa é descritiva, com o estabelecimento de correlações entre variáveis. De acordo com Silva e Menezes (2005), a pesquisa descritiva envolve o uso de técnicas padronizadas de coleta de dados e assume, de modo geral, a forma de levantamento. Desse modo, o método desta pesquisa envolveu o levantamento, classificação, tabulação e análise dos dados coletados.

Esta pesquisa, conforme já destacado, está associada à abordagem metodológica quantitativa, sendo produzidos e analisados indicadores altmétricos e bibliométricos, suas correlações e as coberturas. Cada uma das etapas e respectivos procedimentos metodológicos são descritos a seguir.

3.2 COLETA E TRATAMENTO DOS DADOS

A coleta dos dados altmétricos e bibliométricos, nas diferentes fontes de dados, foi feita de maneira automatizada entre os dias 21 de fevereiro e 02 de março de 2022. As coletas de dados altmétricos, de diferentes fontes de dados e revistas, foram realizadas em datas aproximadas (com diferença máxima de uma semana entre elas), com vistas a minimizar limitações em decorrência da possibilidade de variação nos dados altmétricos em um curto período. Para garantir a consistência da pesquisa, a seleção dos portais e a coleta e tratamento dos dados foram executadas de acordo com as etapas detalhadamente descritas a seguir.

3.2.1 Etapa 1: Seleção dos Portais de Periódicos

Para seleção dos portais de periódicos a serem estudados, foram estabelecidos os seguintes requisitos:

- Publicação em plataforma digital de acesso aberto;
- Armazenamento em servidor institucional, dentro do escopo de portais de periódicos de universidades;
- Divulgação de conteúdo para texto completo;
- Disponibilidade do conteúdo e rastreabilidade por outros bancos de dados;
- Utilizar o software *Open Journal Systems*, permitindo a geração de relatórios como indicadores de acesso e downloads para todas as publicações;
- Atribuição e registro de *DOI* para os artigos.

A seleção dos portais para compor a pesquisa foi feita a partir da análise da quantidade de periódicos de acesso diamante que o portal possuía registrado no *DOAJ*, realizado por meio da interface de busca desta base a partir dos seguintes parâmetros:

- Busca por periódico;
- Utilização da expressão de busca “universidade”;
- A pesquisa do termo (universidade) foi feita no campo “Editoras”, de maneira a recuperar apenas editoras universitárias;
- Refinamento dos resultados para periódicos que não cobrem *APCs* (garantindo que o número listado fosse de periódicos de acesso diamante).

Para os dez portais com maior número de revistas indexadas na *DOAJ*, resultantes desta busca, foram enviados e-mails para o endereço eletrônico da unidade gestora de cada um, convidando-os para participarem da pesquisa. A concordância em participar era necessária, visto que cada portal precisaria disponibilizar os relatórios de acesso e download de seus periódicos. Os portais de periódicos que responderam ao convite e aceitaram participar da pesquisa, se disponibilizando a fornecer os dados solicitados foram:

- Portal de Periódicos da Universidade Federal de Santa Catarina (UFSC)¹¹;
- Portal de Periódicos da Universidade de Brasília (UnB)¹²;
- Portal de Periódicos da Universidade Federal de Goiás (UFG)¹³;
- Portal de Periódicos da Universidade Federal do Rio Grande do Norte (UFRN)¹⁴.

3.2.2 Etapa 2: Identificação de DOIs de publicações

Para o desenvolvimento desta pesquisa, foram consideradas somente publicações com número de *DOI*. Para identificação dos *DOI*'s, foi consultada a *REST API* do *Crossref* (2020), que expõe metadados de conteúdo dos membros (como editoras) que registram conteúdo junto à agência.

Após análise da documentação da *REST API*, foi identificado que o dado mais adequado para recuperação da lista de *DOIs* para as publicações é através do número de ISSN de cada revista. Desse modo, foi feito o levantamento dos números de ISSN para cada revistas dos portais de periódicos do corpus deste estudo diretamente no site oficial do ISSN *International Center*¹⁵, e criada uma lista com estes ISSNs.

A sintaxe utilizada para busca na *REST API* foi: “<https://api.crossref.org/journals/>” + ISSN + “[/works?select=DOI&rows=1000&offset=](https://api.crossref.org/journals/)” + INTERVALO, sendo ISSN a variável que representa o número de ISSN de cada revista. Como a *REST API* restringe os resultados de busca em até 1000 por página, a variável INTERVALO representa o índice inicial para

¹¹ UNIVERSIDADE FEDERAL DE SANTA CATARINA. **Portal de periódicos**. Florianópolis: UFSC, 2021. Disponível em: <https://periodicos.bu.ufsc.br/periodicos-de-a-a-z/>. Acesso em: 10 dez. 2021.

¹² UNIVERSIDADE DE BRASÍLIA. **Portal de periódicos**. Brasília: UnB, 2021. Disponível em: <https://periodicos.unb.br/>. Acesso em: 10 dez. 2021. Acesso em: 10 dez. 2021.

¹³ UNIVERSIDADE FEDERAL DE GOIÁS. **Portal de periódicos**. Goiânia: UFG, 2021. Disponível em: <https://revistas.ufg.br/>. Acesso em: 10 dez. 2021.

¹⁴ UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE. Portal de periódicos eletrônicos. Natal: UFRN, 2021. Disponível em: <https://periodicos.ufrn.br/index/>. Acesso em: 10 dez. 2021.

¹⁵ INTERNATIONAL STANDARD SERIAL NUMBER. **Portal**. [S. l.]: ISSS, 2021. Disponível em: <https://portal.issn.org/>. Acesso em: 10 dez. 2021.

apresentar os resultados por página, ou seja, para os primeiros 1000 resultados, a variável intervalo será igual a zero, uma vez que a *REST API* do *Crossref* inicia a contagem do índice em zero. Para os resultados de 1000 a 1999, a variável INTERVALO será 1000.

Do total de 205 revistas dos portais selecionados com ISSN, constatou-se que 153 revistas possuem DOI registrados pelo menos em um de seus números publicados. O número de revistas por portal de periódico com ISSN registrado e a respectiva cobertura de revistas com DOI dos portais de periódicos selecionados é exibido na Tabela 1.

Tabela 1 - Número de revistas com ISSN e número de revistas com DOI, por portal de periódicos

Portal de Periódicos	ISSN	DOI*	% de Cobertura
UFSC	48	42	88%
UnB	84	52	62%
UFG	34	34	100%
UFRN	39	25	64%
Total	205	153	75%

Fonte: Dados de pesquisa.

* Ao menos um número da revista possuía DOI

Nota-se que a o portal de periódicos com maior cobertura de revistas com *DOI* é o Portal de Periódicos da UFG, com 100% das revistas com *DOI*. Essa cobertura se deve ao fato de que este é o único portal de periódicos que possui uma política de exibir publicamente em seu site apenas as revistas que estão ativas há pelo menos três anos (UFG, 2017), com direito a atribuição de *DOI* às suas publicações. Os demais portais de periódicos listam em seus sites todas as revistas hospedadas no sistema, incluindo revistas desativadas e incubadas. A cobertura total de revistas com *DOI* foi de 75%.

As listas de resultados foram exportadas em formato *Java Script Object Notation (JSON)* e agrupadas em quatro arquivos, um para cada portal de periódicos. Esse agrupamento deu-se por meio da ferramenta *OpenRefine*¹⁶, um aplicativo, de código aberto, de limpeza e transformação de dados para outros formatos. Os quatro arquivos foram então exportados em formato de planilha do Microsoft Excel e tratados de modo a eliminar caracteres de formatação e para que cada número de DOI ocupasse uma linha do arquivo, procedimento necessário para a próxima etapa da coleta de dados.

¹⁶ OPENREFINE. **Home**. [S. l.]: OpenRefine, 2021. Disponível em: <https://openrefine.org/>. Acesso em: 10 dez. 2021.

3.2.3 Etapa 3: Identificação das áreas de conhecimento

A partir da lista de ISSNs, foi consultada e definida a área de conhecimento principal de cada revista, seguindo dos critérios abaixo, em ordem de prioridade:

- A área de conhecimento especificada pelo portal de periódicos;
- O escopo da revista, definido na seção “sobre a revista”;
- Área com maior classificação no Qualis Periódicos 2013-2016.

Assim, foi especificada a área de conhecimento principal de cada revista com base na classificação pré-definida por cada portal de periódico e, caso não houvesse uma classificação pelo portal, a área de conhecimento foi definida de acordo com a descrição dada por cada revista na página “sobre a revista” no site do próprio periódico. Se a própria revista não deixa claro a área de conhecimento, por fim, foi consultado o *ISSN* da revista na plataforma Sucupira¹⁷, para identificação da área de conhecimento da revista que recebeu maior classificação no Qualis Periódicos. As áreas de conhecimento das revistas foram agrupadas em cinco grupos:

- Ciências da Vida (CV);
- Ciências Humanas (CH);
- Ciências Sociais Aplicadas (CSA);
- Linguística, Letras e Artes (LLA); e,
- Ciências Exatas, Tecnológicas e Multidisciplinar (CETM).

Estes cinco grupos foram criados com base nas análises das revistas, considerando-se a afinidade com os Colégios e grandes áreas de conhecimento definidas pela CAPES¹⁸, que subdivide as 49 áreas de avaliação de Programas de Pós-graduação em níveis, sendo o primeiro nível composto por três colégios, e o segundo nível composto por nove grandes áreas, conforme o Quadro 3. Para fins de conhecimento, apresenta-se também no Quadro 3 o número de revistas dos portais de periódicos analisados, por grande área.

¹⁷ BRASIL. Ministério da Educação. Centro de Aperfeiçoamento de Pessoal de Nível Superior. **Plataforma Sucupira**. Brasília: CAPES, 2022. Disponível em: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf>. Acesso em: 10 fev. 2022.

¹⁸ BRASIL. Ministério da Educação. Centro de Aperfeiçoamento de Pessoal de Nível Superior. **Sobre áreas de avaliação**. Brasília: CAPES, 2022. Disponível em: <https://www.gov.br/capes/pt-br/aceso-a-informacao/acoes-e-programas/avaliacao/sobre-a-avaliacao/areas-avaliacao/sobre-as-areas-de-avaliacao/sobre-as-areas-de-avaliacao>. Acesso em: 10 fev. 2022.

Quadro 3 - Colégios e grandes áreas de conhecimento definidos pela CAPES para avaliação de Programas de Pós-graduação, e número de revistas por grande área de conhecimento.

Colégio	Grande área	Número de revistas
Ciências da Vida	Ciências Agrárias	4
	Ciências Biológicas	4
	Ciências da Saúde	9
Humanidades	Ciências Humanas	55
	Ciências Sociais Aplicadas	33
	Linguística, Letras e Artes	35
Ciências Exatas, Tecnológica e Multidisciplinar	Ciências Exatas e da Terra	6
	Engenharias	3
	Multidisciplinar	5

Fonte: Elaborado pela autora, baseado em Plataforma Sucupira, CAPES (BRASIL, 2022).

As Ciências da Vida abrangem áreas como Zootecnia, Ciências Ambientais, Medicina Veterinária, Biologia, Educação Física, Medicina, Enfermagem e Farmácia. As Ciências Humanas abrangem revistas de áreas como Antropologia, Educação, Filosofia, História e Geografia. As Ciências Sociais Aplicadas abrangem áreas como Direito, Biblioteconomia, Economia, Administração, Política, e Arquitetura e Jornalismo. A área de Linguística, Letras e Artes abrangem as áreas de Literatura, Linguística, Letras e Artes. Por fim, as Ciências Exatas, Tecnológicas e Multidisciplinares abrangem as áreas de Matemática, Física, Engenharia Civil e Interdisciplinar.

É possível observar que, entre os grupos de conhecimento definidos para classificar as revistas participantes deste estudo, dois são equivalentes a colégios (CV e CETM), e três equivalentes às grandes áreas do colégio de Humanidades (CH, CSA, LLA). A razão de não ter sido adotado o nível de colégio de Humanidades para classificar as revistas das grandes áreas pertencentes a este se deve ao grande número de revistas destas áreas. Este número é muito superior ao de outras áreas, sendo adequada a subdivisão a nível de grandes áreas (ao invés de colégio), em busca de uma análise mais detalhada dos dados.

3.2.4 Etapa 4: Coleta dos dados altmétricos e bibliométricos

A coleta dos dados do *Altmetric*, do *Mendeley* e das citações do *Crossref* foi feita entre os dias 28 de fevereiro e 02 de março de 2022, utilizando-se a ferramenta *Odisseia Metrics* (RAMOS; MARICATO, 2020). Esta ferramenta é um *script*, elaborado com a linguagem de

programação *Python*, que consulta a *API* pública da Plataforma *Altmetric* a partir da busca pelos números de *DOI* das publicações. Na interface da *Odisseia*, conforme Figura 12, é necessário selecionar a fonte de dados de onde os indicadores serão coletados, e então deve ser feito o *upload* do arquivo contendo a lista de números de *DOI* a serem buscados na base selecionada.

Figura 12 - Interface da ferramenta Odisseia



Fonte: Ramos e Maricato (2020).

Foram realizadas 12 coletas por meio da *Odisseia Metrics*. Para cada um dos quatro arquivos contendo os DOIs de cada portal de periódicos foram realizadas buscas individuais em cada uma das fontes de dados (*Altmetric*, *Mendeley* e *Crossref*). Ao coletar os indicadores altmétricos provenientes da plataforma *Altmetric*, o indicador altmétrico de número de *tweets* e o número de menções em postagens no *Facebook* por publicação científica foram extraídos de maneira individualizada, ou seja, com buscas independentes.

Quanto à coleta dos dados bibliométricos de citação na *Web of Science*, esta foi feita através da interface de busca avançada do site, utilizando-se as listas pré-existentes de números de DOI das publicações, separados pelo operador booleano OR, exemplificado na seguinte sintaxe: “DO= NÚMERO_DOI_1 OR DO= NÚMERO_DOI_2 OR... DO= NÚMERO_DOI_3”.

A coleta dos dados bibliométricos de citação da *Scopus* também foi realizada através da interface de busca avançada do site¹⁹, utilizando-se a lista de números de DOI das publicações separados pelo operador booleano OR, exemplificada com a seguinte sintaxe: DOI (10.1007/s11192-022-04442-2) OR DOI (10.1080/00987913.2022.2066967) OR DOI (10.1590/2318-0889202133e200071)”, e assim sucessivamente.

Como a interface de busca avançada da *Web of Science* e da *Scopus* suportam buscas de no máximo 1000 itens, as listas de *DOIs* dos portais de periódicos foram divididas a cada 1000, e estes intervalos de números de *DOIs* utilizados nas buscas. Uma vez que o total de

¹⁹ SCOPUS. Elsevier. [S. l.]: Scopus, 2021. Disponível em: <https://www.scopus.com/search/form.uri?display=advanced> Aceso em: 12 fev. 2021.

números de *DOIs* coletados foi de 52.354, o total de buscas avançadas realizadas em cada uma destas bases de dados foi de 53 buscas. Os relatórios dos resultados na base *Scopus* foram exportados em arquivos no formato CSV (*Comma Separated Values*) e os relatórios da *Web of Science* exportados em formato de planilha do Microsoft Excel (XLS). Finalmente, os arquivos dos relatórios de cada uma destas bases foram unificados utilizando o *OpenRefine*.

A coleta dos relatórios de acesso ao resumo das publicações científicas e o número de downloads dos arquivos destas publicações foi feita manualmente pelos gestores dos portais de periódicos através da interface do sistema *OJS*, utilizando-se o gerador de relatórios, nativo do sistema. Dentre os tipos de relatórios que o *OJS* permite exportar, o de acessos é o que provê os indicadores necessários de acessos ao resumo e downloads das publicações, sendo possível gerar um relatório por revista, sem limite de linhas ou itens por arquivo.

Os relatórios do Portal de Periódicos da UFSC e do Portal de Periódicos da UFRN foram enviados pela equipe gestora de cada um destes portais via e-mail, e os relatórios do portal de Periódicos da UnB e do Portal de Periódicos da UFG foram gerados pela própria autora desta pesquisa, que teve o acesso ao sistema *OJS* de cada um destes portais. Todos os relatórios de acesso foram gerados dentro do período relativamente curto de tempo (entre os dias 21 e 25 de fevereiro de 2022), com vistas a minimizar vieses. Após a coleta dos dados, os relatórios de acessos e downloads foram unificados em um único relatório/arquivo, contendo informações destes indicadores, individualizados, para todas as revistas dos quatro portais de periódicos selecionados.

No que se refere a análise temporal, baseada na data de publicação do artigo, a partir dos relatórios de citação exportados das fontes de dados *Scopus* e *Web of Science*, foram identificados apenas 28 periódicos que possuíam algum artigo indexado em alguma destas fontes. Para cada um destes periódicos, foram gerados dois relatórios personalizados no *OJS*, diferentes dos relatórios anteriormente mencionados, sendo: um relatório contendo os dados de acessos ao resumo do artigo por ano, para o período de janeiro de 2012 a dezembro de 2021; e, um relatório contendo os downloads de artigos por ano para o mesmo período, correspondente a 10 anos.

Os dados bibliométricos de citação na *Web of Science* e na *Scopus* previamente coletados possuem informações de número de citações por ano recebidas por cada artigo, e estes foram utilizados na análise temporal. Devido a *Crossref* não disponibilizar detalhes da contagem de citações publicamente, sendo que apenas os membros que registram conteúdo junto à *Crossref* aptos a acessar estas informações (AGUIAR, 2022), esta fonte de dados não fez parte da análise temporal dos indicadores.

3.2.5 Etapa 5: Tratamento dos dados coletados

Após a coleta dos dados de indicadores altmétricos e bibliométricos nas fontes de dados *Altmetric*, *Mendeley*, *OJS*, *Crossref*, *Web of Science* e *Scopus*, os diversos relatórios gerados foram unificados, de modo que os metadados contidos em cada arquivo fossem compatíveis, resultando em quatro arquivos das fontes unificados, divididos em: dados de acessos e downloads (RA); dados da plataforma *Altmetric* (*Tweet* e *Facebook*), *Mendeley*, e *Crossref* coletados por intermédio da ferramenta *Odiseia Metrics* (OD); dados de citações da *Web of Science* (CW); dados de citações da *Scopus* (CS).

Os dados dos arquivos coletados por meio do OD, e os dados CW e CS foram cruzados utilizando-se o metadado de número de DOI como chave primária, por ser uma coluna de dado comum a todos os arquivos, exceto aos relatórios de acessos. Essa etapa objetivou integrar todos os dados coletados em uma única planilha do Microsoft Excel, para que os cálculos de cobertura e correlação entre os indicadores coletados fossem executados.

Uma vez que os relatórios de acesso e *download* (RA) não possuem informação do número de DOI, o metadado “Identificador (ID) da publicação” foi utilizado para cruzar os dados destes relatórios com o arquivo unificado resultante da junção de OD, CW e CS. Este metadado refere-se ao identificador atribuído pelo sistema *OJS* no momento da submissão do artigo no sistema, e é utilizado pela maioria das revistas na composição do sufixo do número de DOI. Desse modo, os números de DOI foram tratados de maneira a isolar os IDs em uma nova coluna, sendo possível realizar o cruzamento dos dados desta nova coluna com a coluna ID da publicação, contida nos relatórios de acesso e download.

Para revistas que não utilizam o ID da publicação na composição do DOI, o metadado utilizado para cruzar os arquivos foi o “Título”. Este metadado não foi selecionado como chave primária nos cruzamentos de dados anteriores por dois motivos: por não estar presente em todos os relatórios (somente nos relatórios de acesso e download); e, por se tratar de um campo composto por *strings* mais longas, um tipo de dado que contém caracteres especiais e espaços em brancos, podendo haver incompatibilidades maiores do que para o tipo de dado numérico, como os IDs das publicações.

Sendo assim, o procedimento utilizado para cruzar os dados do RA, que não foram possíveis de serem cruzados através do ID da publicação, foi:

- Identificar, manualmente, as revistas cujos *DOIs* não são compostos pelo ID da publicação;

- Coletar os metadados de título, um a um, na *REST API* do Crossref, utilizando a sintaxe de busca “<https://api.crossref.org/journals/>” + ISSN + “/works?select=DOI,title&rows=1000”, que retorna os metadados de número DOI e título da publicação;
- Reunir os resultados da busca em uma única planilha do Microsoft Excel, utilizando o OpenRefine;
- Cruzar os dados da planilha com o RA, utilizando o campo “Título”.

Finalmente, após este tratamento, todos os dados de indicadores alométricos e bibliométricos, de todas as fontes de dados consideradas, bem como os metadados de identificação das publicações como número de *DOI*, ID e título, foram reunidos em uma única planilha.

A partir do número total de 52.354 publicações com *DOI* coletadas, foram excluídas publicações de conteúdo não-científico através da identificação, no campo de metadados “Título”, das palavras-chave “Apresentação”, “Editorial”, “Expediente”, e “Normas para publicação”. Essa etapa de tratamento garante que o corpus seja composto exclusivamente por artigos científicos, de modo que outros tipos de conteúdo não interfiram nas coberturas e correlações a serem analisadas. Assim, o universo final resultou em 51.200 artigos científicos, das 153 revistas que registram *DOI* para suas publicações.

Para a análise baseada nas áreas de conhecimento, a planilha única contendo os 51.200 artigos científicos foi dividida em cinco planilhas, de acordo com a área de conhecimento de cada artigo.

Para a análise baseada na data de publicação dos artigos, os relatórios são diferentes dos utilizados na análise geral e na análise por área de conhecimento. Na etapa de coleta de dados, foram coletados 28 relatórios de acessos ao resumo e 28 relatórios de downloads de artigos por ano. Os 28 relatórios de acessos ao resumo foram mesclados em um único relatório, e o mesmo foi feito para os relatórios de downloads, restando apenas dois relatórios: o relatório de acessos ao resumo por ano (RAA) e o relatório de downloads por ano (RDA).

Os dados coletados das fontes de dados bibliométricas *Web of Science* e *Scopus* foram cruzados com os relatórios RAA e RDA, considerando o título do artigo como chave primária, por ser o único metadado em comum entre os relatórios de acessos ao resumo, downloads, e citações na *Web of Science* e *Scopus*. Foi observado que nem todos os artigos indexados nas fontes de dados bibliométricas *Web of Science* e *Scopus* possuem informações de acessos ao resumo e downloads, o que reduziu significativamente o número de artigos após o cruzamento

dos dados. Após o cruzamento dos dados, estes foram divididos em seis faixas temporais, de acordo com o ano de publicação dos artigos:

- Artigos publicados antes de 2017;
- Artigos publicados em 2017;
- Artigos publicados em 2018;
- Artigos publicados em 2019;
- Artigos publicados em 2020;
- Artigos publicados em 2021.

3.2.6 Etapa 6: Testes estatísticos e de cobertura

Foi utilizada a correlação de *Spearman* no cálculo das correlações entre os indicadores alométricos e bibliométricos, devido os dados serem tipicamente muito distantes das médias (alto desvio padrão), não sendo recomendável a aplicação de testes de Pearson, sendo este mais comumente utilizado em distribuições normais de dados. Esse tipo de dado comumente apresenta muitos valores nulos para serem transformados em distribuição normal (LI; THELWALL; GIUSTINI, 2012; SUD; THELWALL, 2014; WANG *et al.*, 2014).

Utilizando-se estatística descritiva para obter informações da distribuição, para confirmar se a correlação de Spearman era a mais adequada para o conjunto de dados estudados, foi necessário medir o desvio da normalidade através de testes estatísticos. Para isso, foram utilizados os testes de normalidade Kurtosis e de distorção (*skewness*). O teste Kurtosis mede se há um pico no gráfico de distribuição de valores, ou se estes são mais planos. Valores próximos de zero indicam uma distribuição próxima do normal (-1 e +1 são considerados excelentes para a maioria dos propósitos e análises deste tipo), enquanto valores entre -2 e +2 são aceitáveis a depender da aplicação (GEORGE; MALLERY, 2019).

Já o teste *skewness* mede até que ponto uma distribuição de valores se desvia da simetria em torno da média. Um valor zero representa um equilíbrio simétrico ou uma distribuição uniforme. Assim como no Kurtosis, um valor de assimetria entre -1 e +1 é considerado excelente, mas um valor entre -2 e +2 é aceitável dependendo da sua aplicação (GEORGE; MALLERY, 2019).

Entretanto, não há um consenso sobre qual a faixa de valores para o teste de Kurtosis e *skewness* para uma amostra ser considerada normal, podendo valores entre -3 e +3 ainda serem considerados normais por alguns autores (BYRNE, 2011). Ainda assim, não há dúvidas

de que quanto mais o resultado de ambos os testes de normalidade se desviarem de zero, mais anormal é a distribuição dos valores de uma amostra, sendo recomendável a utilização da correlação de Spearman para esse tipo de amostra de dados (ao invés da correlação de Pearson).

A partir deste referencial, os testes de normalidade *Kurtosis* e *skewness* foram executados para todas as variáveis do corpus deste estudo, tanto para indicadores alométricos quanto para indicadores bibliométricos, para a totalidade dos artigos e para as amostras subdivididas de acordo com a área de conhecimento dos artigos. Foi utilizado o software livre de testes estatísticos Jamovi (2021) para executar os testes. Os resultados obtidos podem ser vistos no Quadro 5 (totalidade dos artigos analisados) e Quadro 6 (por área de conhecimento).

Quadro 5 - Intervalo numérico e testes de normalidade para as variáveis da totalidade dos artigos analisados

	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Minimum</i>	0	0	0	0	0	0	0	0
<i>Maximum</i>	157	1131	61	35	175	113244	293994	6
<i>Skewness</i>	20.4	24.3	32.5	4.76	10.7	39.7	72.4	7.97
<i>Std. error skewness</i>	0.0108	0.0108	0.0108	0.0485	0.0297	0.0127	0.0127	0.0108
<i>Kurtosis</i>	995	1058	1413	30.8	284	3305	8194	92.7
<i>Std. error kurtosis</i>	0.0216	0.0216	0.0216	0.0970	0.0594	0.0254	0.0254	0.0216

Fonte: Dados da pesquisa, calculados com o software Jamovi (2021).

Quadro 6 - Intervalo numérico e testes de normalidade para as variáveis dos artigos analisados, por área de conhecimento

Ciências Exatas, Tecnológicas e Multidisciplinares								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Minimum</i>	0	0	0	NaN	0.00	17	0	0
<i>Maximum</i>	22	307	8	NaN	2.00	25190	293994	1
<i>Skewness</i>	14.6	13.2	21.4	NaN	1.77	10.1	51.4	21.9
<i>Std. error skewness</i>	0.0394	0.0394	0.0394	NaN	0.524	0.0427	0.0427	0.0394
<i>Kurtosis</i>	295	282	610	NaN	2.54	184	2831	479
<i>Std. error</i>	0.0787	0.0787	0.0787	NaN	1.01	0.0853	0.0853	0.0787

<i>kurtosis</i>								
Ciências Humanas								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Minimum</i>	0	0	0	0	0	0	0	0
<i>Maximum</i>	57	1131	61	6	15	113244	66149	3
<i>Skewness</i>	20.6	35.6	26.9	4.06	5.30	40.9	18.8	9.12
<i>Std. error skewness</i>	0.0178	0.0178	0.0178	0.112	0.0624	0.0196	0.0196	0.0178
<i>Kurtosis</i>	879	1861	962	20.3	43.1	2785	572	98.0
<i>Std. error kurtosis</i>	0.0355	0.0355	0.0355	0.224	0.125	0.0393	0.0393	0.0355
Ciências Sociais Aplicadas								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Minimum</i>	0	0	0	0.00	0.00	12	0	0
<i>Maximum</i>	79	645	48	6.00	17.0	24986	162628	6
<i>Skewness</i>	27.5	21.4	23.7	3.87	5.54	9.33	54.4	8.42
<i>Std. error skewness</i>	0.0269	0.0269	0.0269	0.107	0.111	0.0317	0.0317	0.0269
<i>Kurtosis</i>	1344	731	731	18.7	42.6	202	3653	105
<i>Std. error kurtosis</i>	0.0537	0.0537	0.0537	0.214	0.221	0.0634	0.0634	0.0537
Ciências da Vida								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Minimum</i>	0	0	0	0.00	0.00	0	0	0
<i>Maximum</i>	157	890	11	35.0	175	26268	81502	6
<i>Skewness</i>	12.2	14.7	15.7	2.27	9.52	14.9	24.9	5.72
<i>Std. error skewness</i>	0.0235	0.0235	0.0235	0.102	0.0395	0.0313	0.0313	0.0235
<i>Kurtosis</i>	344	442	352	7.00	210	385	900	59.0
<i>Std. error kurtosis</i>	0.0471	0.0471	0.0471	0.203	0.0790	0.0625	0.0626	0.0471
Linguística, Letras e Artes								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Minimum</i>	0	0	0	0.00	0.00	9	0	0
<i>Maximum</i>	19	487	42	8.00	16.0	24564	33233	4
<i>Skewness</i>	11.4	29.8	47.6	5.07	5.09	15.0	14.4	5.93
<i>Std. error skewness</i>	0.0256	0.0256	0.0256	0.0782	0.0812	0.0308	0.0308	0.0256
<i>Kurtosis</i>	219	1155	2694	36.8	36.2	367	298	41.8
<i>Std. error kurtosis</i>	0.0511	0.0511	0.0511	0.156	0.162	0.0616	0.0616	0.0511

Fonte: Dados da pesquisa, calculados com o software Jamovi (2021).

Observa-se que, para a amostra total, os resultados para o teste de Kurtosis variam entre 92,7 e 8194, enquanto para o teste de skewness o intervalo vai de 4,76 a 72,4, com a menor distorção resultante do indicador bibliométrico da base *Web of Science (WoS)*, cujo intervalo numérico de número de citações varia de 0 a 35. Os testes de normalidade indicam, portanto, uma distribuição anormal para todas as variáveis da amostra, confirmando que a utilização da correlação de Spearman nas análises é mais adequada.

Para a interpretação e análise dos valores obtidos com os cálculos do coeficiente de correlação de Spearman, foram considerados os valores apresentados no Quadro 7.

Quadro 7 - Intervalos para a interpretação e análise dos Coeficientes de Correlação de Spearman

Valor de r (+ ou -)	Interpretação da Correlação
0 – 0,2	Nula
0,21 – 0,4	Fraca
0,41 – 0,7	Moderada
0,71 – 0,9	Forte
1	Perfeita

Fonte: Callegari-Jacques (2003) e Vieira (2008), adaptado de Silva Filho (2022, p. 91).

Ao analisar as correlações entre os indicadores das publicações divididas por área de conhecimento, algumas correlações apresentaram erro e não foi possível analisá-las. O erro pode ter se dado devido ao número excessivo de zeros em uma das variáveis, ou devido à condicional no cálculo da correlação de descartar a combinação de valores inteiros com campos inexistentes (em branco), não os transformando em zero. Desse modo, em amostras menores para indicadores de baixa cobertura, as correlações apresentaram erro e não foram analisadas. Este foi o caso das correlações entre citações na *Web of Science* e demais indicadores para artigos da área de conhecimento de Ciências Exatas, Tecnológicas e Multidisciplinares, bem como para citações na *Scopus* e indicadores alométricos do *Twitter* e *Facebook*, para artigos desta mesma área.

3.3 LIMITAÇÕES DA PESQUISA

O presente estudo apresenta algumas limitações decorrentes das escolhas e procedimentos metodológicos adotados para o desenvolvimento da dissertação. Entre os aspectos que podem limitar as análises e conclusões deste estudo estão às fontes de dados

selecionadas para fornecer os dados alométricos e bibliométricos, bem como, as soluções utilizadas para cruzar os dados coletados das diferentes fontes.

Quanto à delimitação das fontes de dados selecionadas, ressalta-se a dificuldade da plataforma *Altmetric* em captar a totalidade de menções feitas no *Facebook*. De acordo com Silva, Maricato e Macêdo (2020), a “*Altmetric* rastreia apenas mensagens de perfis públicos da rede social - excluindo, por exemplo, menções de indivíduos e grupos com perfis fechados” (SILVA; MARICATO; MACÊDO, 2020, p. 284). Portanto, postagens mencionando artigos do universo utilizado na presente pesquisa, feitas por usuários que tenham seus perfis fechados, não são contabilizadas como menções no *Facebook* pelo *Altmetric* em decorrência desta impossibilidade de coleta.

A *Altmetric* rastreia as menções no *Facebook* e *Twitter* através dos identificadores persistentes, como o DOI, o que permite que menções em diferentes línguas sejam captadas, mas menções que não tenham o identificador não são possíveis de serem rastreadas. Isso quer dizer que menções de um artigo feitas apenas pelo seu título, por exemplo, não serão captadas pela *Altmetric*. Esse tipo de limitação da plataforma desperta questionamentos a respeito da cobertura de informações recuperadas das fontes de dados como *Twitter* e *Facebook*, pela *Altmetric* (ROBINSON-GARCÍA *et al.*, 2014).

Outra limitação da plataforma *Altmetric* é que esta possui uma transparência limitada na forma como os dados são coletados das diversas fontes e em como é feita a seleção destas fontes (SILVA; MARICATO; MACÊDO, 2020). Ainda assim, visto que plataformas como o *Altmetric* são a uma forma automatizada e unificada de coletar dados alométricos de diferentes mídias sociais, a análise destes dados a partir dessa plataforma permanece válida e os dados confiáveis.

Outro problema que pode ser encontrado diretamente nos dados coletados provenientes das mídias sociais e que podem gerar incompatibilidade no processo de cruzamento de dados são os erros de digitação dos títulos e dos identificadores DOI dos artigos por parte dos usuários das plataformas, o que impede que a coleta automática feita pela *Altmetric* rastreie as menções ao artigo.

Limitações relativas aos relatórios exportados do Open Journal Systems também podem impactar os resultados obtidos, visto que os relatórios que contêm as estatísticas de uso (acessos ao resumo e downloads) dos artigos são gerados por um plugin (PKP DOCS, c2022), nativo do sistema *OJS* a partir de sua versão 3, e armazenados no banco de dados de cada portal de periódicos. A possibilidade de falhas técnicas ocasionando intervalos de tempo sem

estatísticas de uso coletadas deve ser considerada como possível limitação, ainda que não seja possível comprovar se os dados coletados nesta pesquisa foram sujeitos a este tipo de falha.

Outro tipo de limitação relativa ao sistema *OJS* diz respeito às práticas editoriais das revistas. Editores e demais atores do fluxo editorial podem fazer alterações nos metadados dos artigos após sua publicação (como no título), ainda que esta prática não seja recomendável. Esta ação pode gerar uma incompatibilidade entre os dados das diversas fontes no momento do cruzamento dos dados coletados, especialmente dos relatórios de acessos ao resumo e de downloads, processo que depende da correspondência exata entre as strings no título.

Outra limitação enfrentada se refere às citações provenientes da fonte de dados *Crossref*. No momento da coleta dos dados desta pesquisa, todas as referências das mais de 100 milhões de entidades bibliográficas depositadas na *Crossref* eram divididas em abertas, limitadas ou fechadas, de acordo com a escolha da editora (HEIBI; PERONI; SHOTTON, 2019). Isso quer dizer que parte das listas de referências eram fechadas, não estando visíveis para ninguém que não fosse cadastrado como membro da *Crossref* e participante do serviço de registro de referências *Cited-by* (HEIBI; PERONI; SHOTTON, 2019). O fato de que as revistas que registram conteúdo junto à *Crossref* podiam optar por manter a lista de referências dos artigos científicos fechadas pode, assim como a limitação do *Altmetric*, influir nos resultados aqui apresentados, especialmente nas análises de coberturas, que pode não refletir a real cobertura de citações na base de dados *Crossref*, captando somente as citações definidas como abertas.

A *Crossref* também tem problemas em marcar as referências corretamente, o que, juntamente com a ausência de *DOI* nas referências, pode limitar a contagem de citação (CIOFFI *et al.*, 2022). Os erros mais comuns que impedem que uma referência seja marcada pela *Crossref* são referentes a imprecisões causadas pelos autores ao criar a lista de artigos citados em suas pesquisas (erros de prefixo ou sufixo *DOI*), e erros na entrada de dados, gerando falha na criação do link entre o artigo que cita e o que é citado (BUCHANAN, 2006).

4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados obtidos depois de realizadas as análises necessárias para que os objetivos propostos fossem atendidos. Os resultados são apresentados e discutidos e, sempre que possível contextualizados com pesquisas anteriores publicadas pela comunidade científica.

O capítulo está estruturado em três seções principais, sendo a primeira dedicada à análise das coberturas dos indicadores alométricos e bibliométricos nas diferentes fontes de dados, e das correlações dos indicadores alométricos e bibliométricos para todos os artigos do universo; a segunda seção trata da análise das coberturas e correlações entre os indicadores de acordo com a área de conhecimento a qual os artigos pertencem; e a terceira seção analisa as coberturas e correlações entre os indicadores dos artigos de acordo com a data de publicação dos mesmos.

4.1 COBERTURAS E CORRELAÇÕES DE INDICADORES ALTMÉTRICOS E BIBLIOMÉTRICOS PARA OS ARTIGOS DOS PORTAIS

4.1.1 Coberturas de indicadores alométricos e bibliométricos

Para os 51.200 artigos analisados, o total de menções nas fontes de dados alométricos foi de 37.125.439. Ao todo, 46.294 artigos possuem pelo menos uma menção em alguma das fontes de dados alométricas e ao menos um indicador alométrico dos cinco especificados, sendo eles: leitores no *Mendeley*; menções no *Twitter* (ou *tweets*); número de acessos à página de resumo; número de downloads; e, menções no *Facebook*. A cobertura total de menções (para os 46.294 artigos) equivale a 90,41% do total de artigos analisados. As menções, coberturas por fonte de dados alométricas, médias e medianas podem ser observadas na Tabela 2.

Se considerarmos apenas plataformas de mídias e redes sociais, sendo elas o *Mendeley*, *Twitter* e o *Facebook*, o total de menções é 242.191. Ao considerarmos dados alométricos de acessos à página de resumo e downloads do arquivo do artigo, o total é de 36.883.248. A partir destes totais, é possível observar que o volume total de acessos ao resumo e downloads do artigo é muito maior do que as menções em mídias e redes sociais.

Ao analisarmos a cobertura dos dados alométricos para os 51.200 artigos, observa-se que o indicador de acessos ao resumo é o indicador com maior cobertura por artigo, com

37.263 (72,77%) artigos tendo ao menos um acesso ao resumo. O número de downloads é o segundo indicador altmétrico com maior cobertura, com 37.255 (72,76%) no total. Nota-se que o percentual de cobertura de acessos ao resumo e downloads são próximos, o que sugere uma relação entre os dois indicadores (a diferença entre a cobertura de menções do número de artigos com algum acesso ao resumo e do número de artigos com algum download efetuado é de apenas oito).

Tabela 2 - Total de menções, coberturas, média e mediana dos artigos, por fonte de dados altmétricos

Fonte de Dados	Total de Menções (TM)	Cobertura de menções (CM)	% de Cobertura	Média de menções/Artigo	Mediana de menções/Artigo
<i>Mendeley</i>	236.101	28.532	55,72%	4,61	1
<i>Twitter</i>	4.276	1.649	3,22%	0,08	0
Acesso Resumo	16.632.689	37.263	72,77%	446,36	227
Downloads	20.250.559	37.255	72,76%	543,57	235
<i>Facebook</i>	1.814	1.572	3%	0,04	0
Total	37.125.439	46.294	90,41%	725,05	273

Fonte: Dados da pesquisa.

Legenda: TM é entendido como a soma de menções recebidas por todos os artigos.

CM é entendida como o número de artigos que foram mencionados ou lidos/capturados ao menos uma vez.

O total de menções no *Mendeley*, que se refere ao número de vezes que os artigos (28.532) foram lidos/capturados pelos usuários do *Mendeley*, totalizou 236.101. A cobertura (número de artigos lidos/capturados) foi de 55,72%, sendo a terceira maior cobertura dentre as menções analisadas. Embora superior, a cobertura encontrada no *Mendeley* se aproxima dos resultados obtidos por Mohammadi *et al.* (2014), que encontrou cobertura de 45,6%, para mais de 480 mil artigos de cinco áreas de conhecimento indexados na *Web of Science*. Assim como a cobertura identificada por Haustein *et al.* (2014), de 66,2%, para 1,4 milhão de publicações indexadas na *PubMed* e *Web of Science*; e por Hammarfelt (2014), com 61%, para artigos de 30 universidades suecas indexados na *SwePub*²⁰.

A cobertura de leitores no *Mendeley* de 55,72% descoberta na presente pesquisa, também é similar à de 65,9% encontrada por Haustein e Larivière (2014). Já Zahedi e Haustein (2018) encontram uma cobertura superior de leitores no *Mendeley*, de 84,2%, para mais de 1,3 milhão de artigos indexados na *Web of Science*. Entendemos que esses dados são relevantes, visto que o *Mendeley*, comparativamente ao cenário internacional, apresenta indícios de que os seus indicadores podem ser úteis para monitorar impactos da ciência presente nos portais de periódicos.

²⁰ Disponível em: <https://swepub.kb.se/>. Acesso em: 28 jul. 2022.

Seguindo a ordem decrescente de número total de menções por fonte de dados, o *Twitter* totalizou 4.276 menções/*tweets* para 1.649 artigos, uma cobertura equivalente a apenas 3,22% do total de artigos. Não muito distante deste encontra-se o *Facebook*, com um total de menções de 1.814 para 1.572 artigos, ou seja, uma cobertura correspondente a 3% dos 51.200 artigos, ocupando a posição de fonte de dados com a menor cobertura.

A cobertura de menções no *Twitter* e *Facebook* é menor do que a obtida em estudos anteriores, como o de Haustein, Costas e Larivière (2015), que obtiveram 21,5% de cobertura para o *Twitter*, e 4,7% para o *Facebook*, ao analisarem 1,3 milhão de publicações indexadas na *Web of Science*. Já Robinson-García *et al.* (2014) identificaram uma cobertura no *Twitter* um pouco menor, de 16%, para publicações com DOI indexadas na *Web of Science*. Zahedi, Costas e Wouters (2014) também identificaram uma baixa cobertura no *Twitter*, de 13,3% e no *Facebook*, de 2,5%, para uma amostra de 20.000 publicações indexadas no *ImpactStory*.

Em uma análise de publicações latino-americanas, Alperin (2015) identificou uma cobertura de 8% para o *Twitter*, resultado distante, porém, um pouco mais próximo do obtido no presente estudo, o que parece refletir uma realidade de pouco uso das mídias sociais por acadêmicos de países latinos, ou uma cultura diferente de compartilhamento de pesquisas em mídias sociais entre leitores de pesquisas da América Latina (ALPERIN, 2015).

A baixa cobertura no *Twitter* dos artigos dos portais de periódicos analisados pode, também, estar relacionada às diferenças de divulgação realizadas pelos portais, em comparação com as revistas com fins comerciais. A circulação da ciência no *Twitter*, enquanto ferramenta que pode contar com um público mais diverso, pode ser afetada, ainda, pela temática dos artigos. Assim, é possível que os artigos e temáticas publicadas pelas revistas dos portais possam ser de menor interesse social. Por último, uma hipótese possível poderia estar relacionada a baixa qualidade dos artigos indexados nestes portais. No entanto, acredita-se que esta é questionável, visto que os artigos têm indicadores semelhantes aos do cenário internacional em relação ao *Mendeley*, uma ferramenta que agrupa muitos estudantes e pesquisadores.

A partir desses resultados, que consideram a totalidade dos artigos independente da revista ou área de conhecimento a qual pertençam, podemos perceber que as coberturas mais altas são, respectivamente, as de acessos ao resumo, downloads e leitores no *Mendeley*. Estas três fontes de dados cobrem, cada uma, mais da metade dos artigos analisados. Ainda que não seja possível inferir a partir deste resultado isoladamente que estas seriam as três melhores fontes de dados altmétricas a serem consideradas na análise de indicadores altmétricos de artigos de portais de periódicos brasileiros de acesso aberto. Este resultado concorda com

outros estudos (DEMACHKI; MARICATO, 2022b), indicando alguma possibilidade de adoção destes indicadores altmétricos para monitoramento científico destas publicações.

Ao analisarmos as médias e medianas para as fontes de dados altmétricas, é possível observar que os acessos ao resumo e downloads possuem médias superiores quando comparadas às demais fontes, com médias 446,36 e 543,57, respectivamente. As medianas de acesso ao resumo e downloads são de 227 e 235, respectivamente. Como as médias são numericamente superiores às medianas, evidencia-se que a distribuição numérica da amostra de artigos com algum acesso ao resumo ou downloads possui assimetrias.

Em outras palavras, a comparação da média e da mediana tanto para o indicador de acesso à página de resumo quanto para o indicador de downloads, indica que uma minoria de artigos possui muitos acessos ao resumo e/ou downloads, enquanto a maioria possui poucos. Neste sentido, constatamos que metade dos artigos cobertos por estes indicadores possuem acessos ao resumo e downloads inferiores a 227 e 235, respectivamente.

Comparativamente com resultados anteriores, Moed (2005) obteve média de 342 downloads para artigos da revista *Tetrahedron Letters*, e Chen e Bukhari (2019) obtiveram média de 928 downloads por artigo da revista *Reumathology*. Wang *et al.* (2014) obtiveram média de acessos/visualizações ao artigo de 3.854 para mais de 56 mil artigos do portal PLOS, publicados entre 2009 e 2012.

Entre as três mídias sociais consideradas, o *Mendeley* apresentou a maior média de leitores/capturas, de 4,61. Entretanto, esta média de leitores por artigo no *Mendeley* (4,61) é inferior à de 11,15 encontrada por Zahedi, Costas e Wouters (2017) e a de 11 por Zahedi e Haustein (2018). O *Twitter* apresentou média de menções por artigo de 0,08, e o *Facebook*, de 0,04. Tanto o *Mendeley*, quanto o *Twitter* e *Facebook* apresentaram uma mediana 0. Porém, as médias para estas três fontes de dados parecem indicar que os valores que estão no topo da distribuição (maiores ocorrências) não estão tão distantes do centro quanto para os acessos ao resumo e downloads.

Quanto aos dados bibliométricos de citação, extraídos da fonte de dados *Crossref* e das bases de dados *Scopus* e *Web of Science*, estes podem ser visualizados na Tabela 3. Observa-se que o total de citações nas três bases analisadas é de 44.482, as quais estão distribuídas entre 10.149 artigos, o que equivale a uma cobertura de 19,82%. Em outras palavras, considerando as três bases, aproximadamente 20% dos artigos foram citados ao menos uma vez.

Para a *Crossref*, o total de citações foi de 25.182, para 8.891 artigos, o que corresponde à cobertura de 17,36% do total de artigos. Esta cobertura é menor que a de 28%

encontrada por Martín-Martín *et al.* (2021) para 2.515 publicações altamente citadas, e maior do que a cobertura de 14,7% para 11.955 artigos do Portal de Periódicos da UFG encontrada por Demachki e Maricato (2022b). Já para o *Scopus*, o total de citações é de 16.615, sendo que 3.564 artigos possuem ao menos uma citação nesta base de dados (6,96% do total). A fonte de dados com a menor quantidade de citações coletadas foi a *Web of Science*, com 2.685 citações distribuídas entre 737 artigos, uma cobertura de 1,44% dos artigos totais.

As coberturas na *Scopus* e *Web of Science* também são inferiores às obtidas por Martín-Martín *et al.* (2021), de 57% e 52% respectivamente. Entretanto, devido as publicações estudadas por Martín-Martín *et al.* (2021) serem altamente citadas, é esperada uma alta cobertura de citação para a amostra. Para uma amostra similar à do presente estudo, considerando 12.248 artigos do Portal de Periódicos da UFG, Demachki e Maricato (2022a) encontraram coberturas na *Scopus* e *Web of Science* de 10,9% e 3,3%, respectivamente, valores mais próximos aos resultados obtidos na presente pesquisa.

Ao observarmos que a cobertura de citação nas três bases (19,82%) é menor que a soma dos percentuais de cobertura individual das bases (25,76%), fica perceptível que as fontes de dados captam citações diferentes, evidenciando a complementariedade entre as fontes. Visser, Van Eck e Waltman (2021) identificaram que a *Scopus* cobre um grande número de artigos que não possuem citações captadas pela *Crossref*, embora a *Crossref* cubra um número ainda maior de publicações de revistas sem citações na *Scopus*. Em relação à *Scopus* e a *Web of Science*, os autores identificaram que a primeira (*Scopus*) cobre um número maior de publicações científicas do que a segunda (*Web of Science*). Essas relações entre as fontes de dados evidencia a complementariedade entre elas, reafirmada nos resultados no presente estudo, ainda que com baixas coberturas de citação. Como boa parte das revistas de portais de periódicos de universidades brasileiras não é indexada nestas bases, a *Crossref* tem o potencial de se tornar uma fonte de citação relevante para artigos deste tipo de portal.

Quando analisamos a cobertura de indexação dos artigos nas fontes de dados bibliométricas, a *Crossref* possui uma cobertura de 100% dos artigos da amostra indexados nesta base. Essa alta e perfeita cobertura se deve ao fato de que todos os artigos da amostra possuem um número de *DOI*, que é registrado pela própria *Crossref*. Da totalidade de artigos dos portais, 17,36% possuem alguma citação na *Crossref*. As coberturas de indexação nas três fontes de dados bibliométricos consideradas, assim como as médias e medianas de citações por artigo indexado, podem ser vistas na Tabela 3.

O quantitativo de citações na *Crossref* provavelmente é subestimado, podendo ser influenciada por alguns fatores, como a falta de padronização e marcação das referências por

parte das revistas (CIOFFI *et al.*, 2022; BUCHANAN, 2006); e pelo fato de diversas editoras ainda manterem suas referências fechadas no momento da coleta dos dados. Isso ocorre devido a algumas editoras não terem aderido ao movimento de citação aberta, não fornecendo autorização para a divulgação das citações que suas publicações recebem e que são registradas pela *Crossref*. Algumas pesquisas consideram que o indicador de número de citações na *Crossref* reflete apenas parcialmente as relações de citação dos documentos registrados nessa fonte (DEMACHKI; MARICATO, 2022b).

Tabela 3 - Coberturas de indexação, médias e medianas de citações nas fontes de dados bibliométricas

Fonte de Dados	Total de artigos	Total de citações	CC	%CC	CI	% CI	%CCAI	MeCAI	MdCAI
<i>Crossref</i>		25.182	8.891	17,36%	51.200	100%	17,36%	0,49	0
<i>Scopus</i>		16.615	3.564	6,96%	6.787	13,25%	52,51%	2,44	1
<i>WoS</i>		2.685	737	1,44%	2.546	4,97%	28,95%	1,05	0
Total	51.200	44.482	10.149	19,82%	51.200	100%	19,87%	0,86	0

Fonte: Dados da pesquisa.

CC = Cobertura de citação total; CI = Cobertura de indexação; CCAI = Cobertura de citação entre artigos indexados; MeCAI = Média de citações por artigo indexado; MdCAI = Mediana de citações por artigo indexado.

A *Scopus* é a segunda base com maior cobertura de indexação, com 6.787 artigos indexados, o que corresponde a uma cobertura de 13,25% do total de artigos do universo analisado. Percebe-se que mais da metade dos artigos indexados na *Scopus* possuem alguma citação capturada por esta base (52,51%), e comparando as médias e as medianas de citações para os artigos indexados, de 2,44 e 1, respectivamente, indicam que esta fonte de dados possui distribuição mais assimétrica. De fato, dos 6.787 artigos indexados na *Scopus*, apenas 1.947 possuem mais citações na base do que a média (2,44). Ainda que a cobertura de citação na *Crossref* seja a maior entre as fontes, as médias de citação por artigo nas bases de dados *Scopus* e *Web of Science* são as maiores entre as três fontes bibliométricas, o que pode estar relacionado a fatores como a qualidade das revistas e ao grande potencial de divulgação que estas duas bases proporcionam aos artigos indexados (DEMACHKI; MARICATO, 2022a).

A *Web of Science (WoS)* se apresenta como a fonte de dados bibliométricos com menor cobertura, tanto em citação quanto em indexação. Dos 51.200 artigos analisados, somente 2.546 estão indexados na *Web of Science*, o que representa uma cobertura percentual de 4,97%. Destes artigos indexados na *Web of Science*, apenas 28,95% possuem alguma citação nesta base. A média de citação para os artigos indexados na *Web of Science* é de 1,05, enquanto a mediana é 0.

Estes resultados posicionam a *Scopus* e a *Web of Science* em segundo e terceiro lugar, respectivamente, em termos de cobertura entre as três fontes de dados bibliométricas consideradas neste estudo. A diferença entre as coberturas de citação na *Crossref* (17,36%) e na *Scopus* (6,96%) evidencia que a *Crossref* é a fonte de dados que abrange a maior quantidade de artigos citados, sendo a mais promissora para ser considerada em estudos futuros de indicadores de citação para artigos de portais de periódicos brasileiros de acesso aberto.

As baixas coberturas de indexação e citação na *Scopus* e *Web of Science* aqui apresentadas, confirmam que há uma sub-representação da ciência brasileira nestas fontes, considerando que os periódicos do país são responsáveis por publicar parte importante da ciência que desenvolve (PACKER, 2011). Resultados semelhantes de baixa cobertura de revistas e artigos brasileiros indexados em bases de dados bibliográficas comerciais foram previamente obtidos por Melo, Trinca e Maricato (2021), que identificaram que mais da metade (51,66%) dos artigos, presentes na plataforma Sucupira, de pesquisadores afiliados a instituições brasileiras, não estavam indexados na *Web of Science*.

Ao compararmos os resultados de cobertura de citação de artigos indexados com pesquisas realizadas no cenário internacional, a baixa cobertura para os artigos dos portais de periódicos brasileiros, participantes deste estudo, em grandes bases comerciais, fica ainda mais evidente. Enquanto Haustein, Costas e Larivière (2015) encontraram uma cobertura de citações de 66,8% para 1,3 milhão de artigos indexados na *Web of Science*, a cobertura de citação obtida no presente estudo para artigos indexados na base foi de 29%.

A menor cobertura dos indicadores bibliométricos para artigos da América Latina é justificada por um menor uso destas publicações (ALPERIN, 2015), comparativamente com artigos publicados em inglês ou em revistas da América do Norte e Europa, pelas quais as bases de dados *Web of Science* e *Scopus* parecem ter uma preferência de indexação (TENNANT, 2020). De modo geral, a *Crossref* se mostrou a fonte de dados bibliométricos com maior potencial de cobertura de publicação e citação para o cenário de publicações em portais de periódicos brasileiros de acesso aberto, ainda que com a possibilidade desta cobertura ter sido reduzida devido a referências limitadas de algumas revistas, como as do Portal de Periódicos da UFG, comprometendo a visualização do cenário completo das relações de citação captadas pela *Crossref*.

Com a nova política de referências abertas da *Crossref*, adotada a partir de junho de 2022, todas as referências fornecidas pelos membros como parte de seus metadados serão automaticamente abertas (BLOG CROSSREF, 2022). Desse modo, revistas e editoras

científicas devem adicionar a lista de referências de seus artigos aos metadados enviados à *Crossref* para registro do DOI, com vistas a aumentar a cobertura de citação captada por esta base, contribuindo para que o cenário integral das citações captadas por esta fonte de dados se reflita nos indicadores de citação.

4.1.2 Correlações entre indicadores alométricos e bibliométricos

A análise das correlações entre indicadores alométricos e bibliométricos abre caminho para entendimento se os indicadores alométricos podem resultar em citações futuras. Por outro lado, a inexistência ou correlações baixas indicam para a complementaridade dos dois tipos de indicadores. Diversos estudos anteriores consideram interessante o uso conjunto destas métricas para analisar o impacto da produção científica argumentando que elas podem medir outro tipo de impacto. Assim, esta seção é dedicada à análise das correlações entre os indicadores alométricos e bibliométricos para os artigos dos portais de periódicos das universidades participantes deste estudo.

Ao analisarmos as correlações de Spearman dispostas na Tabela 4, é possível observar que as correlações entre todos os indicadores bibliométricos de citação são fortes. A correlação entre citações da *Crossref* e da *Web of Science* é de 0,743, bem como entre as citações da *Crossref* e da *Scopus*. De fato, estes resultados estão de acordo com o encontrado por Liu *et al.* (2013), que obtiveram uma correlação forte de 0,738 entre citações da *Crossref* e da *Scopus* para mais de 33 mil artigos extraídos do portal de periódicos *PLOS*. Isso sugere que, em grande medida, as citações são captadas de maneira semelhante pelas duas bases.

A matriz de correlação entre os indicadores da *Crossref* e da *Scopus* pode ser visualizada na Figura 18, e a matriz de correlação entre os indicadores da *Crossref* e da *Web of Science* é ilustrada na Figura 19. A correlação entre as citações da *Scopus* e da *Web of Science* é de 0,901, indicando uma correlação quase perfeita entre estes indicadores, com a matriz desta correlação ilustrada na Figura 20. As Figuras 13 a 41 ilustram as matrizes de correlação entre os indicadores alométricos e bibliométricos para todos os artigos analisados.

Tabela 4 - Correlações de Spearman entre indicadores altmétricos e bibliométricos para todos os artigos

	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Crossref</i>	—							
<i>Mendeley</i>	0.429	—						
<i>Twitter</i>	0.060	0.073	—					
<i>WoS</i>	0.743	0.496	-0.001	—				
<i>Scopus</i>	0.743	0.520	0.054	0.901	—			
Acesso Resumo	0.132	0.325	0.059	-0.020	0.045	—		
Downloads	0.123	0.347	0.002	0.074	0.103	0.806	—	
<i>Facebook</i>	0.097	0.104	0.224	-0.057	0.136	0.041	0.007	

Fonte: Dados da pesquisa, calculados com o software Jamovi (2021).

Correlações fortes entre as citações, assim como as encontradas neste estudo, indicam que há uma influência entre os indicadores envolvidos, com uma tendência positiva de que, quando as citações em uma fonte de dados aumentam, as citações capturadas por outra fonte de dados também aumentam para o mesmo artigo. Deve-se ressaltar que as baixas coberturas de citação nas fontes de dados *Web of Science* e *Scopus* não influenciam na força das correlações encontradas, indicando, provavelmente, similitudes entre as citações provindas das três bases de dados bibliométricas analisadas.

De fato, estudos anteriores identificaram uma sobreposição de 42% entre citações da *Crossref* e da *Scopus*, e de 44% entre citações da *Crossref* e *Web of Science*, para 3 milhões de artigos altamente citados de diferentes áreas de conhecimento e publicados em 2006 (MARTÍN-MARTÍN *et al.*, 2021). Entre as citações da *Scopus* e da *Web of Science*, a sobreposição encontrada por Martín-Martín *et al.* (2021) foi de 78%. Em um estudo anterior, Martín-Martín *et al.* (2018) também encontraram uma sobreposição de 54,6% das citações da *Scopus* e da *Web of Science*, para uma amostra de mais de 2 milhões de artigos altamente citados de diferentes áreas.

Outra correlação que se destaca pela sua força é entre os indicadores altmétricos de acesso ao resumo e de downloads, com $\rho = 0,806$. Essa correlação indica que o número de downloads de um artigo é altamente influenciado pela quantidade de acessos à página de resumo do mesmo, o que nos leva a considerar que as informações básicas do artigo contidas na página de resumo, como o seu título e o próprio resumo, são um fator decisivo para os usuários na tomada de decisão de acessar o conteúdo completo do artigo. Desse modo, a partir

do momento em que o usuário decide por fazer o download do artigo, pressupõe-se que o conhecimento científico contido no artigo é do seu interesse.

Uma vez que aos indicadores altmétricos refletem o impacto acadêmico para além da comunidade científica, espera-se que os downloads de um artigo não sejam feitos exclusivamente por acadêmicos e pesquisadores que virão a citá-lo (se o acharem relevante) em seus trabalhos. De acordo com resultados de estudos anteriores, a correlação entre downloads e citações varia entre nula a fraca (LIU *et al.*, 2013, CHEN *et al.*, 2020, MOED, 2005, MOED; HALEVI, 2016). Ao se analisar as correlações entre o indicador altmétrico de downloads e as citações para a amostra, o valor do coeficiente de Spearman varia de $\rho = 0,074$ a $\rho = 0,123$, com a correlação mais fraca entre downloads e citações da *Web of Science*, e a mais forte (porém, ainda de força nula) entre downloads e citações da *Crossref*. Estes resultados estão de acordo com a correlação fraca de 0,22 encontrada por Moed (2005), entre citações extraídas do *Institute for Scientific Information* (ISI) e downloads para artigos da revista *Tetrahedron Letters*, ainda que as do presente estudo sejam ligeiramente mais fracas. Entretanto, a correlação entre downloads e citações da *Web of Science* difere da correlação forte de 0,7 encontrada por Vaughan, Tang e Yang (2017), que analisaram a correlação entre número de downloads e citações para artigos dos 150 periódicos com os maiores JIFs, de cinco disciplinas diferentes.

Ao compararmos a correlação obtida de 0,123 entre o indicador de número de downloads e de citações da *Crossref* com a de 0,378 obtida por Liu *et al.* (2013), entre estes mesmos indicadores, é possível perceber que a correlação obtida no presente estudo é mais fraca. Quanto às citações na *Scopus*, Moed e Halevi (2016) encontraram uma correlação moderada de 0,387 entre downloads de artigos publicados em 62 revistas do portal *ScienceDirect* e citações no *Scopus*, enquanto a correlação obtida no presente estudo foi de 0,103. A correlação nula de 0,103 entre downloads e citações na *Scopus* também se distancia da correlação moderada de 0,444 obtida por Liu *et al.* (2013); e da correlação de 0,694 obtida por Wang *et al.* (2014).

Há vários fatores que influenciam a força da correlação entre downloads e citações, como o tempo entre a citação e o download (MOED, 2005), a relação entre a frequência de downloads e a frequência de citações (XUE-LI; HONG-LING; MEI-YING, 2011), e a qualidade do artigo (THELWALL; SUD, 2016). Além destes, Moed e Halevi (2016) descrevem outros 10 fatores que podem causar divergências nas análises entre número de downloads e citações, o que nos mostra que a correlação entre estes indicadores necessita de uma análise detalhada levando em consideração os diversos fatores envolvidos, antes de se

afirmar se a correlação entre os indicadores pode ser efetivamente utilizada para mensurar impacto científico.

Ainda que resultados anteriores tenham apontado para uma relação de causa e efeito entre indicadores altmétricos de uso, como acessos e downloads, e citações, a presente amostra não demonstra a existência desta relação, demonstrando a complexidade das relações entre os indicadores altmétricos e de citação (DEMACHKI; MARICATO, 2022b). Para Thelwall e Sud (2016), uma correlação perfeita entre indicadores altmétricos e bibliométricos de citação é teoricamente impossível de acontecer, devido estes serem indicadores indiretos de medição do fenômeno de interesse, o impacto da pesquisa.

Thelwall e Sud (2016) também consideram que a força das correlações pode ser reduzida simplesmente pela mistura de diferentes disciplinas em uma mesma amostra de dados. Junto a estes fatores, diversas outras propriedades de um documento científico podem influenciar na força da correlação entre indicadores altmétricos e de citação, como a nacionalidade dos autores, idioma do texto e a legibilidade do resumo (THELWALL; SUD, 2016).

No caso do universo de artigos do presente estudo, os fatores citados por Thelwall e Sud (2016) são relevantes para compreender e, em parte, justificar as correlações nulas encontradas entre o indicador altmétrico de acesso ao resumo e todos os indicadores de citação, assim como entre o indicador altmétrico de downloads e todos os indicadores de citação das bases de dados consideradas. Uma vez que os artigos são publicações de periódicos brasileiros, espera-se que a maioria destes seja publicada em português, restringindo o alcance das pesquisas pelo fator idioma.

Assim, o fato de que a maioria dos artigos são publicações em português pode influenciar na relação existente entre leitura e citação. Para serem citados, artigos precisam ser lidos, o que demanda que sejam baixados primeiro (MOED; HALEVI, 2016). Se o artigo está em português, os downloads deste serão feitos, em sua maioria, por usuários que compreendam português, com uma provável chance de levarem a futuras publicações também em português do que em outros idiomas.

Estes fatores acima discutidos influenciam toda a amostra do presente estudo, assim como influenciaram os resultados de estudos anteriores, e não devem nunca ser interpretados como fatores que invalidam as correlações obtidas. Afinal, os indicadores altmétricos são justamente uma forma dinâmica e diversificada de se medir o impacto científico para além das fronteiras da comunidade acadêmica, então é natural que as dinâmicas de usabilidade, as

características inerentes ao artigo, e o tempo de publicação influenciem na força das correlações entre os indicadores.

A correlação entre posts no *Facebook* e todos os outros indicadores altmétricos e de citação analisados é nula, exceto para o *Twitter*, cuja correlação obtida é de 0,224. Esta correlação fraca é similar à de 0,256 obtida por Costas, Zahedi e Wouters (2014) para 1,3 milhão de publicações indexadas no Altmetric. Esta correlação sofre influência de diversos fatores que podem limitar a identificação de menções em redes sociais pela plataforma Altmetric, como a identificação de menções apenas penas em perfis públicos no *Facebook*, e erros como título da publicação mencionada incompleto ou impreciso, problemas com o link que direciona para o artigo e DOI inválido (YU *et al.*, 2021).

A correlação entre o indicador de leitores no *Mendeley* e citações na *Crossref* é moderada, com $\rho = 0,429$. A correlação moderada de 0,496 entre o *Mendeley* e citações da *Web of Science* se equipara a correlação de 0,456 encontrada por Haustein *et al.* (2014) para artigos indexados em ambas as fontes de dados. A correlação aqui encontrada também se assemelha a outras correlações moderadas de estudos anteriores, como de Zahedi e Haustein (2018), que encontraram correlações variando entre 0,452 e 0,638 para mais de 1 milhão de artigos divididos por área e indexados na *Web of Science*; e a de 0,512 obtida por Haustein e Larivière (2014), para cerca de 1 milhão de artigos extraídos através da API do *Mendeley*.

Os leitores no *Mendeley* também se correlacionam moderadamente com as citações na *Scopus*, com uma correlação de Spearman de 0,520. Autores como Bar-Ilan (2012) e Yang, Xing e Wolfram (2018) também encontraram correlações moderadas entre estes indicadores, de 0,448 e 0,563, respectivamente. A correlação obtida também é similar às obtidas por Thelwall e Sud (2016), que subdividiram as correlações entre estes indicadores por área de conhecimento, variando entre 0,30 e 0,72 para 52 áreas de conhecimento, sendo a maioria (94,23%) das correlações encontradas moderadas.

As correlações moderadas entre os leitores no *Mendeley* e as citações das três fontes de dados bibliométricas consideradas no presente estudo reforçam que estes dois indicadores estão medindo um tipo similar de impacto, ainda que sejam processos diferentes (ZAHEDI; COSTAS; WOUTERS, 2017). Correlações moderadas entre estes indicadores são um resultado que se repete nos estudos altmétricos, indicando que os leitores no *Mendeley* são uma forma de identificar artigos com citações. O estudo de Zahedi e Haustein (2018) também afirma que existe relação entre o quantitativo de citações e de leitores no *Mendeley*, uma vez que estes dois indicadores são afetados por variáveis semelhantes.

As correlações entre os leitores no *Mendeley* e os acessos ao resumo e downloads são fracas, de 0,325 e 0,347 respectivamente, sendo estas as únicas correlações não nulas entre estes dois indicadores altmétricos de uso demais indicadores analisados. Entretanto, a forte correlação de 0,803 entre os indicadores de acessos ao resumo e de downloads indica que estes indicadores possuem uma relação de causa e efeito relacionada ao comportamento do usuário. É esperado que, quando o usuário acessa a página de resumo do artigo, este tome a decisão de acessar o arquivo completo (ação de *download*) se a prévia disponível na página de resumo indicar que o conteúdo é do seu interesse.

Além deste comportamento, espera-se que a correlação entre estes indicadores seja mais forte à medida que a ação de fazer o *download* do artigo seja um efeito causado pelo acesso à página de resumo do mesmo. Sendo assim, a correlação entre acessos ao resumo e downloads indica que as informações contidas na página de resumo são um fator que desperta o interesse do usuário e o leva a fazer o download do artigo.

Quando comparamos as correlações entre indicadores de fontes de dados de maior cobertura, nota-se que o volume de menções e cobertura não influencia na força da correlação entre indicadores altmétricos e as citações. Isso pode ser observado para o número de downloads e de acessos ao resumo do artigo que, apesar de terem alta cobertura, possuem correlação nula com as citações das três fontes de dados consideradas.

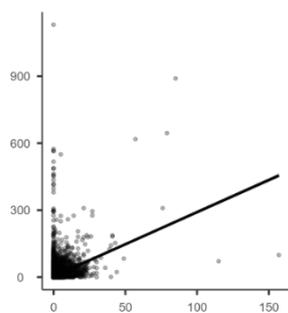
Essa observação nos ajuda a perceber que a relação entre leitores no *Mendeley* e citações merece, a princípio, ser mais bem compreendida por este ser o único indicador altmétrico que se correlacionou com força moderada com as citações para os artigos do universo. É interessante observar que a correlação entre os leitores no *Mendeley* e as citações é mais forte do que entre os leitores no *Mendeley* e os demais indicadores altmétricos. Ainda que indicadores bibliométricos se correlacionem melhor entre si, e que o mesmo valha para indicadores altmétricos (COSTAS; ZAHEDI; WOUTERS, 2014). As correlações obtidas no presente estudo indicam um comportamento diferente para o indicador de leitores do *Mendeley*, também percebida por outros estudos (BARCELOS; MARICATO, 2021).

Assim como Barcelos e Maricato (2021) identificaram um número maior de leitores no *Mendeley* do que de menções em redes sociais digitais para artigos publicados na revista *Scientific Data*, também observamos um número maior de leitores no *Mendeley* do que de menções no *Twitter* e no *Facebook* para os artigos da presente pesquisa. Este resultado, juntamente com as análises de coberturas e correlações moderadas entre leitores no *Mendeley* e citações, nos mostra que há uma relação entre estes indicadores que merece ser ainda mais investigada em análises que considerem outros fatores que influenciem nesta relação.

Ainda que o indicador de leitores no *Mendeley* possua correlações mais fortes com citações do que os demais indicadores altmétricos, indicando, em certa medida, a possibilidade de explicar citações futuras, há a necessidade de análises levando em consideração outros fatores, como a influência do tempo de publicação do artigo e a área de conhecimento a qual ele pertence.

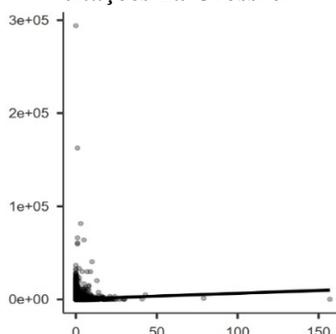
Para isso, as próximas seções realizam análises considerando diferenças entre as áreas de conhecimento (seção 4.2) e entre o tempo de publicação dos artigos (seção 4.3), para analisar as relações entre os indicadores altmétricos e de citação.

Figura 13 - Matriz de correlação entre leitores do Mendeley e citações na Crossref



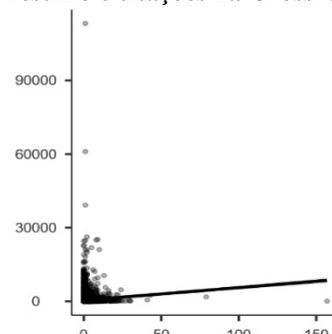
Fonte: Dados da pesquisa (2022).

Figura 15 - Matriz de correlação entre downloads e citações na Crossref



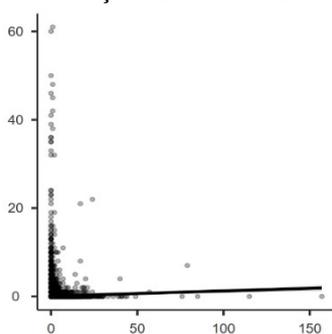
Fonte: Dados da pesquisa (2022).

Figura 17 - Matriz de correlação entre acessos ao resumo e citações na Crossref



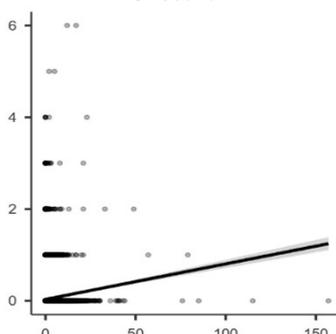
Fonte: Dados da pesquisa (2022).

Figura 14 - Matriz de correlação entre tweets e citações na Crossref



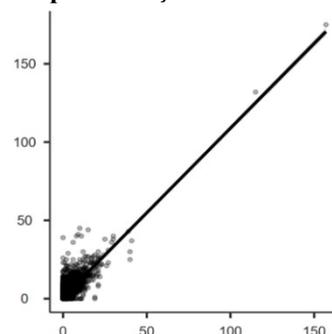
Fonte: Dados da pesquisa (2022).

Figura 16 - Matriz de correlação entre menções no Facebook e citações na Crossref



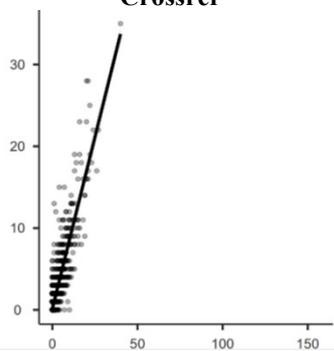
Fonte: Dados da pesquisa (2022).

Figura 18 - Matriz de correlação entre citações Scopus e citações na Crossref



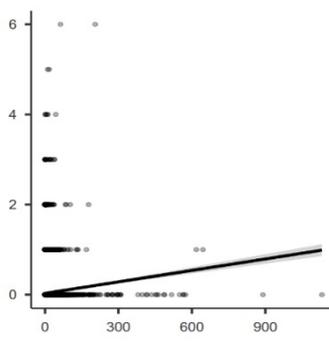
Fonte: Dados da pesquisa (2022).

Figura 19 - Matriz de correlação entre citações na Web of Science e citações na Crossref



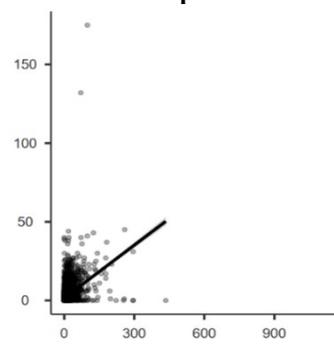
Fonte: Dados da pesquisa (2022).

Figura 22 - Matriz de correlação entre leitores no Mendeley e menções no Facebook



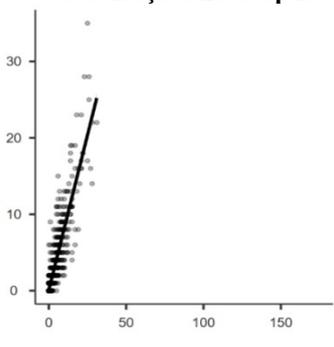
Fonte: Dados da pesquisa (2022).

Figura 25 - Matriz de correlação entre leitores no Mendeley e citações na Scopus



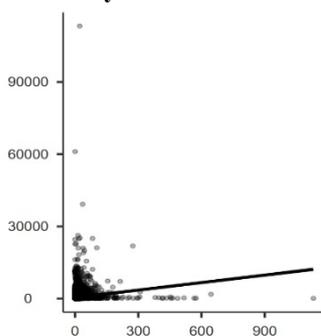
Fonte: Dados da pesquisa (2022).

Figura 20 - Matriz de correlação entre citações na WoS e citações na Scopus



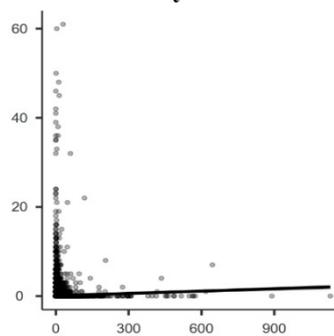
Fonte: Dados da pesquisa (2022).

Figura 23 - Matriz de correlação entre leitores no Mendeley e acesso ao resumo



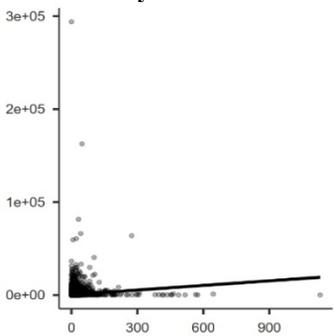
Fonte: Dados da pesquisa (2022).

Figura 26 - Matriz de correlação entre leitores no Mendeley e tweets



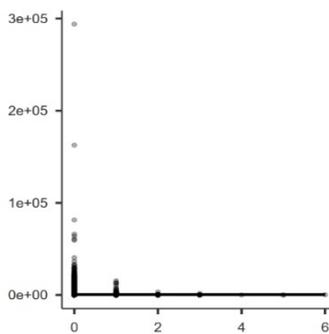
Fonte: Dados da pesquisa (2022).

Figura 21 - Matriz de correlação entre leitores no Mendeley e downloads



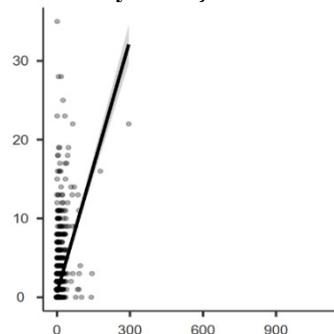
Fonte: Dados da pesquisa (2022).

Figura 24 - Matriz de correlação entre downloads e Facebook



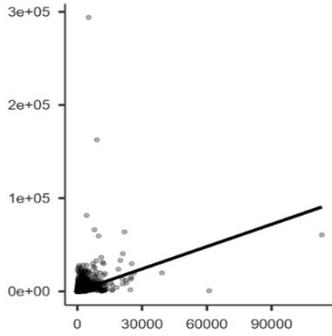
Fonte: Dados da pesquisa (2022).

Figura 27 - Matriz de correlação entre leitores no Mendeley e citações na WoS



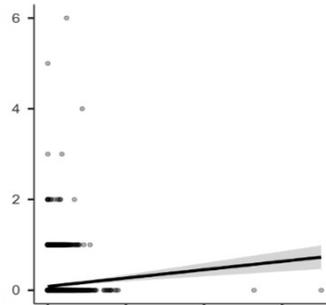
Fonte: Dados da pesquisa (2022).

Figura 28 - Matriz de correlação entre acessos ao resumo e downloads



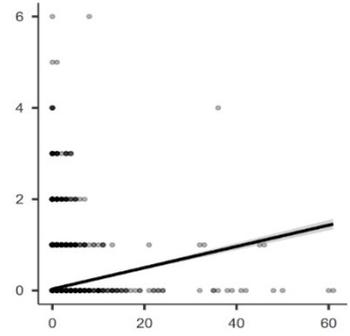
Fonte: Dados da pesquisa (2022).

Figura 31 - Matriz de correlação entre menções no Facebook e citações na Scopus



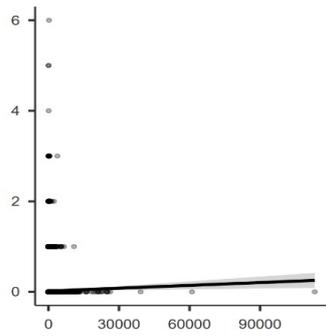
Fonte: Dados da pesquisa (2022).

Figura 34 - Matriz de correlação entre tweets e menções no Facebook



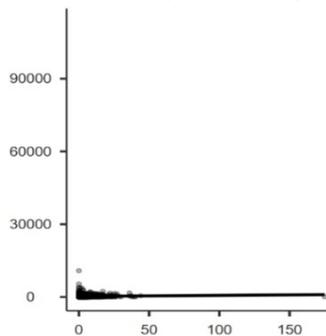
Fonte: Dados da pesquisa (2022).

Figura 29 - Matriz de correlação entre acessos ao resumo e menções no Facebook



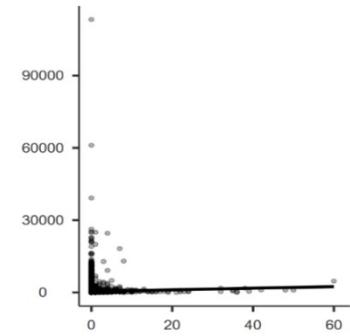
Fonte: Dados da pesquisa (2022).

Figura 32 - Matriz de correlação entre acessos ao resumo e citações na Scopus



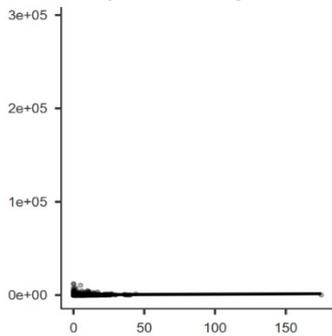
Fonte: Dados da pesquisa (2022).

Figura 35 - Matriz de correlação entre acessos ao resumo e tweets



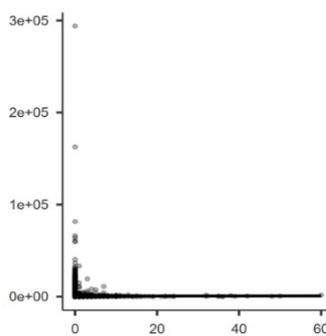
Fonte: Dados da pesquisa (2022).

Figura 30 - Matriz de correlação entre downloads e citações na Scopus



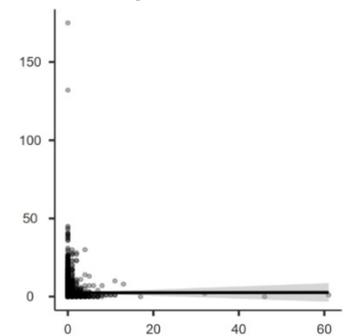
Fonte: Dados da pesquisa (2022).

Figura 33 - Matriz de correlação entre downloads e tweets



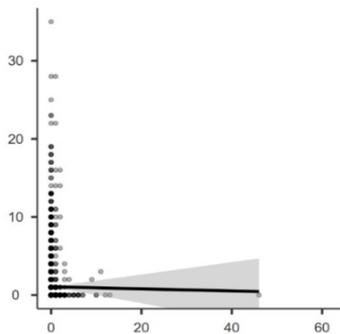
Fonte: Dados da pesquisa (2022).

Figura 36 - Matriz de correlação entre citações na Scopus e tweets



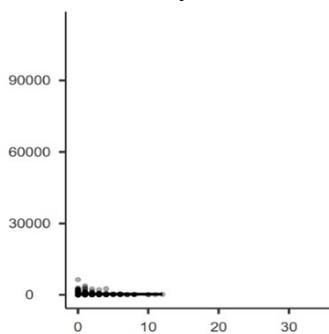
Fonte: Dados da pesquisa (2022).

Figura 37 - Matriz de correlação entre citações na WoS e tweets



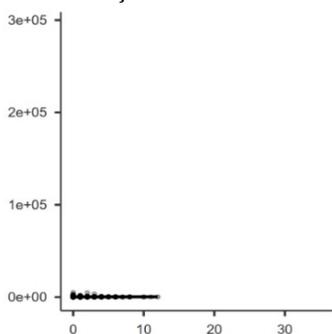
Fonte: Dados da pesquisa (2022).

Figura 40 - Matriz de correlação entre acessos ao resumo e citações na WoS



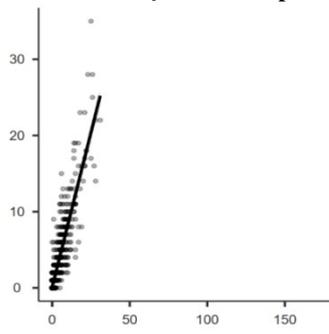
Fonte: Dados da pesquisa (2022).

Figura 38 - Matriz de correlação entre downloads e citações na WoS



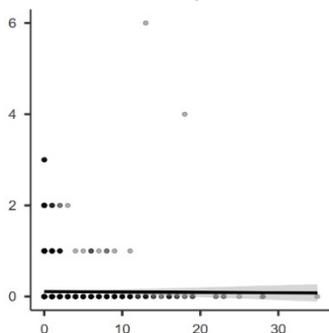
Fonte: Dados da pesquisa (2022).

Figura 41 - Matriz de correlação entre citações na WoS e citações na Scopus



Fonte: Dados da pesquisa (2022).

Figura 39 - Matriz de correlação entre menções no Facebook e citações na WoS



Fonte: Dados da pesquisa (2022).

4.2 COBERTURAS E CORRELAÇÕES DE INDICADORES ALTMÉTRICOS E BIBLIOMÉTRICOS PARA OS ARTIGOS DOS PORTAIS POR ÁREA DE CONHECIMENTO

Nesta seção são analisadas as coberturas e correlações dos indicadores alométricos e bibliométricos de acordo com a área de conhecimento dos artigos, conforme áreas pré-definidas anteriormente nos procedimentos metodológicos. As análises desta seção visam atender aos objetivos específicos de verificar a cobertura das fontes de dados alométricos e bibliométricos de artigos dos portais nas diferentes áreas do conhecimento; e investigar as correlações dos indicadores alométricos e bibliométricos dos artigos dos portais nas diferentes áreas do conhecimento.

4.2.1 Coberturas de indicadores alométricos e bibliométricos para artigos dos portais por área de conhecimento

Em busca de investigar como as práticas de comunicação científica e uso de plataformas de mídias sociais influenciam os indicadores alométricos e de citação, é necessário analisar os artigos de acordo com a área de conhecimento a qual pertencem. Conforme especificadas anteriormente, as áreas de conhecimento nas quais os artigos foram categorizados são: Ciências da Vida (CV); Ciências Humanas (CH); Ciências Sociais Aplicadas (CSA); Linguística, Letras e Artes (LLA); e Ciências Exatas, Tecnológicas e Multidisciplinares (CETM).

Dos 51.200 artigos do universo, 10.832 são pertencentes a CV, o que equivale a 21,15% do total de artigos da amostra. A área de CH contempla 19.001 artigos, sendo 37,11% do total; 8.316 artigos são de CSA, equivalente a 16,24% do total. A área de LLA possui 9.183 artigos, o que representa 17,94% do total; e a área de CETM abrange 3.868 artigos, equivalente a 7,56% do total do universo. O número de artigos (n) por área de conhecimento podem ser vistos na Tabela 5.

Comparativamente com as demais áreas, a CH possui a maior cobertura de artigos, seguida pela CV, LLA, CSA e CETM. Numa análise simplista que considerasse apenas o total de menções e cobertura nas fontes de dados por área, poderia se esperar que as menções fossem proporcionais ao número de artigos, onde a área de conhecimento com mais artigos possuiria o maior número de menções nas fontes de dados alométricas. Entretanto, devido a diversos fatores, como problemas na coleta de relatórios de acessos ao resumo e downloads para algumas revistas; a influência das práticas de comunicação em cada área de

conhecimento; e a preferência por determinadas plataformas de mídias sociais de acordo com o perfil dos usuários, o número de menções varia nas áreas de conhecimento de modo a não ser diretamente proporcional ao número de artigos por área.

A área de conhecimento com maior quantitativo de menções em fontes de dados altmétricas é a CH, com 15.374.177 menções. Destas, a maior parte são downloads e acessos ao resumo, com 81,83% dos artigos de CH com ao menos um acesso ao resumo e 80,43% com ao menos um download efetuado. A área de CETM possui as maiores coberturas de acessos ao resumo e downloads, com 85,08% e 83,89%, respectivamente. Estes percentuais de cobertura podem indicar que os indicadores de acessos ao resumo e downloads são complementares, e que o processo do usuário acessar a página de resumo e decidir fazer o download do artigo é concluído para a maioria dos artigos cobertos. Isso evidencia que artigos de CETM e CH são mais acessados e baixados do que artigos das demais áreas de conhecimento, indicando uma possível diferença nos padrões de uso (acessos e downloads) dos artigos para usuários de diferentes áreas, conclusão esta semelhante a apresentada por outros autores (SCHLÖGL *et al.*, 2014).

A CETM, apesar de reunir a menor parcela de artigos por área de conhecimento, possui a maior cobertura total de menções entre as áreas, com 93,67% dos artigos com ao menos uma menção em alguma das fontes de dados altmétricas. Já para os 10.832 artigos de Ciências da Vida, 10.081 artigos possuem pelo menos uma menção em alguma das bases de dados altmétricas, o que equivale a 93,07% do total de artigos de CV. As coberturas e menções por base de dados altmétricos podem ser observadas na Tabela 5.

Tabela 5 - Indicadores alométricos de menções, coberturas, média e mediana de menções por artigos por área de conhecimento dos portais de periódicos

Fonte de Dados	Total de Menções	Artigos cobertos (com menções)	% de Artigos cobertos	Média (Menções/artigo)	Mediana (Menções/artigo)
Ciências da Vida (CV) n = 10832					
<i>Mendeley</i>	106.530	8.922	82,37%	9,83	5
<i>Twitter</i>	458	300	2,77%	0,04	0
Acesso Resumo	2.457.822	6.131	56,60%	400,82	209
Downloads	3.393.653	6.095	56,27%	554,16	243
<i>Facebook</i>	569	547	5,05%	0,05	0
Total**	5.959.032	10.081	93,07%	-	-
Ciências Humanas (CH) n = 19001					
<i>Mendeley</i>	56.197	9.122	48,01%	2,96	0
<i>Twitter</i>	2.174	729	3,84%	0,11	0
Acesso Resumo	7.073.795	15.549	81,83%	454,91	205
Downloads	8.241.713	15.282	80,43%	530,01	214
<i>Facebook</i>	298	285	1,50%	0,02	0
Total**	15.374.177	17.540	92,31%	-	-
Ciências Sociais Aplicadas (CSA) n = 8316					
<i>Mendeley</i>	43.543	5.150	61,93%	5,24	2
<i>Twitter</i>	1.192	400	4,81%	0,14	0
Acesso Resumo	2.876.059	5.963	71,71%	482,32	262
Downloads	3.225.760	5.907	71,03%	540,96	255
<i>Facebook</i>	371	300	3,61%	0,04	0
Total**	6.146.925	7.524	90,48%	-	-
Linguística, Letras e Artes (LLA) n = 9183					
<i>Mendeley</i>	14.735	3.117	33,94%	1,60	0
<i>Twitter</i>	382	173	1,88%	0,04	0
Acesso Resumo	2.247.264	6.323	68,86%	355,41	207
Downloads	2.726.989	6.323	68,86%	431,28	210
<i>Facebook</i>	568	432	4,70%	0,06	0
Total**	4.989.938	7.522	81,91%	-	-
Ciências Exatas, Tecnológicas e Multidisciplinares (CETM) n = 3868					
<i>Mendeley</i>	15.096	2.221	57,42%	3,90	1
<i>Twitter</i>	70	47	1,22%	0,02	0
Acesso Resumo	1.977.205	3.291	85,08%	600,79	370
Downloads	2.662.032	3.245	83,89%	808,88	376
<i>Facebook</i>	8	8	0,21%	0,00	0
Total**	4.654.411	3.623	93,67%	-	-

Fonte: Dados da pesquisa.

**O Total é calculado a partir da totalidade de artigos da amostra.

Comparativamente com o universo de artigos, onde a maior cobertura de menções foi de 72,77% para acessos ao resumo, para os artigos de CV a cobertura de acessos ao resumo é menor, de 56,6%. Para a área de CV, o *Mendeley* é a fonte de dados alométricos de maior

cobertura, com 82,37%.Essa diferença entre a fonte de dados altmétricos de maior cobertura do universo para a de maior cobertura da amostra de CV pode ser justificada por dois motivos: devido aos procedimentos de coleta de dados, uma vez que artigos de 6 das 17 revistas de CV não tiveram os indicadores de acessos ao resumo e de downloads coletados (problemas técnicos no momento da coleta); e devido às diferenças entre práticas de comunicação científica nas áreas de conhecimento indicada em alguns estudos (MUELLER, 2005).

Esta última justificativa é reforçada quando comparamos as coberturas por fonte de dados altmétrica entre as duas áreas de maior cobertura de menções totais, CETM e CV. Observa-se que a distribuição das coberturas é diferente para CV e CETM. A fonte de dados altmétrica de maior cobertura para artigos de CV é o *Mendeley*, com 82,37%, enquanto para artigos de CETM são os acessos ao resumo, com 85,08%. A menor cobertura altmétrica para artigos de CV é o *Twitter*, com 2,77%, e para os artigos de CETM é o *Facebook*, com apenas 0,21% de cobertura.

O *Mendeley* se destaca na CV não apenas pela cobertura, mas pela média de 9,83 leitores por artigo ser a maior entre as áreas de conhecimento, e maior do que a média de leitores obtida no universo de artigos (4,61). A cobertura de leitores no *Mendeley* para a área de CV está de acordo com Haustein e Larivière (2014), que encontraram coberturas que variam de 62,8% a 81% para artigos de quatro subáreas de CV, e uma média de 9,6 leitores por artigo.

O fato da área de CV ter a maior cobertura de leitores no *Mendeley* reafirma o resultado obtido por Zahedi e Haustein (2018), cuja amostra de publicações de Ciências da Vida e da Terra teve cobertura de 91,4%, a maior entre as cinco áreas estudadas pelos autores. Zahedi, Costas e Wouters (2014) também identificaram uma maior cobertura para a área de Ciências da Vida e Medicina (73%), assim como Zahedi, Costas e Wouters (2017), com 93,06%. A cobertura de leitores no *Mendeley* para a área de Ciências da Vida pode estar relacionada a fatores como o tamanho do artigo (ZAHEDI; HAUSTEIN, 2018), ao comportamento dos leitores e às práticas por área de conhecimento (ZAHEDI; COSTAS; WOUTERS, 2014; ALPERIN, 2015).

As áreas de CH, CSA e CETM tiveram coberturas intermediárias de leitores no *Mendeley*, de 48,01%, 61,93% e 57,42%, respectivamente, similar às coberturas das áreas correspondentes analisadas por Zahedi, Costas e Wouters (2014), variando de 51% a 68%. Já Hammarfelt (2014) obteve uma cobertura de 61% de leitores no *Mendeley* para artigos da base SwePub da área de Humanidades, que abrange disciplinas correspondentes as áreas de CH e LLA deste estudo.

A baixa cobertura de 33,94% de leitores no *Mendeley* para a área de LLA também é próxima à cobertura de 39% encontrada por Zahedi, Costas e Wouters (2014). Uma possível explicação para esta área ter a cobertura mais baixa é que o *Mendeley* ainda não é amplamente utilizado por leitores de conteúdo científico desta área, comportamento que ilustra as diferenças comportamentais de usuários na prática de consumir conteúdo científico entre as diferentes áreas de conhecimento.

Quanto ao *Twitter*, os artigos de CV possuem apenas 2,77% de cobertura. Este resultado está abaixo do encontrado por Haustein, Costas e Larivière (2015) para a área correspondente de Ciências da Vida e da Terra, com cobertura no *Twitter* de 21,6%. A área com a maior cobertura no *Twitter* é CSA, com 4,81%, e a de menor cobertura é a CETM, com 1,22% dos artigos mencionados no *Twitter*. Essas coberturas são inferiores a todas as coberturas encontradas por Haustein, Costas e Larivière (2015) e Costas, Zahedi e Wouters (2014) ainda que, comparativamente, as áreas relacionadas às Ciências Exatas e Engenharias tenham as menores coberturas de *tweets* entre as áreas.

As coberturas no *Facebook* variam de 0,21% a 5,05%, sendo a maior cobertura para a área de CV, e a menor para a CETM. Haustein, Costas e Larivière (2015) também encontraram as menores coberturas no *Facebook* para as áreas de Matemática e Ciência da Computação (1,5%) e Ciências Naturais e Engenharias (2,3%), e as maiores coberturas para as Ciências da Vida e da Terra (5,7%) e Ciências da Saúde e Biomédicas (7,5%). Alperin (2015) também encontrou coberturas de *tweets* (1%) e menções no *Facebook* (2,8%) para artigos brasileiros indexados na *SciELO* menores que outros estudos, o que, de acordo com o autor, pode ser explicado por três possibilidades: um baixo uso dos artigos em geral, refletido no uso das mídias sociais; o uso dessas mídias sociais no Brasil ser baixo entre acadêmicos; ou, há uma cultura diferente no compartilhamento de pesquisas nas mídias sociais no Brasil (ALPERIN, 2015).

As baixas coberturas de menções nas mídias sociais *Twitter* e *Facebook* para artigos dos portais da área de conhecimento de Ciências Exatas, Tecnológicas e da Terra indica que a prática de compartilhar ou debater conteúdo científico desta área nestas mídias não é um costume da comunidade.

Brigham (2014) afirma que algumas disciplinas são mais ativas do que outras online, e o engajamento também pode variar de acordo com a plataforma de mídia social utilizada. Por isso a análise dos indicadores altmétricos, especialmente quando comparados com os indicadores de citação, deve levar em consideração a área de conhecimento das publicações (BRIGHAM, 2014).

Na Tabela 6 são apresentadas as coberturas de citação e indexação, assim como médias e medianas de citações para os artigos divididos por área de conhecimento. Quanto às citações, a diferença nas coberturas entre as áreas do conhecimento é acentuada ao compararmos a CV com as demais áreas. Enquanto 49,09% dos artigos de CV possuem pelo menos uma citação nas três fontes de dados bibliométricas, as demais áreas possuem uma cobertura menor, que varia de 9,10% a 14,77%. Comparativamente ao universo de artigos, que apresentou cobertura de 19,82%, a área de CV se destaca com o maior número de artigos cobertos nas bases bibliométricas, e com o maior volume total de citações, com 34.853 citações recebidas.

O indicador de citação de maior cobertura para todas as áreas é o de citações na *Crossref*, seguido pela *Scopus* e por fim, pela *Web of Science*. Na área de CV, que possui as maiores coberturas em todas as fontes de dados bibliométricas, 43,84% dos artigos possuem alguma citação na *Crossref*, 26,07% na *Scopus*, e 3,98% na *Web of Science*. Essa cobertura de citações na *Crossref* para artigos de CV concorda com a cobertura de 35,9% obtida por Demachki e Maricato (2022b), ao analisarem artigos de revistas do Portal de Periódicos da UFG. Para os autores, “a cobertura de citações significativamente maior da área de Ciências da Vida em comparação com as demais áreas [...] pode estar relacionada ao volume de pesquisadores e de produção científica” (DEMACHKI; MARICATO, 2022b).

Na *Web of Science*, Haustein, Costas e Larivière (2015) identificaram uma cobertura de 74,3% citação para a área de Ciências da Vida e da Terra, 54,6% para Matemática e Ciência da Computação, 73,7% para Ciências Naturais e Engenharias, e 45,8% para Ciências Sociais e Humanidades. A menor cobertura de citação na *Web of Science* identificada pelos autores foi de 45,8% para as Ciências Sociais e Humanidades, muito maior que a encontrada neste estudo para as áreas correspondentes de CH e CSA, com 0,39% e 1,08% de cobertura na *Web of Science*, respectivamente.

As coberturas de indexação das publicações na *Web of Science* também são baixas, variando de 0% para CETM a 10,65% para LLA. É interessante observar que, ainda que a área de CV tenha a maior cobertura de citação na *Web of Science*, ela não é a área de maior cobertura de indexação nesta fonte (a de maior cobertura é a LLA). Existem duas possibilidades para essa diferença: pode haver um número maior de revistas de LLA indexadas na *Web of Science* do que de CV; e, independentemente do número de revistas de cada área que estão indexadas na *Web of Science*, as revistas indexadas de LLA possuem um número maior de artigos publicados do que as demais da mesma área. Entretanto, a segunda justificativa é menos provável, uma vez que revistas de uma mesma área de conhecimento

seguem padrões de publicação semelhantes, incluindo o número de publicações por edição e ano (SCIELO, 2014, p. 15).

Ainda que a cobertura de indexação para a área de CV não seja a maior entre as áreas, a sua média de citações recebidas para artigos indexados na *Web of Science* é a maior, sendo a média de 3,82 citações por artigo, e uma mediana de 2 citações. Entre os artigos indexados, a maior cobertura de citação dentre todas as fontes de dados bibliométricas também é para a área de CV, com a maior cobertura para a *Web of Science*, com 74,83% dos artigos indexados com alguma citação recebida. A segunda maior cobertura de citação entre artigos indexados de CV é na *Scopus*, com 73,58%, e por fim na *Crossref*, com 43,84%. Esses dados sugerem que, para a área CV, as citações da *Crossref* não são tão promissoras. Isso pode estar relacionado ao nível de internacionalização da área de CV e ao número de revistas da área indexadas em grandes bases de dados comerciais.

A média de citações recebidas na *Scopus* para artigos de CV é 3,95, ligeiramente maior do que a média para a *Web of Science*, apesar da cobertura de citação para artigos indexados na *Web of Science* ser um pouco maior do que na *Scopus*, o que indica que estas duas bases estão captando algumas citações diferentes. De fato, Visser, Van Eck e Waltman (2021) identificaram que a sobreposição de artigos entre a *Web of Science* e a *Scopus* foi a menor entre as fontes analisadas (*Web of Science*, *Crossref*, *Dimensions* e *Microsoft Academic*).

A *Crossref* possui uma cobertura de indexação perfeita para toda a amostra de artigos estudados, superando a *Scopus* e *Web of Science*. Como relatado anteriormente, essa cobertura se deve ao fato de todos os artigos possuírem número de DOI, o que implica no registro destes artigos na *Crossref*. Isso nos leva a questionar se, caso os artigos não possuíssem DOI, a cobertura de indexação na *Crossref* seria inferior à das demais fontes de dados bibliométricos. De fato, ainda que os artigos indexados na *Crossref* se resumissem somente aos artigos com alguma citação, a cobertura de indexação na *Crossref* ainda seria a maior entre as três bases de dados bibliométricas para quatro das cinco áreas de conhecimento, exceto para a área de LLA.

Nesse sentido, em um comparativo entre produções científicas (artigos e resenhas) de 20 países nas fontes de dados bibliométricas *Scopus*, *Web of Science* e *Dimensions*, Singh *et al.* (2021) identificaram que para 90% dos países, a fonte que indexa um maior volume de publicações é a *Dimensions*, com o maior volume total de documentos pertencentes aos Estados Unidos, China e Reino Unido. Uma vez que a *Dimensions* depende fortemente dos dados da *Crossref*, estas duas fontes cobrem os mesmos dados (VISSER; VAN ECK;

WALTMAN, 2021), podendo-se considerar que a *Crossref* supera, em cobertura de indexação, a *Scopus* e *Web of Science*.

Tabela 6 - Total de citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais por área de conhecimento, de acordo com as fontes de dados bibliométricas

Fonte de Dados	Total de citações	CC	% CC	CI	% CI	% CCAI	MeCAI	MdCAI
Ciências da Vida (CV) n = 10832								
<i>Crossref</i>	17.485	4.749	43,84%	10.832	100,00%	43,84%	1,61	0
<i>Scopus</i>	15.168	2.824	26,07%	3.838	35,43%	73,58%	3,95	2
<i>WoS</i>	2.200	431	3,98%	576	5,32%	74,83%	3,82	2
Total**	34.853	5.317	49,09%	10.832	100,00%	-	-	-
Ciências Humanas (CH) n = 19001								
<i>Crossref</i>	3.830	2.060	10,84%	19.001	100,00%	10,84%	0,20	0
<i>Scopus</i>	726	401	2,11%	1.538	8,09%	26,07%	0,47	0
<i>WoS</i>	126	75	0,39%	472	2,48%	15,89%	0,27	0
Total**	4.682	2.406	12,66%	19.001	100,00%	-	-	-
Ciências Sociais Aplicadas (CSA) n = 8316								
<i>Crossref</i>	2.158	1.053	12,66%	8.316	100,00%	12,66%	0,26	0
<i>Scopus</i>	267	133	1,60%	486	5,84%	27,37%	0,55	0
<i>WoS</i>	140	90	1,08%	520	6,25%	17,31%	0,27	0
Total**	2.565	1.228	14,77%	8.316	100,00%	-	-	-
Linguística, Letras e Artes (LLA) n = 9183								
<i>Crossref</i>	1.075	672	7,32%	9.183	100,00%	7,32%	0,12	0
<i>Scopus</i>	448	202	2,20%	906	9,87%	22,30%	0,49	0
<i>WoS</i>	219	142	1,55%	978	10,65%	14,52%	0,22	0
Total**	1.742	836	9,10%	9.183	100,00%	-	-	-
Ciências Exatas, Tecnológicas e Multidisciplinares (CETM) n = 3868								
<i>Crossref</i>	634	357	9,23%	3.868	100,00%	9,23%	1,78	1
<i>Scopus</i>	6	5	0,13%	19	0,49%	26,32%	0,32	0
<i>WoS</i>	0	0	0,00%	0	0,00%	0,00%	0	0
Total**	640	362	9,36%	3.868	100,00%	-	-	-

Fonte: Dados da pesquisa.

**O Total é calculado a partir da totalidade de artigos da amostra.

CC = Cobertura de citação total; CI = Cobertura de indexação; CCAI = Cobertura de citação entre artigos indexados; MeCAI = Média de citações por artigo indexado; MdCAI = Mediana de citações por artigo indexado.

Ao compararmos as coberturas alométricas e bibliométricas, a área de CV possui as maiores coberturas de leitores no *Mendeley* (82,37%) e citações em todas as fontes (totalizando 49,09%). Entretanto, a menor diferença entre cobertura de leitores no *Mendeley* e citações totais é para a área de LLA (24,84), seguida pela CV (33,28). Já a área de CETM possui a maior diferença (48,06) entre as coberturas no *Mendeley* (57,42%) e de citações

totais (9,36%). Áreas com mais leitores ou com equilíbrio proporcional entre leitores e citações podem se beneficiar do uso do *Mendeley*, por terem mais impacto de leitura do que impacto de citação (ZAHEDI; COSTAS; WOUTERS, 2014).

Tendo em vista as diferenças entre as coberturas nas áreas do conhecimento, observou-se que a área de CV possui as maiores coberturas de leitores no *Mendeley*, menções no *Facebook*, e citações nas três fontes bibliométricas. Enquanto isso, a área de CETM possui as menores coberturas de indicadores altmétricos do *Twitter* e *Facebook*, e de indicadores de citações na *Web of Science* e *Scopus*, em contraste com a maior cobertura de acessos ao resumo e downloads.

4.2.2 Correlações entre indicadores altmétricos e bibliométricos para artigos dos portais por área de conhecimento

Ao analisarmos as correlações de Spearman entre os indicadores por área de conhecimento que estão dispostas no Quadro 11, podemos observar que a correlação mais forte entre leitores no *Mendeley* e citações nas três fontes bibliométricas se dá na área de CV, com o coeficiente de Spearman variando entre 0,432, para citações na *Scopus*, e 0,491 para citações na *Web of Science*. Haustein *et al.* (2014) também obtiveram correlações moderadas entre leitores no *Mendeley* e citações na *Web of Science* de 0,530, 0,448 e 0,439 para áreas de Pesquisa Biomédica, Biologia, e Medicina, respectivamente. Da mesma forma, Thelwall e Sud (2016) identificaram correlações moderadas (entre 0,54 a 0,67) para citações na *Scopus* para diversas subáreas de Ciências da Vida.

As correlações entre leitores no *Mendeley* e citações na *Web of Science* são um pouco mais fracas para as áreas de LLA (0,250) e CH (0,251) do que para CSA (0,285), assim como ocorreu em estudos anteriores, para estas mesmas fontes de dados (MOHAMMADI; THELWALL, 2014, HAUSTEIN *et al.*, 2014, THELWALL, 2019). Uma vez que a área de CETM não possui nenhuma das revistas dos portais indexada na *Web of Science*, a correlação para esta fonte não foi calculada.

Mohammadi e Thelwall (2014) observaram uma correlação mais fraca para as Humanidades (0,428), comparativamente com as Ciências Sociais (0,516) para publicações indexadas na *Web of Science* em 2008. Já Haustein *et al.* (2014) obtiveram correlações de 0,431 para Ciências Sociais, 0,227 para Humanidades e -0,209 para Artes, com diferença na força das correlações ainda maiores do que as observadas no presente estudo. Thelwall (2019)

identificou correlações variando de fraca a moderada para Literatura (0,382) e Artes Visuais (0,477), e de moderada a forte para História (0,622) e Filosofia da Ciência (0,779).

Considerando que a *Crossref* possui as maiores coberturas de citação para diversas áreas de conhecimento, a análise das correlações entre o número de citações provindo desta base com os leitores do *Mendeley* pode trazer indicadores mais robustos, melhorando os resultados das análises destas correlações por área. Nesse sentido, a única correlação de força moderada entre leitores no *Mendeley* e citações na *Crossref* é na área de CV (0,445), com as demais correlações variando entre 0,254 e 0,316, o que indica que o *Mendeley* tem maior capacidade de prever citações para artigos da área de CV.

Assim como a correlação entre leitores no *Mendeley* e citações na *Web of Science*, a correlação entre leitores no *Mendeley* e citações na *Crossref* para a área de LLA é a mais fraca entre as áreas, com $\rho = 0,254$, o que indica que, para artigos de LLA, as leituras/capturas no *Mendeley* convergem menos em citações futuras do que para áreas como a CV. Já entre leitores no *Mendeley* e citações na *Scopus*, a correlação se mantém fraca (0,391), mas já é superior à de outras áreas, como CH (0,140) e CSA (0,006).

Áreas com correlações mais fracas entre leitores e citações podem indicar que os leitores não são necessariamente, pesquisadores, por isso podem não ter o interesse em citar o artigo lido. Nesse sentido, Zahedi e Van Eck (2018) identificaram que artigos de áreas com menor densidade de citação na *Web of Science* possuem mais leitores no *Mendeley*, como foi o caso de publicações de Ciências Sociais e Humanidades. Ao analisarem os perfis dos usuários no *Mendeley*, os autores observaram que publicações das Ciências Sociais e Humanidades foram mais lidas por bibliotecários, estudantes e professores. Os autores também identificaram que os pesquisadores estão mais concentrados em leituras de artigos da área de Ciências da Saúde e Biomédicas. De acordo com Zahedi, Costas e Wouters (2017), os artigos podem ser lidos para serem usados no ensino, por curiosidade ou por interesses individuais e não acadêmicos. Essas diferenças de uso dos artigos no *Mendeley* podem ajudar a explicar diferenças entre as áreas de conhecimento (ZAHEDI; VAN ECK, 2018).

É interessante observar que, apesar da correlação mais fraca entre os leitores no *Mendeley* e citações na *Scopus* ser para a área de CSA, com $\rho = 0,006$, a CETM apresenta uma correlação negativa fraca de $-0,304$ entre estes indicadores, o que indica que o aumento no número de leitores no *Mendeley* possui uma influência contrária no número de citações. Em outras palavras, ainda que a correlação seja fraca, artigos de CETM que possuem mais leitores no *Mendeley* tendem a ser menos citados na *Scopus*, o que pode estar relacionado ao tempo de publicação dos artigos ser recente, tendendo a ter substancialmente mais leitores do

que citações (THELWALL; SUD, 2016, ZAHEDI; COSTAS; WOUTERS, 2017), e a prática de pesquisadores de subáreas de CETM, assim como o tato das Engenharias publicarem preferencialmente em anais de congressos, ao invés de artigos em revistas científicas (MUELLER, 2005).

A correlação entre leitores no *Mendeley* e acessos ao resumo é moderada para as áreas de CSA (0,527) e CETM (0,514), e nula para a LLA (0,190), enquanto nas demais áreas a correlação é fraca. Quanto aos downloads, a correlação com os leitores no *Mendeley* é moderada para CSA (0,489) e CETM (0,581), sendo mais fortes que a correlação entre leitores e citações em todas as fontes bibliométricas para estas áreas. Comparativamente, o estudo de Schlogl *et al.* (2014) também identificou correlações maiores (0,73 e 0,66) entre downloads no ScienceDirect e leitores no *Mendeley* para artigos de duas revistas da área de CSA, e correlações menores entre leitores e citações na *Scopus* (0,51 e 0,59).

As correlações moderadas entre leitores no *Mendeley* e acessos ao resumo para as áreas de CSA e CETM são maiores do que a correlação para o universo (0,325), assim como entre os leitores do *Mendeley* e downloads, com $\rho = 0,347$ para o universo de artigos. Esse aumento na força destas correlações quando comparadas com o universo pode indicar uma relação causal de que mais acessos a página de resumo e mais downloads são feitos por usuários que salvaram artigos de CSA e CETM em sua biblioteca no *Mendeley*. Essas correlações também podem ser hipoteticamente relacionadas devido aos leitores darem maior atenção a artigos práticos e de alto nível acadêmico (LIU *et al.*, 2013), que geralmente são baixados com mais frequência.

Podemos observar que, para as áreas de CH, LLA e CETM, as correlações entre downloads e leitores no *Mendeley* são ligeiramente mais fortes do que entre acessos ao resumo e leitores no *Mendeley*, assim como ocorre com estas correlações da totalidade dos dados. Isso indica que parte dos downloads de artigos dessas áreas, que são provenientes de usuários que salvaram este mesmo artigo em sua biblioteca de referências no *Mendeley* (ação aqui considerada como leitura na plataforma), não tiveram a página de resumo previamente acessada por este usuário. Este fluxo de navegação aponta que a página do arquivo do artigo está sendo a página de destino para cliques provindos do *Mendeley*.

Tabela 7 - Correlações de Spearman entre indicadores altmétricos e bibliométricos para os artigos dos portais por área de conhecimento.

Ciências da Vida								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Crossref</i>	—							
<i>Mendeley</i>	0.445	—						
<i>Twitter</i>	0.098	0.110	—					
<i>WoS</i>	0.871	0.491	0.197	—				
<i>Scopus</i>	0.767	0.432	0.070	0.901	—			
Acesso Resumo	0.085	0.343	0.050	0.050	0.163	—		
Downloads	0.016	0.328	0.022	0.449	0.185	0.825	—	
<i>Facebook</i>	0.117	0.159	0.313	0.148	0.022	0.032	0.028	—
Ciências Humanas								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Crossref</i>	—							
<i>Mendeley</i>	0.316	—						
<i>Twitter</i>	0.048	0.065	—					
<i>WoS</i>	0.403	0.251	-0.053	—				
<i>Scopus</i>	0.237	0.140	0.097	0.769	—			
Acesso Resumo	0.166	0.296	0.077	0.084	0.125	—		
Downloads	0.160	0.324	0.002	0.177	0.092	0.795	—	
<i>Facebook</i>	0.019	0.063	0.073	NaN	0.016	0.114	0.088	—
Ciências Sociais Aplicadas								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Crossref</i>	—							
<i>Mendeley</i>	0.377	—						
<i>Twitter</i>	0.079	0.093	—					
<i>WoS</i>	0.231	0.285	0.044	—				
<i>Scopus</i>	0.033	0.006	0.095	0.330	—			
Acesso Resumo	0.199	0.527	0.014	0.213	-0.064	—		
Downloads	0.172	0.489	-0.020	0.251	-0.036	0.828	—	
<i>Facebook</i>	0.162	0.137	0.411	0.055	0.076	-0.151	-0.149	—
Linguística, Letras e Artes								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Crossref</i>	—							
<i>Mendeley</i>	0.254	—						
<i>Twitter</i>	0.073	0.067	—					
<i>WoS</i>	0.409	0.250	-0.001	—				
<i>Scopus</i>	0.571	0.391	-0.002	0.671	—			
Acesso Resumo	0.152	0.190	0.084	0.185	0.077	—		
Downloads	0.141	0.193	0.029	0.223	0.134	0.784	—	
<i>Facebook</i>	0.024	0.089	0.219	-0.006	0.036	0.103	0.015	—
Ciências Exatas, Tecnológicas e Multidisciplinares								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	<i>Facebook</i>
<i>Crossref</i>	—							
<i>Mendeley</i>	0.255	—						
<i>Twitter</i>	0.038	0.036	—					
<i>WoS</i>	NaN	NaN	NaN	—				
<i>Scopus</i>	NaN	-0.304	NaN	NaN	—			
Acesso Resumo	0.203	0.514	0.085	NaN	-0.414	—		
Downloads	0.218	0.581	0.046	NaN	-0.414	0.793	—	
<i>Facebook</i>	0.047	0.009	0.150	NaN	NaN	0.036	0.040	—

Fonte: Dados da pesquisa, calculados com o software Jamovi (2021).

Na análise entre indicadores de citação, a correlação entre citações na *Web of Science* e na *Scopus* é quase perfeita (0,901) para a área de CV. Enquanto isso, para a área de CSA, esta correlação é fraca, com $\rho = 0,330$. Isso indica que, para a área de CV, provavelmente haja um número maior de citações sobrepostas entre as fontes *Scopus* e *Web of Science*, com o oposto ocorrendo para artigos da área de CSA, onde a maioria das citações para um mesmo artigo são diferentes (VISSER; VAN ECK; WALTMAN, 2021). Isso reforça a ideia de que as citações da Crossref têm potencial de ser utilizada em áreas menos internacionalizadas e sub-representadas nas bases de dados comerciais internacionais, a exemplo da CSA (em oposição à área de CV).

As correlações entre os acessos ao resumo e os downloads com as citações na *Scopus* são fracas para quatro das cinco áreas de conhecimento, variando de -0,064 a 0,185. Entretanto, para a área de CETM, essas correlações são moderadas e negativas, ambas de -0,414. Isso nos leva a investigar a causa destas correlações que diferem de resultados anteriores, que apontam para uma relação positiva entre indicadores de acesso e citações (MOED, 2005, LIU *et al.*, 2013, WANG *et al.*, 2014, CHEN *et al.*, 2020). Ao observarmos a cobertura de citações para a CETM, vemos que apenas 19 dos 3868 artigos da área estão indexados na fonte, com apenas 5 deles tendo alguma citação. Ainda que a amostra de artigos para a CETM não seja suficientemente ampla para se tirar conclusões aplicáveis para todas as áreas do conhecimento, as análises de cobertura e correlações foram executadas para os artigos desta área. Entretanto, a pequena amostra de artigos citados influencia na força da correlação entre os indicadores.

Assim como para o universo de artigos, a correlação entre os indicadores de acesso ao resumo e citações da *Crossref* é quase nula para todas as áreas de conhecimento, variando de 0,085 a 0,203. Entre os downloads e citações da *Crossref*, as correlações são quase nulas para quatro das cinco áreas de conhecimento, variando de 0,016 a 0,172, e fraca para a área de CETM, sendo 0,218.

De modo similar às citações da *Crossref*, as citações na *Web of Science* apresentam correlações nulas e fracas tanto com os acessos ao resumo quanto com os downloads para as áreas CH, CSA e LLA. Como a área CETM não possui nenhuma citação na *Web of Science*, não há correlação entre os indicadores. Para a CV, apesar da correlação entre acessos ao resumo e citações na *Web of Science* ser próxima a nula, de 0,050, a correlação entre downloads e citações na *Web of Science* é moderada, de 0,449.

A correlação moderada de 0,411 entre os indicadores altmétricos de posts no Facebook e no *Twitter* para a área de CSA possui força superior à universal (0,224) e das demais áreas

do conhecimento, que possuem correlações fracas. Isso indica que as postagens nessas mídias sociais possuem uma ligação maior, provavelmente relacionada ao público que usa as plataformas.

Entre as postagens no *Facebook* e as citações na *Crossref*, a área de CSA também tem a correlação mais forte entre todas as áreas, porém, ainda é considerada uma correlação próxima a nula, de 0,162. Em todas as análises, tanto por área quanto no universo de artigos, fica evidente que as postagens no Facebook não exercem influência no número de citações que os artigos recebem. As correlações entre o *Twitter* e as citações na *Crossref* são próximas às correlações entre o *Facebook* e estas citações, sendo praticamente nulas para todas as áreas. Estes resultados estão em concordância com os de Chen *et al.* (2020), que encontraram uma correlação quase nula de 0,004 entre o indicador do *Altmetric* (composto em sua maioria por menções no *Twitter* e *Facebook*) e citações.

4.3 COBERTURAS E CORRELAÇÕES DE INDICADORES ALTMÉTRICOS E BIBLIOMÉTRICOS PARA OS ARTIGOS DOS PORTAIS DE ACORDO COM A DATA DE PUBLICAÇÃO

4.3.1 Coberturas de indicadores altmétricos e bibliométricos para artigos dos portais de acordo com a data de publicação

Na análise de cobertura por fonte de dados bibliométricas, feita na seção 4.1.1 deste estudo, foram considerados artigos de todas as revistas dos portais de periódicos de acesso aberto participantes do universo. Entretanto, para a análise das coberturas de acordo com a data de publicação, foram considerados exclusivamente artigos de revistas que estão indexadas em alguma das seguintes fontes de dados bibliométricas: *Scopus* e *Web of Science*. Foram coletados indicadores de acessos ao resumo e downloads para 16.439 artigos, sendo: 14.739 artigos de revistas indexadas na *Scopus*, e 8.280 artigos de revistas indexadas na *Web of Science*. Entende-se por revista indexada em uma fonte de dados toda revista que possui ao menos um artigo indexado na fonte. As informações de citações e coberturas de citação e indexação para estes artigos podem ser vistas na Tabela 8.

Considerando que a presença de dados altmétricos para publicações antigas é insignificante (COSTAS; ZAHEDI; WOUTERS, 2015), os artigos analisados foram divididos de acordo com o ano de publicação, de 2017 a 2021, com artigos publicados antes de 2017 agrupados na mesma faixa temporal (>2017).

O total de citações na *Scopus* é de 4.906, para 1.308 artigos, equivalente a uma cobertura de 8,87% do total de artigos com indicadores de acessos ao resumo e downloads publicados em revistas indexadas na *Scopus*. A cobertura de indexação nesta fonte de dados é de 18,12%, ou seja, quase metade (48,95%) dos artigos indexados na *Scopus* possuem alguma citação nesta mesma fonte.

O total de citações na *Web of Science* é menor do que na *Scopus*, com apenas 467 citações para 225 artigos, uma cobertura de 2,72% do total. O percentual de artigos indexados na *Web of Science* é 10,66%, equivalente a 883 artigos. É interessante observar que, apesar de que 8.280 artigos com indicadores de acessos ao resumo e downloads tenham sido coletados exclusivamente de revistas indexadas na *Web of Science*, a indexação de uma revista não garante uma alta cobertura de indexação de seus artigos (DEMACHKI; MARICATO, 2022a), assim como ocorre para revistas indexadas na *Scopus*. De fato, Demachki e Maricato (2022a) identificaram, para o Portal de Periódicos da UFG, que as bases *Scopus* e *Web of Science* cobrem poucas revistas (16%) e, destas nenhuma possui uma cobertura integral de seus artigos.

Tabela 8 - Total de citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais com algum indicador altmétrico de acessos ao resumo ou downloads, de acordo com as fontes de dados bibliométricas nas quais as revistas estão indexadas

Fonte de dados	Total de artigos	Total de citações	CC	%CC	CI	% CI	%CCAI	MeCAI	MdCAI
<i>Scopus</i>	14.739	4.906	1.308	8,87%	2.672	18,12%	48,95%	1,83	0
<i>WoS</i>	8.280	467	225	2,72%	883	10,66%	25,48%	0,51	0

Fonte: Dados da pesquisa.

CC = Cobertura de citação total; CI = Cobertura de indexação; CCAI = Cobertura de citação entre artigos indexados; MeCAI = Média de citações por artigo indexado; MdCAI = Mediana de citações por artigo indexado.

Ainda que a *Crossref* não tenha sido considerada na análise temporal dos indicadores bibliométricos, foram consideradas as duas fontes que são utilizadas na maioria dos estudos bibliométricos, *Web of Science* e *Scopus* (CÉSPEDES, 2021), ainda que estas não sejam as de maior cobertura de indexação e citação para artigos brasileiros. A Tabela 9 possui informações de citações e coberturas de citação para os artigos indexados nas fontes de dados bibliométricas, de acordo com o ano de publicação do artigo. Ao analisarmos estas coberturas, é possível observar que a maioria dos artigos (1213) indexados na *Scopus* foram publicados antes de 2017. A cobertura de citação para estes artigos foi de 69,41%, com uma média de 3,14 citações por artigo, sendo a maior cobertura e média entre as faixas temporais analisadas.

Ao comparar as coberturas de citação de artigos na *Scopus* ano a ano, nota-se que a cobertura diminui à medida que o ano de publicação do artigo é mais recente, chegando à menor cobertura de citação no ano de 2021, com apenas 8,72% dos artigos publicados neste ano com alguma citação. Esse resultado reforça o conhecimento de que citações demoram mais tempo para ocorrer do que menções (THELWALL *et al.*, 2013), e que indicadores de citação tendem a ser menores para anos mais recentes e próximos à data de publicação do artigo (WANG *et al.*, 2014).

Quanto à *Web of Science*, não há informações de citações nesta fonte de dados para artigos publicados antes de 2017, indicando que somente artigos mais recentes possuem indexação nesta fonte. Artigos publicados em 2018 possuem a maior cobertura de citação, com 88 de 224 artigos com alguma citação na *Web of Science*, equivalente a 39,28% do total para o ano. A média de citação na *Web of Science* para artigos publicados em 2018 também é a maior entre os anos de publicação, sendo 1,05 citação por artigo.

Exceto para o ano de 2021, todas as coberturas de citação na *Web of Science* são inferiores às coberturas de citação na *Scopus*, o que reafirma a sub representatividade de artigos brasileiros em bases de dados internacionais como a *Web of Science* (MELO; TRINCA; MARICATO, 2021; ALPERIN, 2014; AVENA & BARBOSA, 2017). Comparativamente com estudos anteriores, como o de McGillivray e Astell (2019), as médias e medianas de citação também são inferiores às encontradas pelos autores, de 8,9 e 6, respectivamente, para os dois anos após a publicação de artigos da revista *Scientific Reports*.

A cobertura de 24,17% de artigos publicados em 2021 com alguma citação na *Web of Science* pode indicar uma maior imediatez de citação para artigos indexados nesta fonte de dados, possivelmente relacionada à visibilidade que artigos ganham ao serem indexados na *Web of Science*.

Tabela 9 - Total de citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais por ano de publicação, de acordo com as fontes de dados bibliométricas nas quais estão indexados

Fonte de Dados	Total de artigos indexados	Total de citações	CCAI	% CCAI	MeCAI	MdCAI
<2017						
<i>Scopus</i>	1.213	3.820	842	69,41%	3,14	2
<i>WoS</i>	0	0	0	0,00%	0	0
2017						
<i>Scopus</i>	153	389	103	67,32%	2,54	1
<i>WoS</i>	174	92	42	24,13%	0,52	0
2018						
<i>Scopus</i>	257	338	136	52,92%	1,31	1
<i>WoS</i>	224	236	88	39,28%	1,05	0
2019						
<i>Scopus</i>	317	228	114	35,96%	0,72	0
<i>WoS</i>	133	53	25	18,79%	0,39	0
2020						
<i>Scopus</i>	388	120	83	21,39%	0,31	0
<i>WoS</i>	170	41	29	17,05%	0,24	0
2021						
<i>Scopus</i>	344	43	30	8,72%	0,12	0
<i>WoS</i>	182	75	44	24,17%	0,41	0

Fonte: Dados da pesquisa.

CCAI = Cobertura de citação entre artigos indexados; MeCAI = Média de citações por artigo indexado; MdCAI = Mediana de citações por artigo indexado.

Quanto à cobertura de acessos ao resumo e downloads, todos os artigos coletados possuem indicadores de número de acessos ao resumo ou número de downloads, resultando em uma cobertura de 100% para todos os anos. Dos 16.439 artigos coletados, 16.122 estão presentes nos relatórios de downloads, e 14.608 nos relatórios de acessos ao resumo, indicando que alguns artigos que possuem acessos ao resumo não possuem nenhum download.

O total de menções por ano são maiores para artigos publicados antes de 2017, o que pode estar relacionado ao fato de o número de artigos ser maior para este período. Entretanto, as médias de acessos ao resumo e downloads também são as maiores entre os períodos analisados, com média de 926,65 downloads por artigo publicado antes de 2017. Quanto a artigos publicados recentemente, em 2021, estes possuem as menores médias de acessos ao resumo e downloads. Estes números indicam que, apesar dos indicadores altmétricos serem

considerados relevantes por possibilitarem a mensuração do impacto social de artigos científicos imediatamente após a publicação, estes indicadores seguem sendo relevantes ao longo dos anos.

As médias e medianas de acessos ao resumo, downloads e citações são maiores para artigos mais antigos, publicados antes de 2017, assim como o resultado encontrado por Wang *et al.* (2014), que ao analisar artigos do portal *PLOS* publicados entre 2004 e 2012, encontrou as maiores médias e medianas para o ano de 2004.

Tabela 10 - Total de menções, coberturas, médias e medianas de menções para artigos dos portais com algum indicador altmétrico de acessos ao resumo ou downloads

Fonte de Dados	Total de artigos	Total de menções	Artigos cobertos*	% de artigos cobertos	Média	Mediana
<2017						
Acessos	7.767	4.671.554	7.767	100,00%	601,46	305
Downloads	4.494	4.164.407	4.494	100,00%	926,65	404
2017						
Acessos	3.272	1.723.936	3.272	100,00%	526,87	312
Downloads	1.170	597.856	1.170	100,00%	510,98	343
2018						
Acessos	1.434	369.559	1.434	100,00%	257,71	150
Downloads	1.590	322.252	1.590	100,00%	202,67	67
2019						
Acessos	869	284.681	869	100,00%	327,59	227
Downloads	7.361	2.646.905	7361	100,00%	359,58	161
2020						
Acessos	685	203.684	685	100,00%	297,35	228
Downloads	874	216.766	874	100,00%	248,01	188
2021						
Acessos	581	107.689	581	100,00%	185,67	146
Downloads	633	89.778	633	100,00%	141,82	112
Total						
Acessos	14.608	7.361.103	14.608	100,00%	366,10	227,5
Downloads	16.122	8.037.964	16.122	100,00%	398,28	174,5

Fonte: Dados da pesquisa.

*Artigos com algum acesso ao resumo ou algum download.

Em 2019 o número de artigos com indicador de downloads é o maior entre os anos, com a maior diferença entre o número de artigos com acessos ao resumo (869) e número de downloads (7.361) entre todos os períodos de tempo analisados. Essa diferença entre o número de artigos com downloads e acessos ao resumo pode ter duas justificativas: uma falha técnica nos relatórios gerados pelo sistema *OJS*, onde nem todos os artigos com downloads tiveram a contagem de acessos ao resumo contabilizada; ou, artigos publicados em 2019

tiveram uma atenção social elevada, com muitos downloads do artigo feitos através de links externos ao sistema *OJS*, sem acessarem antes a página de resumo, não sendo contabilizados nos relatórios de acesso ao resumo. Estas duas possíveis justificativas não são passíveis de comprovação com os dados disponíveis, visto que a primeira demandaria uma análise no software *OJS*, e para a segunda seriam necessárias análises qualitativas de menções em diversas fontes de dados.

Assim, percebe-se que as coberturas são, de modo geral, maiores para artigos mais antigos, um indicativo de que a obsolescência dos indicadores altmétricos de acessos ao resumo e downloads não ocorre da mesma forma que para os indicadores altmétricos de mídias sociais, que devem ser coletados para períodos mais próximos da data de publicação (ALPERIN, 2015). As análises de cobertura serão complementadas a seguir com a análise das correlações entre acessos ao resumo, downloads e citações na *Web of Science* e *Scopus*.

4.3.2 Correlações entre indicadores altmétricos e bibliométricos para artigos dos portais de acordo com a data de publicação

Analisando as correlações entre downloads e citações na *Scopus* dispostas na Tabela 11, é possível observar que todas as correlações são de força nula, variando de $-0,185$ para artigos publicados em 2017, a $0,187$, para artigos publicados em 2020. Entre downloads e citações na *Web of Science*, a correlação mais forte foi de $-0,264$ para o ano de 2021. O fato desta correlação ser negativa indica que, quanto mais downloads do artigo são feitos, menos citações este artigo tem.

Quanto às correlações entre acessos ao resumo e citações, os acessos ao resumo de artigos publicados em 2021 se correlacionam fraca e negativamente com as citações na *Web of Science* ($-0,226$). Este resultado indica que há influência do tempo de publicação na correlação entre indicadores altmétricos e bibliométricos de citação. Artigos publicados em 2021 possuem poucas citações devido serem publicações recentes, mas, também possuem muitos downloads e acessos ao resumo, indicadores disponíveis imediatamente após a publicação e que variam de acordo com a atenção social recebida pelo artigo (WANG *et al.*, 2021).

Além disso, artigos tendem a ter mais leitores do que citações nos primeiros anos após a publicação (THELWALL & SUD, 2016), e estas leituras provenientes dos downloads do artigo serão responsáveis por uma parte das citações futuras. Citações possuem um atraso de cerca de dois anos em relação aos downloads (WATSON, 2009), o que nos leva a considerar

que as correlações para artigos publicados antes de 2020 refletem mais adequadamente a real influência entre os indicadores.

Para artigos publicados em 2017, a correlação entre acessos ao resumo e citações na *Web of Science* é fraca e positiva, de 0,259, sendo a correlação mais forte entre os anos para estes indicadores, e confirmando a influência do tempo de publicação do artigo nas correlações entre indicadores altmétricos e de citação.

Tabela 11 - Correlações de Spearman entre indicadores altmétricos de acessos e downloads, e indicadores bibliométricos para os artigos dos portais por ano de publicação

Ano	Fonte	ND	Downloads	NA	Acessos
<2017	<i>Scopus</i>	1213	0,152	1059	0,001
	<i>WoS</i>	0	-	0	-
2017	<i>Scopus</i>	153	-0,185	153	-0,195
	<i>WoS</i>	174	0,171	167	0,259
2018	<i>Scopus</i>	257	-0,157	222	0,057
	<i>WoS</i>	223	0,152	224	-0,007
2019	<i>Scopus</i>	317	-0,028	251	-0,124
	<i>WoS</i>	131	0,134	133	-0,008
2020	<i>Scopus</i>	388	0,187	329	0,061
	<i>WoS</i>	170	0,094	170	-0,062
2021	<i>Scopus</i>	344	-0,073	294	0,147
	<i>WoS</i>	182	-0,264	182	-0,226

Fonte: Dados da pesquisa, calculados com o software Jamovi (2021).

Legenda: ND = número de artigos indexados na fonte de dados e com algum indicador de downloads. NA = número de artigos com algum indicador de acessos ao resumo.

Entretanto, observa-se que as correlações dos indicadores de downloads e de acessos ao resumo com os indicadores bibliométricos de citação na *Web of Science* e *Scopus* não seguem o mesmo padrão de força a cada ano. Em 2017, ano de publicação de maior força nas correlações positivas dos indicadores altmétricos com a *Web of Science*, não é o ano de maior força (0,187, em 2020) para as correlações positivas entre os indicadores altmétricos e as citações na *Scopus*. Isso indica que grande parte das citações nestas duas fontes de dados não se sobrepõem, possivelmente por não serem referentes ao mesmo artigo, visto que nem todas as revistas indexadas na *Scopus* estão indexadas na *Web of Science*.

Diferente das correlações moderadas e fortes, variando de 0,517 a 0,772, encontradas por Wang *et al.* (2014) para todos os anos de publicação dos artigos (2004 a 2012), as correlações entre downloads e citações do presente estudo são, em sua maioria, nulas, assim como entre acessos ao resumo e citações. Os presentes resultados também diferem dos obtidos por McGillivray e Astell (2019), que encontraram correlações moderadas entre downloads e citações (0,49); e visualizações e citações (0,47), para artigos publicados entre 2012 e 2014, com os indicadores altmétricos de downloads e visualizações referentes ao primeiro ano após a publicação, e as citações após dois anos da publicação.

As correlações nulas encontradas neste estudo podem ter sido influenciadas pelo tamanho reduzido da amostra, não refletindo uma correlação real, causado pela baixa compatibilidade obtida no cruzamento entre artigos com algum indicador altmétrico de acessos ao resumo e downloads, com artigos indexados nas fontes de dados bibliométricas. Outra possibilidade é que, mesmo com amostras variando entre 131 e 1.213 artigos a cada ano, as correlações sejam nulas devido a atenção social recebida pelos artigos não estar impactando nas citações recebidas ao longo dos anos. Isso indica que, para artigos de revistas de portais de periódicos de acesso aberto brasileiros indexados nas fontes *Web of Science* e *Scopus*, os acessos ao resumo e downloads não resultam em citações futuras nas fontes *Web of Science* e *Scopus*.

A segunda hipótese é reforçada quando comparamos as correlações por ano de publicação com as correlações entre acessos ao resumo, downloads e citações para o universo de 51.200 artigos. Todas as correlações entre estes indicadores altmétricos e as citações nas fontes *Web of Science* e *Scopus* foram nulas para os artigos do universo.

De acordo com Chen *et al.* (2020), um artigo atinge seu pico de citações três anos após sua publicação, enquanto o pico de downloads é atingido após cerca de cinco meses da publicação, reduzindo drasticamente após 10 meses da publicação. Porém, os downloads continuam frequentes ao longo dos anos para artigos de alto nível acadêmico (XUE-LI; HONG-LING; MEI-YING, 2011), o que mantém o número de downloads e citações consistentes entre si. De acordo com os autores, mais downloads para um artigo em um curto período é um indicador de mais citações em longo prazo. Entretanto, para os artigos dos portais de periódicos de acesso aberto brasileiros, as correlações indicam não haver consistência entre os indicadores de acessos ao resumo e downloads com as citações nas fontes de dados *Web of Science* e *Scopus*.

Desse modo, e, considerando que outras fontes de dados bibliométricas diferentes das internacionalmente consolidadas devem ser consideradas ao analisar citações de artigos

publicados no Brasil, somado aos resultados de correlações analisadas nas seções anteriores, conclui-se que as fontes de dados *Mendeley* e *Crossref* seriam as mais promissoras em estudos de correlações entre indicadores alométricos e bibliométricos de acordo com o ano de publicação dos artigos de portais de periódicos brasileiros.

5 CONCLUSÃO

Foram analisados indicadores para 51.200 artigos oriundos de portais de periódicos de universidades brasileiras, em diferentes fontes de dados bibliométricas e altmétricas. As análises foram conduzidas de modo a satisfazer os objetivos específicos de verificar a cobertura das fontes de dados altmétricos e bibliométricos de artigos de portais de periódicos nas diferentes áreas do conhecimento, bem como investigar as correlações entre estes indicadores, além de analisar o fator temporal nas correlações entre indicadores bibliométricos e altmétricos. A seguir são apresentados os principais resultados e respectivas conclusões da pesquisa, em função dos objetivos propostos.

Dentre as fontes de dados bibliométricos (WoS, *Crossref* e Scopus), a *Crossref* se mostrou a fonte de dados que, em geral, possui maior cobertura de citação e indexação para o universo estudado, confirmando a relevância de que novas fontes de dados sejam estudadas, indo além das bases de dados de maior prestígio internacional. Este resultado abre caminho para que pesquisas futuras explorem indicadores de citação da *Crossref* para artigos de regiões sub-representadas em bases de dados comerciais de renome, mostrando ser uma alternativa, especialmente para áreas do conhecimento com baixa inserção internacional. Embora a *Crossref* apresente um conjunto grande de limitações, acredita-se que esta fonte de dados tem enorme potencial para o monitoramento e avaliação da pesquisa científica em diferentes contextos.

Porém, existem diferenças entre as áreas em relação às coberturas de citação nas diferentes fontes. Na área de Ciências da Vida, por exemplo, os artigos apresentam, correlações fortes entre as fontes de dados *Web of Science*, *Scopus* e *Crossref*, quando comparadas com outras áreas do conhecimento. Isso indica que, muito provavelmente, existe sobreposição das citações. Ou seja, em Ciências da Vida, diferentemente de algumas áreas, as coberturas de citação não aparentam ser complementares. As coberturas encontradas estão em concordância com resultados anteriores e indicam haver diferenças nas práticas acadêmicas entre as áreas do conhecimento.

As coberturas de dados altmétricos encontradas indicam diferenças de uso das mídias sociais de acordo com a área de conhecimento, com o *Mendeley* sendo amplamente utilizado por pesquisadores na área de Ciências da Vida. O *Twitter* e o *Facebook* possuem as menores coberturas para todas as áreas, indicando pouco uso e compartilhamento de conteúdo acadêmico nestas mídias sociais por usuários interessados em artigos publicados no Brasil. As coberturas encontradas estão de acordo com estudos anteriores.

Comparativamente com o cenário internacional, a cobertura de leitores no *Mendeley* é boa, estando de acordo com estudos anteriores para artigos de todas as áreas de conhecimento. Considerando as análises de coberturas e correlações nas amostras geral e por área de conhecimento, é desejável que estudos futuros explorem o potencial do *Mendeley*, para o desenvolvimento e implantação de novas formas de avaliação da qualidade de publicações científicas. A composição de indicadores de avaliação da qualidade deve se adequar às variações existentes dentro da comunicação científica em cada área de conhecimento, e considerar métodos qualitativos para melhor compreender as potencialidades e limitações dos indicadores de leitores do *Mendeley*.

Constatou-se que, de maneira geral, as correlações entre os indicadores altmétricos e bibliométricos se mostraram fracas na maioria das fontes de dados, confirmando que os indicadores de citação e altmétricos medem fenômenos diferentes, e que, portanto, são complementares. Por outro lado, a correlação entre leitores do *Mendeley* e citações se mostrou moderada no universo de artigos dos portais analisados, indicando que o *Mendeley* é a fonte de indicadores altmétricos que possui certa relação de causa e efeito nas citações, com potencial para ser utilizada na composição de novos indicadores para avaliação do impacto acadêmico, especialmente para países com produção científica sub-representada em bases de dados comerciais internacionais.

Assim, os dados apontaram que indicadores altmétricos de leitores no *Mendeley* são os que melhor podem explicar citações futuras, especialmente em determinadas áreas do conhecimento, como a Ciências da Vida. Por outro lado, indicadores altmétricos provenientes das mídias sociais *Twitter* e *Facebook* demonstraram não exercer influência nas citações de artigos no âmbito de portais de periódicos de universidades brasileiras de acesso aberto. Indicadores altmétricos de uso, como acessos ao resumo e downloads, demonstraram ter relação com o número de leitores no *Mendeley* para artigos das áreas de Ciências Sociais Aplicadas e Ciências Exatas, Tecnológicas e Multidisciplinares, entretanto, estas leituras aparentemente não resultam em citações para artigos destas áreas de conhecimento. Portanto, pode-se concluir que, os indicadores do *Mendeley* apresentam menor complementaridade com os indicadores bibliométricos. E, por outro lado, as demais fontes de dados altmétricas, ao apresentarem indicadores com correlações fracas, se apresentam como complementares aos indicadores de citação, visto que aparentemente medem outros tipos de impacto.

Na análise temporal era esperado que houvesse correlações mais fortes, que confirmassem relações entre os indicadores altmétricos (acessos ao resumo e downloads) e os de citações, como observado por estudos anteriores. Entretanto, a maioria das correlações

obtidas foram praticamente nulas. Apesar disso, a correlação mais forte foi identificada para artigos mais antigos, anteriores a 2017, o que indica que há, em algum nível, influência do tempo de publicação do artigo nas correlações entre indicadores altmétricos e de citação. Entretanto, não se descarta a possibilidade de que o agrupamento dos artigos publicados antes de 2017 pode ter exercido alguma influência na correlação apresentada.

Finalmente, este estudo contribuiu nas análises de correlações e coberturas, estando em consonância, muitas vezes, com estudos anteriores, que recomendam o uso conjunto de determinados indicadores altmétricos e indicadores bibliométricos na avaliação da ciência. Constatou-se que alguns dos indicadores possuem altas coberturas, demonstrando grande potencial de utilização para países com produção sub-representada em grande bases de dados comerciais internacionais, a exemplo de parte da ciência publicada pelo Brasil em portais de periódicos de acesso aberto de universidades. Além disso, considera-se que os resultados sejam utilizados para a tomada de decisões pelos gestores de portais de periódicos de universidades brasileiras, implementando e fortalecendo estratégias de comunicação e divulgação científica (a exemplo da participação em mídias sociais), e melhor descrição e organização das revistas e dos artigos hospedados nos portais, proporcionando melhores indicadores de citação em bases como a Crossref.

REFERÊNCIAS

- ABNT. Associação Brasileira de Normas Técnicas. Informação e documentação- citações em documentos- apresentação. **NBR 10520**. Rio de Janeiro: ABNT, 2002.
- ADIE, E. Attention! A study of open access vs non-open access articles. **Altmetric blog**, [S. l.], 23, out. 2014. Disponível em: <http://dx.doi.org/10.6084/m9.figshare.1213690>. Acesso em: 05 jan. 2022.
- AGUIAR, E. D. [Correio eletrônico]. Destinatário: Suporte Crossref. Goiânia, 15, agosto, 2022. 1 e-mail. Endereço: support@crossref.org
- AGUILLO, I.; ORTEGA, J.; FERNÁNDEZ, M.; UTRILLA, A. Indicators for a webometric ranking of open access repositories. **Scientometrics**, [S. l.], v. 82, n. 3, p. 477-486, 2010. Disponível em: <https://akjournals.com/view/journals/11192/82/3/article-p477.xml>. Acesso em: 14 nov. 2021.
- AHMED, A. Open access towards bridging the digital divide—policies and strategies for developing countries. **Information Technology for Development**, v. 13, n. 4, p. 337-361, 2007.
- ALPERIN, J. P. Exploring altmetrics in an emerging country context. *In*: AN ACM WEB SCIENCE CONFERENCE, 2014, Bloomington, Indiana. **Workshop** [...]. New York: Association for Computing Machinery, 2014. Disponível em: <http://altmetrics.org/altmetrics14/>. Acesso em: 15 nov. 2021.
- ALPERIN, J. P. Geographic variation in social media metrics: an analysis of Latin American journal articles. **Aslib Journal of Information Management**, [S. l.], v. 67, n. 3, 2015.
- ALTMETRIC. **Facebook**. [S. l.]: Altmetric, [2020]. Disponível em: <https://help.altmetric.com/support/solutions/articles/6000235936-facebook>. Acesso em: 28 jul. 2022.
- ALTMETRIC. **Who's talking about your research?** [S. l.]: Altmetric, [202-]. Disponível em: <https://www.altmetric.com/>. Acesso em: 15 nov. 2021.
- ARCHAMBAULT, É.; AMYOT, D.; DESCHAMPS, P.; NICOL, A.; PROVENCHER, F.; REBOUT, L.; ROBERGE, G. **Proportion of open access papers published in peer-reviewed journals at the European and world levels: 1996–2013**. [S. l.]: European Commission, 2014. Disponível em: https://science-metrix.com/sites/default/files/science-metrix/publications/d_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11p.pdf. Acesso em: 14 nov. 2021.
- ATKISON-BONASIO, Alice. **Ten years of Mendeley**: and what's next. The revolutionary reference management platform continues to evolve to meet the changing needs of researchers. [S. l.]: Elsevier, 2018.
- AVENA, M. J.; BARBOSA, D. A. Indicadores bibliométricos das Revistas de Enfermagem sob a ótica das bases indexadoras. **Revista da Escola de Enfermagem da USP**, [S. l.], v. 51, 2017.

BARBA, B. M. **Los indicadores bibliométricos**: fundamentos y aplicación al análisis de la ciencia. Gijón, Espanha: Trea, 2003.

BARCELOS, J.; MARICATO J. M.; João; MACÊDO, D. J. Medição da informação científica na Web 2.0: explorando as possibilidades e limitações da plataforma Altmetric. **Ciência da Informação**, v. 49, n. 3, 2020. Disponível em: <https://revista.ibict.br/ciinf/article/view/5657> Acesso em: 27 out. 2022.

BARCELOS, J.; MARICATO, J. M. Visibilidade e engajamento público na web 2.0: um estudo altmétrico a partir dos artigos publicados na Scientific Data. **Em Questão**, v. 27, n. 1, p. 263-285, 2021. Disponível em: <https://doi.org/10.19132/1808-5245271.263-285>. Acesso em: 14 jul. 2022.

BAR-ILAN, J. Citations to the “Introduction to informetrics” indexed by Web of Science, Scopus and Google Scholar. **Scientometrics**, v. 82, n. 3, p. 495-506, 2010. Disponível em: <https://doi.org/10.1007/s11192-010-0185-9>. Acesso em: 16 jan. 2022.

BAR-ILAN, J.; HAUSTEIN, S.; PETERS, I.; PRIEM, J.; SHEMA, H.; TERLIESNER, J. Beyond citations: scholars' visibility on the social web. **arXiv**, [S. l.], 2012. Disponível em: <https://arxiv.org/abs/1205.5611>. Acesso em: 27 nov. 2021.

BJÖRNEBORN, L.; INGWERSEN, P. Toward a basic framework for webometrics. **Journal of the American Society for Information Science and Technology**, [S. l.], v. 55, n. 14, p. 1216–1227, 2004. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/asi.20077>. Acesso em: 15 out. 2021.

BOAI. BUDAPEST OPEN ACCESS INITIATIVE. Leia a declaração. Budapest: Hungary, 2002. Disponível em: <http://www.budapestopenaccessinitiative.org/read>. Acesso em: 19 nov. 2021.

BLOG CROSSREF, 2022. Amendments to membership terms to open reference distribution and include UK jurisdiction. Disponível em: <https://www.crossref.org/blog/amendments-to-membership-terms-to-open-reference-distribution-and-include-uk-jurisdiction/>. Acesso em: 01 nov. 2022.

BORNMANN, L. Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. **Scientometrics**, [S. l.], v. 103, n. 3, p. 1123-1144, 2015. Disponível em: <https://link.springer.com/article/10.1007%2Fs11192-015-1565-y>. Acesso em: 13 ago. 2021.

BORNMANN, L. Validity of altmetrics data for measuring societal impact: a study using data from Altmetric and F1000 Prime. **Journal of informetrics**, [S. l.], v. 8, n. 4, p. 935-950, 2014.

BORNMANN, L.; MUTZ, R. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. **Journal of the Association for Information Science and Technology**, v. 66, n. 11, p. 2215-2222, 2015. Disponível em: <https://doi.org/10.1002/asi.23329>. Acesso em: 15 jan. 2022.

BRASIL. Ministério da Educação. Centro de Aperfeiçoamento de Pessoal de Nível Superior. **Quem somos**. Brasília: CAPES, ©2020. Disponível em: <https://www-periodicos-capes-gov-br.ez1.periodicos.capes.gov.br/index.php/sobre/quem-somos.html>. Acesso em: 12 dez. 2021.

BRASIL. Ministério da Educação. Centro de Aperfeiçoamento de Pessoal de Nível Superior. 2022. **Sobre as áreas de avaliação**. Brasília: CAPES, 2022. Disponível em: <https://www.gov.br/capes/pt-br/aceso-a-informacao/acoes-e-programas/avaliacao/sobre-a-avaliacao/areas-avaliacao/sobre-as-areas-de-avaliacao/sobre-as-areas-de-avaliacao>. Acesso em: 28 abr. 2022.

BRASIL. Ministério da Educação. **Plataforma Sucupira**. Brasília: MEC, c2016b. Disponível em: <https://sucupira.capes.gov.br/sucupira/public/index.jsf>. Acesso em: 12 dez. 2021.

BRASIL. Ministério da Educação. Qualis. **Qualis periódicos**. Brasília: MEC, c2016a. Disponível em: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf>

BRIGHAM, T. J. An introduction to altmetrics. **Medical Reference Services Quarterly**, [S. l.], v. 33, n. 4, p. 438-447, 2014. DOI: 10.1080/02763869.2014.957093.

BUCHANAN, R. A. Accuracy of cited references: The role of citation databases. **College & Research Libraries**, v. 67, n. 4, p. 292-303, 2006. Disponível em: <https://doi.org/10.5860/crl.67.4.292>. Acesso em: 27 out. 2022.

BUENO; C. O. M.; AGUIAR; E. D.; SILVA, J. T. A. **Diretrizes para atribuição do DOI nos artigos de periódicos hospedados no portal de periódicos da UFG**. Goiânia: UFG, 2018. Disponível em: https://portal.revistas.ufg.br/revistas_ufg/wp-content/uploads/2021/08/Diretrizes-maio-2019-v2.pdf. Acesso em: 08 ago. 2021.

CAFÉ, L.; BRÄSCHER, M. Organização da informação e bibliometria. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, n. Esp., p. 54-75, 2008. Disponível em: <https://www.redalyc.org/pdf/147/14709806.pdf>. Acesso em: 7 nov. 2021.

CÉSPEDES, L. Revistas latino-americanas e línguas hegemônicas para publicação acadêmica no Scopus e Web of Science. **Trabalhos em Linguística Aplicada**, v. 60, n. 1, p. 141-154, 2021. Disponível em: <https://doi.org/10.1590/010318138901311520201214>. Acesso em: 15 jan. 2022.

CHEN, W. M.; BUKHARI, M.; COCKSHULL, F.; GALLOWAY, J. The relationship between citations, downloads and alternative metrics in rheumatology publications: a bibliometric study. **Rheumatology**, [S. l.], v. 59, n. 2, p. 277-280, 2020.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Porto Alegre: Artmed editora, 2010.

CHO, J. A comparative study of the impact of Korean research articles in four academic fields using altmetrics. **Performance Measurement and Metrics**, [S. l.], v. 16, n. 1, p. 38-51, 2017. DOI: 10.1108/PMM-02-2016-0005.

CHUDLARSKÝ, T.; DVOŘÁK, J. Can crossref citations replace web ofscience for research evaluation? The share of open citations. **Journal of Data and Information Science**, [S. l.], v. 2, n. 2, Sep. 2020. Disponível em: <https://sciendo.com/article/10.2478/jdis-2020-0037>. Acesso em: 12 dez. 2021.

CINTRA, Paulo Roberto; FURNIVAL, Ariadne Chloe; MILANEZ, Douglas Henrique. Vantagens de citação do acesso aberto em periódicos selecionados da Ciência da Informação: uma análise ampliada aos indicadores altmétricos. **Informação & Informação**, [S. l.], v. 22, n. 1, p. 129-149, jun. 2017. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/27146>. Acesso em: 12 dez. 2021.

CIOFFI, A.; COPPINI, S.; MASSARI, A.; MORETTI, A.; PERONI, S.; SANTINI, C.; SHAHIDZADEH ASADI, N. Identifying and correcting invalid citations due to DOI errors in Crossref data. **Scientometrics**, p. 1-20, 2022. Disponível em: <https://doi.org/10.1007/s11192-022-04367-w>. Acesso em: 27 out. 2022.

CLARIVATE. Web of Science todas as bases de dados - ajuda. [S. l.]: Clarivate, 2020. Disponível em: https://images.webofknowledge.com/WOKRS516B3/help/pt_BR/WOK/hp_database.html#dsy366-TRS_wos. Acesso em: 12 dez. 2021.

COUNTER. Counter provides the standard that enables the knowledge community to count the use of electronic resources. [S. l.], 2021. Disponível em: <https://www.projectcounter.org/>. Acesso em: 12 dez. 2021.

COSTAS, R.; ZAHEDI, Z.; WOUTERS, P. Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. **JASIST**, [S. l.], v. 66, n. 10, p. 2003-2019, jul. 2014. DOI: <https://doi.org/10.1002/asi.23309>. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.23309>. Acesso em: 13 mar. 2020.

COSTA, S. M. S. **The impact of computer usage on scholarly communication amongst academic social scientists**. 1999. 318 f. Tese (Doutorado em Ciência da Informação) - Loughborough University, Inglaterra, 1999.

CROSSREF. **About us**. [S. l.]: Crossref, 2021. Disponível em: <https://www.crossref.org/about/>. Acesso em: 1 dez. 2021.

CROSSREF. **Documentation**. [S. l.]: Crossref, 2020. Disponível em: <https://www.crossref.org/documentation/retrieve-metadata/rest-api/>. Acesso em: 5 nov. 2021.

CROSSREF SUPPORT. Questions about Crossref citation counts [mensagem pessoal]. Mensagem recebida por <erikademachki@gmail.com> em 26 out. 2021.

DELGADO LÓPEZ-CÓZAR, E.; ORDUÑA-MALEA, E.; MARTÍN-MARTÍN, A. Google Scholar as a data source for research assessment. In: GLAENZEL, W.; MOED, H.; SCHMOCH, U.; THELWALL, M. **Springer handbook of science and technology indicators**. Berlin: Springer, 2019.

DEMACHKI, E.; MARICATO, J. M. Cobertura de revistas, artigos e citações em bases de dados: uma análise a partir do Portal de Periódicos da UFG. In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 8., 2022, Maceió. **Proceedings** [...]. Maceió: Universidade Federal de Alagoas, 2022a. p. 110-117. Disponível em: <https://ebbc.inf.br/ojs/index.php/ebbc/article/view/52>. Acesso em: 15 set. 2022.

DEMACHKI, E.; MARICATO, J. M. Coverage of Data Sources and Correlations Between Altmetrics and Citation Indicators: The Case of a Brazilian Portal of Open Access Journals. **Serials Review**, p. 1-16, 2022b. Disponível em: <https://doi.org/10.1080/00987913.2022.2066967>

DEMETER, M. A. **Academic knowledge production and the global south: questioning inequality and under-representation**. Basingstoke: Palgrave Macmillan, 2020.

DEMETER, M. A.; ISTRATII, R. Scrutinising what open access journals mean for global inequalities. **Publishing Research Quarterly**, [S. l.], v. 36, n. 4, p. 505-522, 2020. Disponível em: <https://link.springer.com/article/10.1007/s12109-020-09771-9>. Acesso em: 5 jan. 2022.

DOAJ. Directory of Open Access Journals. **Find open access journals & articles**. [S. l.]: DOAJ, c2020. Disponível em: <https://doaj.org/>. Acesso em: 15 nov. 2021.

EJIKEME, A. N.; EZEMA, I. J. The potentials of open access initiative and the development of institutional repositories in Nigeria: implications for scholarly communication. **Publishing Research Quarterly**, [S. l.], v. 35, n. 1, p. 6-21, 2019. Disponível em: <https://doi.org/10.1007/s12109-018-09626-4>. Acesso em: 5 jan. 2022.

ELLEGAARD, O.; WALLIN, J. A. The bibliometric analysis of scholarly production: How great is the impact?. **Scientometrics**, v. 105, n. 3, p. 1809-1831, 2015.

ELSE, H. How unpaywall is transforming open science. **Nature**, [S. l.], 2018. Disponível em: <https://unpaywall.org/>. Acesso em: 10 dez. 2021.

ELSEVIER. **Scopus content guide**. [S. l.]: Elsevier, 2020. Disponível em: <https://www.elsevier.com/?a=69451> Acesso em: 6 jan. 2022.

ERFANMANESH, M. The presence of Iranian information science and library science articles in social media: an altmetric study. **Iranian Journal of Information Processing & Management**, [S. l.], v. 32, n. 2, p. 349-373, 2017. Disponível em: <https://www.researchgate.net/publication/284869680> The Presence of Iranian Information Science and Library Science Articles in Social Media An Altmetric Study. Acesso em: 12 dez. 2021.

EYSENBACH, G. Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. **Journal of medical Internet research**, [S. l.], v. 13, n. 4, p. 123, 2011.

EZEMA, I. J.; UGWU, C. I. Correlating research impact of library and information science journals using citation counts and altmetrics attention. **Information Discovery and Delivery**, [S. l.], v. 47, n. 3, p. 143-153, 2019.

FARIAS, S. A. Internacionalização dos periódicos brasileiros. **Revista de Administração de Empresas**, [S. l.], v. 57, p. 401-404, 2017.

FREITAS, M. H. Considerações acerca dos primeiros periódicos científicos brasileiros. **Ciência da Informação**, [S. l.], v. 35, p. 54-66, 2006. Disponível em: <https://www.scielo.br/j/ci/a/RRqQp5h4xm5FSn7dSK99gTG/?lang=pt&format=pdf>. Acesso em: 13 dez. 2022.

GALLIGAN, F.; DYAS-CORREIA, S. Altmetrics: rethinking the way we measure. **Serials Review**, [S. l.], v. 39, n. 1, p. 56-61, 2013. DOI: <https://doi.org/10.1016/j.serrev.2013.01.003>

GARFIELD, E. Journal impact factor: a brief review. **Cmaj**, [S. l.], v. 161, n. 8, p. 979-980, 1999. Disponível em: <https://www.cmaj.ca/content/161/8/979.short>. Acesso em: 21 out. 2021.

GARRIDO, I. S.; RODRIGUES, R. S. Portais de periódicos científicos online: organização institucional das publicações. **Perspectivas em Ciência da Informação**, [S. l.], v. 15, p. 56-72, 2010. Disponível em: <https://www.scielo.br/j/pci/a/3DY77hcvCtFvsc85r4SYrjs/abstract/?lang=pt>. Acesso em: 21 out. 2021.

GEORGE, D.; MALLERY, P. **IBM SPSS statistics 26 step by step: A simple guide and reference**. Routledge, 2019.

GOUVEIA, F. C. Estudos altmétricos no Brasil: uma análise a partir dos currículos da Plataforma Lattes-CNPq. **Transinformação**, [S. l.], v. 31, 2019.

GOUVEIA, F. C. Altmetria: métricas de produção científica para além das citações | Altmetrics: scientific production metrics beyond citations. **Liinc em revista**, [S. l.], v. 9, n. 1, 2013. Disponível em: <http://revista.ibict.br/liinc/article/view/3434>. Acesso em: 13 ago. 2021.

HAMMARFELT, B. Using altmetrics for assessing research impact in the humanities. **Scientometrics**, [S. l.], v. 101, n. 2, p. 1419-1430, 2014. Disponível em: <https://doi.org/10.1007/s11192-014-1261-3>. Acesso em: 22 jul. 2020.

HAUSTEIN, S.; COSTAS, R.; LARIVIÈRE, V. Characterizing social media metrics of scholarly papers: the effect of document properties and collaboration patterns. **PloS one**, [S. l.], v. 10, n. 3, 2015. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120495>. Acesso em: 12 dez. 2021.

HAUSTEIN, S.; LARIVIÈRE, V. Mendeley as a source of readership by students and postdocs? Evaluating article usage by academic status. *In: ANNUAL INTERNATIONAL ASSOCIATION OF TECHNOLOGICAL UNIVERSITY LIBRARIES*, 35., 2014, Finland. **Proceedings** [...]. Helsinki: Finland, 2014. Disponível em: <https://docs.lib.purdue.edu/iatul/2014/altmetrics/2/>. Acesso em: 22 jul. 2020.

HAUSTEIN, S.; LARIVIÈRE, V. The use of bibliometrics for assessing research: possibilities, limitations and adverse effects. *In: WELPE; I. M.; WOLLERSHEIM, J.;*

RINGELHAN, S.; OSTERLOH, Margit. **Incentives and performance: governance of research organizations.** Springer, Cham, 2015. p. 121-139.

HAUSTEIN, S.; LARIVIÈRE, V.; THELWALL, M.; AMYOT, D.; PETERS, I. Tweets vs. Mendeley readers: how do these two social media metrics differ? **arXiv**, [S. l.], v. 56, n. 5, p. 207-215, 2014.

HEIBI, I.; PERONI, S.; SHOTTON, D. Software review: COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. **Scientometrics**, v. 121, n. 2, p. 1213-1228, 2019. Disponível em: <https://doi.org/10.1007/s11192-019-03217-6>. Acesso em: 27 out. 2022.

HOLMBERG, K.; HEDMAN, J.; BOWMAN, T. D.; DIDEGAH, F.; LAAKSO, M. Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. **Scientometrics**, [S. l.], v. 122, n. 1, p. 645-659, 2020.

HOLMBERG, K.; THELWALL, M. Disciplinary differences in Twitter scholarly communication. In: INTERNATIONAL SOCIETY OF SCIENTOMETRICS AND INFORMETRICS CONFERENCE, 14., 2013, Vienna. **Proceedings** [...]. Vienna: Austrian Institute of Technology GmbH, 2013. p. 567-592. Disponível em: https://www.issi-society.org/proceedings/issi_2013/ISSI_Proceedings_Volume_I.pdf. Acesso em: 12 dez. 2021.

HOOD, W. W.; WILSON, C. S. The literature of bibliometrics, scientometrics, and informetrics. **Scientometrics**, [S. l.], v. 52, n. 2, p. 291-314, 2001. Disponível em: <https://link.springer.com/article/10.1023/A:1017919924342>. Acesso em: 20 out. 2021.

HUANG, W.; WANG, P.; WU, Q. A correlation comparison between Altmetric Attention Scores and citations for six PLOS journals. **PLoS One**, [S. l.], Apr. 2018. <https://doi.org/10.1371/journal.pone.0194962>

IMPACTSTORY. **Discover the online impact of your research.** [S. l.]: Impactstory, [202-]. Disponível em: <https://profiles.impactstory.org/>. Acesso em: 10 dez. 2021.

ISTRATII, R.; DEMETER, M. Plan S and the ‘opening up’ of scientific knowledge: a critical commentary. **Decolonial Subversions**, [S. l.], p. 13–21, 2020. Disponível em: <https://eprints.soas.ac.uk/32525/>. Acesso em: 5 jan. 2022.

JAMOVI. The Jamovi Project. Jamovi, versão 1.6. 2021. Disponível em: <https://www.jamovi.org/download.html> Acesso em: 26 de abril de 2022.

JOKIĆ, M.; MERVAR, A.; MATELJAN, S. Scientific potential of European fully open access journals. **Scientometrics**, [S. l.], v. 114, n. 3, p. 1373-1394, 2018.
KEELE, B. J. Law libraries as publishers: counting things, with altmetrics. **ALL-SIS Newsletter**, [S. l.], v. 35, n. 2, p. 15-16, 2016.

KONKIEL, S. What Jeffrey Beall gets wrong about altmetrics. **OurResearch blog**, [S. l.], 9 Sep. 2014. Disponível em: <https://blog.ourresearch.org/beall-altmetrics/>. Acesso em: 12 dez. 2021.

KONKIEL, S.; SCHERER, D. New opportunities for repositories in the age of altmetrics. *Asis&t*, v. 39, n. 4, 2013. Disponível em: <https://doi.org/10.1002/bult.2013.1720390408>. Acesso em: 12 dez. 2021.

LAGOTTO. **About**. [S. l.]: Lagotto, [202-]. Disponível em: <https://www.lagotto.io/>. Acesso em: 20 nov. 2021.

LANCASTER, F. W. The evolution of electronic publishing. *Library Trends*, [S. l.], v. 43, n. 4, p. 518-527, 1995.

LATOUR, B.; WOOLGAR, S. **Laboratory life**. [S. l.]: Princeton University Press, 2013.

LI, X.; THELWALL, M.; GIUSTINI, D. Validating online reference managers for scholarly impact measurement. *Scientometrics*, [S. l.], v. 91, n. 2, p. 461-471, 2012. Disponível em: <https://doi.org/10.1007/s11192-011-0580-x>. Acesso em: 12 dez. 2021.

LIU, C. L.; XU, Y. Q.; WU, H.; CHEN, S. S.; GUO, J. J. Correlation and interaction visualization of altmetric indicators extracted from scholarly social network activities: dimensions and structure. *Journal of Medical Internet Research*, [S. l.], v. 15, n. 11, 2013. Disponível em: <https://www.jmir.org/2013/11/e259>. Acesso em: 14 abr. 2021.

MACHLUP, F. **The production and distribution of knowledge in the United States**. New Jersey: Princeton University, 1962.

MAFLAHI, N.; THELWALL, M. When are readership counts as useful as citation counts? Scopus versus Mendeley for LIS journals. *J Assn Inf Sci Tec*, [S. l.], v. 67, p. 191-199, 2016. <https://doi.org/10.1002/asi.23369>. Acesso em: 12 dez. 2021.

MALIN, M. V. The science citation index: a new concept in indexing. *Library Trends*, [S. l.], v. 16, p. 374-387, 1968. Disponível em: https://www.ideals.illinois.edu/bitstream/handle/2142/6392/librarytrendsv16i3h_opt.pdf. Acesso em: 21 out. 2021.

MARICATO, J. M. **Dinâmica das relações entre Ciência e Tecnologia: estudo bibliométrico e cientométrico de múltiplos indicadores de artigos e patentes em biodiesel**. Tese (Doutorado em Ciência da Informação) - Universidade de São Paulo. São Paulo, 2010. 359 p.

MARICATO, J. M.; MARTINS, D. L. Altmetrics: complexities, challenges and new forms of measuring and comprehending scientific communication in the social. *Biblios*, [S. l.], v. 68, n. 68, p. 48-68, 2017. Disponível em: <https://doi.org/10.5195/biblios.2017.358>. Acesso em 03 mar. 2021.

MARICATO, J. M.; SANTOS, P. R. Highly Cited Publications Index (HCP) for individual evaluation: applications in Brazilian Information Metric Studies. *Library Philosophy and Practice (e-journal)*. 6394, 2021. Disponível em: <https://www.researchgate.net/publication/356065311>. Acesso em: 15 jan. 2022.

MARICATO, J. M.; VILAN FILHO, J. L. The potential for altmetrics to measure other types of impact in scientific production: academic and social impact dynamics in social media and

networks. **Information Research**, [S. l.], v. 23, n. 1, p. 780, 2018. Disponível em: https://repositorio.unb.br/bitstream/10482/32082/3/ARTIGO_PotencialAltmtriaMedir.pdf. Acesso em: 21 out. 2021.

MARCHIORI, P. Z.; BETTONI, E. M.; CARVALHO, M. B.; APPEL, A. L. Geração de indicadores para periódicos científicos abertos. **Transinformação**, [S. l.], v. 30, p. 324-335, 2018.

MARTÍN-MARTÍN, A.; ORDUNA-MALEA, E.; THELWALL, M.; LÓPEZ-CÓZAR, E. D. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. **Journal of informetrics**, [S. l.], v. 12, n. 4, p. 1160-1177, 2018.

MARTÍN-MARTÍN, A.; THELWALL, M.; ORDUNA-MALEA, E.; LÓPEZ-CÓZAR, E. D. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. **Scientometrics**, [S. l.], v. 126, n. 1, p. 871-906, 2021. Disponível em: <https://link.springer.com/article/10.1007/s11192-020-03690-4>. Acesso em: 12 dez. 2021.

MCCABE, M. J.; SNYDER, C. M. Identifying the effect of open access on citations using a panel of science journals. **Economic inquiry**, [S. l.], v. 52, n. 4, p. 1284-1300, 2014. Disponível em: <https://doi.org/10.1111/ecin.12064>. Acesso em: 25 out. 2021.

MCGILLIVRAY, B.; ASTELL, M. The relationship between usage and citations in an open access mega-journal. **Scientometrics**, v. 121, n. 2, p. 817-838, 2019. Disponível em: <https://doi.org/10.1007/s11192-019-03228-3>. Acesso em: 12 out. 2022.

MEADOWS, A. J. **A comunicação científica**. Reino Unido: Loughborough University, 1999.

MELO, J. H. N.; TRINCA, T. P.; MARICATO, J. M. Limites dos indicadores bibliométricos de bases de dados internacionais para avaliação da Pós-Graduação brasileira: a cobertura da *Web of Science* nas diferentes áreas do conhecimento. **Transinformação**, [S. l.], v. 33, 2021. Disponível em: <https://doi.org/10.1590/2318-0889202133e200071>. Acesso em: 3 jan. 2022.

MENDELEY: Simplify your workflow, accelerate your research. [S. l.]: Elsevier, c2021. Disponível em: <https://www.elsevier.com/solutions/mendeley>. Acesso em: 10 dez. 2021.

MINGERS, J.; YANG, L. Evaluating journal quality: a review of journal citation indicators and ranking in business and management. **European Journal of Operational Research**, [S. l.], v. 257, n. 1, p. 323-337, 2016.

MOED, H. F. Statistical relationships between downloads and citations at the level of individual documents within a single journal. **Journal of the American Society for Information Science and Technology**, [S. l.], v. 56, n. 10, p. 1088-1097, 2005.

MOED, H. F.; HALEVI, G. On full text download and citation distributions in scientific-scholarly journals. **Journal of the Association for Information Science and Technology**, v. 67, n. 2, p. 412-431, 2016.

MOHAMMADI, E.; THELWALL, M. Assessing the Mendeley readership of social science and humanities research. *In*: INTERNATIONAL SOCIETY OF SCIENTOMETRICS AND

INFORMETRICS CONFERENCE, 14., 2013, Vienna. **Proceedings** [...]. Vienna: Austrian Institute of Technology GmbH, 2013. p. 200–214. Disponível em: https://www.issi-society.org/proceedings/issi_2013/ISSI_Proceedings_Volume_I.pdf. Acesso em: 12 dez. 2021.

MOHAMMADI, E.; THELWALL, M. Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. **Journal of the American Society for Information Science and Technology**, [S. l.], v. 65, n. 8, 2014. Disponível em: <https://doi.org/10.1002/asi.23071>. Acesso em: 14 abr. 2020.

MOHAMMADI, E.; THELWALL, M.; HAUSTEIN, S.; LARIVIÈRE, V. Who reads research articles? An altmetrics analysis of Mendeley user categories. **Journal of the Association for Information Science and Technology**, [S. l.], v. 66, n. 9, p. 1832-1846, 2015.

MOREIRA, L. F. P. Is there any influence of journal impact factor on the citation index of articles published in Brazilian journals with international scope? **Arquivos Brasileiros de Cardiologia**, [S. l.], v. 105, n. 1, p. 1-2, 2015. DOI: <https://doi.org/10.5935/abc.20150091>. Disponível em: <https://www.scielo.br/j/abc/a/JG4SJGbrwQ5Dd8dxmqrxZZM/?lang=en>. Acesso em: 10 dez. 2021.

MUELLER, S. P. M. A publicação da ciência: áreas científicas e seus canais preferenciais. **DataGramZero**, Brasília, v. 6, n. 1, 2005. Disponível em: <https://repositorio.unb.br/handle/10482/980>. Acesso em: 5 dez. 2021.

MUGNAINI, R. **Caminhos para adequação da avaliação da produção científica brasileira: impacto nacional versus internacional**. 2006. 254 f. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2006.

MUGNAINI, R.; DAMACENO, R. J. P.; DIGIAMPIETRI, L. A.; MENA-CHALCO, J. P. Panorama da produção científica do Brasil além da indexação: uma análise exploratória da comunicação em periódicos. **Transinformação**, Campinas, SP, v. 31, 2019. Disponível em: <https://doi.org/10.1590/2318-0889201931e190033>. Acesso em: 04 jan. 2022.

NORRIS, M.; OPPENHEIM, C. Comparing alternatives to the Web of Science for coverage of the social sciences' literature. **Journal of informetrics**, v. 1, n. 2, p. 161-169, 2007. Disponível em: <https://doi.org/10.1016/j.joi.2006.12.001>. Acesso em: 16 jan. 2022.

OLIVEIRA, É. B. P. M. Periódicos científicos eletrônicos: definições e histórico. **Informação & Sociedade**, João Pessoa, v. 18, n. 2, 2008. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/1701>. Acesso em: 10 dez. 2021.

OLIVEIRA, É. B. P. M. **Uso de periódicos científicos eletrônicos por docentes e pós-graduandos do Instituto de Geociências da USP**. 2006. 140 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de São Paulo, São Paulo, 2006.

ORTEGA, J. L. Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event. **Scientometrics**, [S. l.], v. 116, p. 2123-2138, 2018.

OUCHI, A.; SABERI, M. K.; ANSARI, N.; HASHEMPOUR, L.; ISFANDYARI-MOGHADDAM, A. Do altmetrics correlate with citations? A study based on the 1,000 most-cited articles. **Information Discovery and Delivery**, [S. l.], v. 47, n. 4, p. 192-202, 2019. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/IDD-07-2019-0050/full/pdf?title=do-altmetrics-correlate-with-citations-a-study-based-on-the-1000-most-cited-articles>. Acesso em: 19 jul. 2021.

PACKER, A. L. A eclosão dos periódicos do Brasil e cenários para o seu porvir. **Educação e Pesquisa**, v. 40, p. 301-323, 2014.

PACKER, A. L. Os periódicos brasileiros e a comunicação da pesquisa nacional. **Revista USP**, São Paulo, n. 89, p. 26-61, 2011.

PASSOS, P. C. S. J.; PASSOS, J. E.; CAREGNATO, S. E.; SILVA, T. L. K. D. Critérios de qualidade em periódicos científicos. **Informação & Sociedade**, João Pessoa, v. 28, n. 2, p. 209-226, maio/ago. 2018.

PERNEGER, T. V. Relation between online “hit counts” and subsequent citations: prospective study of research papers in the BMJ. **Bmj**, [S. l.], v. 329, n. 7465, p. 546-547, 2004. Disponível em: <https://www.bmj.com/content/bmj/329/7465/546.full.pdf>. Acesso em: 26 nov. 2021.

PINTO, C. S.; COSTA, J. L. Padrões de comunicação em diferentes comunidades científicas. *In*: COSTA, Sely Maria de Souza; LEITE, Fernando César Lima; TAVARES, Rosemeire Barbosa (org.). **Comunicação da informação, gestão da informação e gestão do conhecimento**. Brasília: IBICT, 2018. p. 145-159.

PIWOWAR, H.; PRIEM, J.; LARIVIÈRE, V.; ALPERIN, J. P.; MATTHIAS, L.; NORLANDER, B.; FARLEY, A.; WEST, J.; HAUSTEIN, S. The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles. **PeerJ**, [S. l.], v. 6, 2018. DOI <https://doi.org/10.7717/peerj.4375>

PKP. Public Knowledge Project. **Journals Publishing with Open Journal Systems (OJS) by Year and Region**. [S. l.]: PKP, c2021. Disponível em: [Pkp 2021 https://pkp.sfu.ca/ojs/stats](https://pkp.sfu.ca/ojs/stats). Acesso em: 15 jan. 2021a.

PKP. Public Knowledge Project. **Download**. [S. l.]: PKP, c2021. Disponível em: [Pkp 2021 https://pkp.sfu.ca/ojs/](https://pkp.sfu.ca/ojs/). Acesso em: 8 dez. 2021b.

PLUM ANALYTICS. **About artifacts**. [S. l.]: PLUM, c2021. Disponível em: <https://plumanalytics.com/learn/about-artifacts/>. Acesso em: 8 dez. 2021.

POST list. **Research blogging**, [S. l.], c2015. Disponível em: <http://researchblogging.org/>. Acesso em: 20 dez. 2021.

PRICE, D. S. A general theory of bibliometric and other cumulative advantage processes. **Journal of the American society for Information science**, [S. l.], v. 27, n. 5, p. 292-306, 1976.

PRIEM, J. Altmetrics. In: CRONIN, B.; SUGIMOTO, C. R. (Ed.) **Beyond bibliometrics: harnessing multidimensional indicators of scholarly impact**. Cambridge, Massachusetts: MIT Press, 2014.

PRIEM, J. **I like the term #article level metrics, but it fails to imply *diversity* of measures**. Lately, I'm liking #altmetrics. [S. l.], 28 set. 2010. Twitter: @jasonpriem. Disponível em: <https://twitter.com/jasonpriem/status/25844968813>. Acesso em: 10 dez. 2021.

PRIEM, J.; GROTH, P.; TARABORELLI, D. The altmetrics collection. **PloS one**, [S. l.], v. 7, n. 11, p. e48753, 2012.

PRIEM, J.; PIWOWAR, H. A.; HEMMINGER, B. M. Altmetrics in the wild: using social media to explore scholarly impact. **arXiv.org**, [S. l.], 2011. Disponível em: <https://arxiv.org/abs/1203.4745>. Acesso em: 6 dez. 2021.

PRIEM, J.; TARABORELLI, D.; GROTH, P.; NEYLON, C. **Altmetrics: a manifesto**. London: Altmetric, 2011. Disponível em: <http://altmetrics.org/manifesto/>. Acesso em: 9 dez. 2021.

PRITCHARD, A. Statistical bibliography or bibliometrics. **Journal of documentation**, [S. l.], v. 25, n. 4, p. 348-349, 1969.

PURKAYASTHA, A.; PALMARO, E.; FALK-KRZESINSKI, H. J.; BAAS, J. Comparison of two article-level, field-independent citation metrics: Field-Weighted Citation Impact (FWCI) and Relative Citation Ratio (RCR). **Journal of Informetrics**, v. 13, n. 2, p. 635-642, 2019. Disponível em: <https://doi.org/10.1016/j.joi.2019.03.012>. Acesso em: 15 jan. 2022.

RAISIG, L. Miles. Statistical bibliography in the health sciences. **Bulletin of the Medical Library Association**, [S. l.], v. 50, n. 3, p. 450, 1962.

RAMOS, T. B. C.; MARICATO, J. M. **Odisseia métrics: ferramenta de extração de dados**. Versão 1.0. [S. l.: s. n.], 2020. Disponível em: https://colab.research.google.com/drive/1L75mJbqteVE0kMoTMK1_kLVXWFZpQumi#scrollTo=ohRqrEb7RXlm

RASMUSSEN, P. G.; ANDERSEN, J. P. Altmetrics: An alternate perspective on research evaluation. **Sciecom info**, [S. l.], v. 9, n. 2, 2013. Disponível em: <https://journals.lub.lu.se/sciecominfo/article/view/7292>. Acesso em: 22 out. 2021.

ROBINSON-GARCÍA, N.; TORRES-SALINAS, D.; ZAHEDI, Z.; COSTAS, R. New data, new possibilities: exploring the insides of altmetric.com. **Profesional de la Informacion**, [S. l.], v. 23, n. 4, p. 359-366, 2014. Disponível em: <https://arxiv.org/abs/1408.0135>. Acesso em: 6 dez. 2021.

RODRIGUES, R. S.; FACHIN, G. R. B. A comunicação científica e o uso de portais: estudo. ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais [...]**. São Paulo: ANCIB, 2008.

RODRIGUES, R. S.; FACHIN, G. R. B. Portal de periódicos científicos: um trabalho multidisciplinar. **Transinformação**, [S. l.], v. 22, p. 33-45, 2010.

ROEMER, R. C.; BORCHARDT, R. Issues, controversies, and opportunities for altmetrics. **Library Technology Reports**, [S. l.], v. 51, n. 5, p. 20-30, 2015.

SANTOS, G. C. Panorama dos Portais de Periódicos no Brasil. **Encontro Nacional de Portais de Periódicos**, Campinas, SP, v. 2, p. 32 slides, 2019. Disponível em: <https://econtents.bc.unicamp.br/eventos/index.php/enapp/article/view/1557>. Acesso em: 15 jan. 2022.

SANTOS, T. H. **Comunicação da ciência e web 2.0**. 2011. Dissertação (Mestrado em História e Filosofia das Ciências) - Universidade de Lisboa, Lisboa, 2011.

SCIELO. Critérios, política e procedimentos para a admissão e a permanência de periódicos científicos na Coleção SciELO Brasil. São Paulo, 2014. Disponível em: <https://wp.scielo.org/wp-content/uploads/20140900-Criterios-SciELO-Brasil.pdf>. Acesso em: 19 ago. 2022.

SCIELO ANALYTICS. **Collection composition**. São Paulo: Scielo, 2021. Disponível em: <https://analytics.scielo.org/>. Acesso em: 15 dez. 2021.

SENGUPTA, I. N. Bibliometrics, informetrics, scientometrics and librametrics: an overview. **LIBRI**, [S. l.], v. 42, n. 2, p. 75-9, 1992. Disponível em: <https://www.degruyter.com/document/doi/10.1515/libr.1992.42.2.75/html>. Acesso em: 6 dez. 2021.

SHEMA, H.; BARILAN, J.; THELWALL, M. Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. **Journal of the Association for information science and technology**, [S. l.], v. 65, n. 5, p. 1018-1027, 2014.

SILVA FILHO, R. C. **Influência das métricas de mídias sociais nas citações dos periódicos brasileiros de Enfermagem indexados na Scopus**. 2022. Tese (Doutorado) - Programa de Pós-Graduação em Comunicação e Informação, Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022.

SILVA, E. L., MENEZES, E. M. Metodologia da pesquisa e elaboração de dissertação. UFSC, Florianópolis, 4ª edição, v. 123, 2005. Disponível em: https://www.academia.edu/download/33206387/metodologia_da_pesquisa_e_elaboracao_de_dissertacao.pdf. Acesso em: 28 dez. 2022.

SILVA, R. C.; VANZ, S. A.S. Impacto de altmetrics sobre a visibilidade de artigos em acesso aberto da enfermagem brasileira: um estudo de caso. **Transinformação**, v. 31, 2019. Disponível em: <https://doi.org/10.1590/2318-0889201931e190025>. Acesso em: 12 out. 2022.

SINGH, V. K.; SINGH, P.; KARMAKAR, M.; LETA, J.; MAYR, P. The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. **Scientometrics**, v. 126, n. 6, p. 5113-5142, 2021. Disponível em: <https://doi.org/10.1007/s11192-021-03948-5>. Acesso em 05 set. 2022.

SOLOMON, D.; BJÖRK, B. Article processing charges for open access publication—the situation for research intensive universities in the USA and Canada. **PeerJ**, [S. l.], v. 4, p. e2264, 2016.

SOUTO, P. N. E-publishing development and changes in the scholarly communication system. **Ciência da Informação**, Brasília, v. 36, n. 1, p. 158-166, jan./abr. 2007. Disponível em: <https://www.scielo.br/j/ci/a/T3XwR6VN5BsPHYXrgTBPnhG/?lang=en>. Acesso em: 11 dez. 2021.

SOUZA, M. Produção científica brasileira: caminhos norteadores para instituições de fomento à pesquisa. **Brazilian Journal of Information Science: research trends**, [S. l.], v. 12, n. 1, 2018.

SUBER, P. Open access overview. **Open Access: Unrestricted Access to Published Research**, p. 7-12, 2006.

SUD, P.; THELWALL, M. Evaluating altmetrics. **Scientometrics**, [S. l.], v. 98, n. 2, p. 1131-1143, 2014.

SUN, X.; LI, B.; DING, K.; LIN, Y. Who says what about the most-discussed articles of Altmetric? *In*: INTERNATIONAL CONFERENCE ON SCIENCE AND TECHNOLOGY INDICATORS, 23., 2018. **Proceedings** [...]. Leiden, Holanda: Centre for Science and Technology Studies, 2018. p. 504-509. Disponível em: <https://scholarlypublications.universiteitleiden.nl/access/item%3A2732941/view>. Acesso em: 6 dez. 2021.

SWEPUB. **Sök**. [Suécia]: Kungliga biblioteket, c2018. Disponível em: <https://swepub.kb.se/>. Acesso em: 4 nov. 2021.

TAGUE-SUTCLIFFE, J. An introduction to informetrics. **Information processing & management**, [S. l.], v. 28, n. 1, p. 1-3, 1992. Disponível em: [https://doi.org/10.1016/0306-4573\(92\)90087-G](https://doi.org/10.1016/0306-4573(92)90087-G). Acesso em: 21 out. 2021.

TENNANT, J. Web of Science and scopus are not global databases of knowledge. **European Science Editing**, [S. l.], v. 46, 2020. Disponível em: <https://osf.io/preprints/socarxiv/qhvgr>. Acesso em: 04 jan. 2022.

THELWALL, M. Do Mendeley reader counts indicate the value of arts and humanities research? **Journal of Librarianship and Information Science**, [S. l.], v. 51, n. 3, p. 781-788, 2019.

THELWALL, M. Interpreting correlations between citation counts and other indicators. **Scientometrics**, v. 108, n. 1, p. 337-347, 2016. Disponível em: <https://link.springer.com/article/10.1007/s11192-016-1973-7>. Acesso em: 21 jun. 2022.

THELWALL, M. The Pros and cons of the use of altmetrics in research assessment. **Scholarly Assessment Reports**, [S. l.], v. 2, n. 1, 2020. DOI: <https://doi.org/10.29024/sar.10>

THELWALL, M.; HAUSTEIN, S.; LARIVIÈRE, V.; SUGIMOTO, C. R. Do altmetrics work? Twitter and ten other social web services. **PLoS One**, [S. l.], maio 2013. DOI <https://doi.org/10.1371/journal.pone.0064841>.

THELWALL, M.; SUD, P. Mendeley readership counts: An investigation of temporal and disciplinary differences. **Journal of the Association for Information Science and Technology**, [S. l.], v. 67, n. 12, p. 3036-3050, 2016.

TKACZYK, Dominika. Matchmaker, matchmaker, faça-me uma combinação. **Blog Crossref**, [S. l.], 12 nov. 2018. Disponível em: <https://www.crossref.org/blog/matchmaker-matchmaker-make-me-a-match/>. Acesso em: 4 dez. 2021.

TRZESNIAK, P. A estrutura editorial de um periódico científico. In: SABADINI, A. A. Z. P.; SAMPAIO, M. I. C.; KOLLER, S. H. (org.). **Publicar em psicologia: um enfoque para a revista científica**. São Paulo: ABECiP; IPUSP, 2009. p. 87-102. Disponível em: <http://www.livrosabertos.sibi.usp.br/portaldelivrosUSP/catalog/view/16/12/70>. Acesso em: 4 dez. 2021.

UNIVERSIDADE FEDERAL DE GOIÁS. **Política e diretrizes para os periódicos científicos da UFG**. Goiânia, 2017. Disponível em: https://portal.revistas.ufg.br/revistas_ufg/wp-content/uploads/2018/08/Politica_Diretrizes_Periodicos_UFG-05-1.pdf. Acesso em: 26 fev. 2022.

VAUGHAN, L.; TANG, J.; YANG, R. Investigating disciplinary differences in the relationships between citations and downloads. *Scientometrics*, v. 111, n. 3, p. 1533-1545, 2017. DOI <https://doi.org/10.1108/AJIM-12-2014-0176>.

VERA-BACETA, M. A; THELWALL, M.; KOUSHA, K. Web of Science and Scopus language coverage. **Scientometrics**, [S. l.], v. 121, n. 3, p. 1803-1813, 2019.

VILÁN FILHO, J. L. **Autoria múltipla em artigos de periódicos científicos das áreas de informação no Brasil**. 2010. 215 p. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2010. Disponível em: <https://repositorio.unb.br/handle/10482/7468>. Acesso em: 12 dez. 2021.

VISSER, M.; VAN ECK, N. J.; WALTMAN, L. Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. **Quantitative Science Studies**, v. 2, n. 1, p. 20-41, 2021. Disponível em: https://doi.org/10.1162/qss_a_00112. Acesso em: 14 jul. 2022.

WANG, X.; LIU, C.; FANG, Z.; MAO, W. From attention to citation, what and how does altmetrics work? **arXiv.org**, [S. l.], Sep. 2014. Disponível em: <https://arxiv.org/abs/1409.4269>. Acesso em: 11 dez. 2021.

WATSON, A. B. Comparing citations and downloads for individual articles at the Journal of Vision. **Journal of vision**, [S. l.], v. 9, n. 4, p. 1-4, 2009.

WOUTERS, P.; COSTAS, R. **Users, narcissism and control: tracking the impact of scholarly publications in the 21st century**. Holanda: Leiden University, 2012.

WILLINSKY, J. The nine flavours of open access scholarly publishing. **Journal of Postgraduate Medicine**, [S. l.], v. 49, n. 3, p. 263, 2003.

WHITE, R. K.; ANGELO, A.; FITCHETT, D.; FRASER, M.; HAYES, L.; HOWIE, J.; WHITE, B. Only two out of five articles by New Zealand researchers are free-to-access: a multiple API study of access, citations, cost of Article Processing Charges (APC), and the potential to increase the proportion of open access. **PeerJ**, [S. l.], v. 9, e11417, 2021. Disponível em: <https://peerj.com/articles/11417/>. Acesso em: 04 jan. 2022.

WILSON, V. Research methods: altmetrics. **Evidence Based Library and Information Practice**, [S. l.], v. 11, n. 1, p. 53-55, 2016.

XUE-LI, L.; HONG-LING, F.; MEI-YING, W. Correlation between download and citation and download-citation deviation phenomenon for some papers in Chinese medical journals. **Serials review**, v. 37, n. 3, p. 157-161, 2011. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/00987913.2011.10765377>. Acesso em: 14 jul 2022.

YAN, K. K.; GERSTEIN, M. K. The spread of scientific information: insights from the web usage statistics in PLoS article-level metrics. **PLoS One**, [S. l.], v. 6, n. 5, p. e19917, 2011.

YANG, S.; XING, X.; WOLFRAM, D. Difference in the impact of open-access papers published by China and the USA. **Scientometrics**, [S. l.], v. 115, n. 2, p. 1017-1037, 2018.

YU H.; MURAT, B.; LI, L.; XIAO, T. How accurate are Twitter and Facebook altmetrics data? A comparative content analysis. **Scientometrics**, [S. l.], v. 126, p. 4437–4463, 2021. Disponível em: <https://link.springer.com/article/10.1007/s11192-021-03954-7>. Acesso em: 5 dez. 2021.

ZAHEDI, Z.; COSTAS, R. General discussion of data quality challenges in social media metrics: extensive comparison of four major altmetric data aggregators. **PloS one**, [S. l.], v. 13, n. 5, p. e0197326, 2018.

ZAHEDI, Z.; COSTAS, R.; WOUTERS, P. How well developed are altmetrics? A cross disciplinary analysis of the presence of “alternative metrics” in scientific publications. **Cientometrics**, [S. l.], v. 101; n. 2, p. 1491–1513, 2014. DOI <https://doi.org/10.1007/s11192-014-1264-0>

ZAHEDI, Z.; COSTAS, R.; WOUTERS, P. Mendeley readership as a filtering tool to identify highly cited publications. **J. Ass. Inf. Sci. Tech.**, [S. l.], v. 68, n. 10, p 2511–2521, 2017. Disponível em: <https://dl.acm.org/doi/10.5555/3204593.3204611>. Acesso em: 11 dez. 2021.

ZAHEDI, Z.; HAUSTEIN, S. On the relationships between bibliographic characteristics of scientific documents and citation and Mendeley readership counts: A large-scale analysis of Web of Science publications. **Journal of Informetrics**, [S. l.], v. 12, n. 1, p. 191-202, 2018.

ZAHEDI, Z.; VAN ECK, N. J. P. Exploring topics of interest of Mendeley users. **Journal of altmetrics**, v. 1, n. 1, p. 5, 2018. Disponível em: <https://doi.org/10.29024/joa.7>. Acesso em: 19 nov. 2022.

ZIMAN, J. M. **An introduction to science studies**: the philosophical and social aspects of science and technology. United Kingdom: Cambridge University Press, 1987.