

COBERTURA DE FONTES DE DADOS E CORRELAÇÕES ENTRE INDICADORES ALTMÉTRICOS E CITAÇÕES EM PORTAIS DE PERIÓDICOS BRASILEIROS: COMPARAÇÕES NAS DIFERENTES ÁREAS DO CONHECIMENTO

COVERAGE OF DATA SOURCES AND CORRELATIONS BETWEEN ALTMETRIC INDICATORS AND CITATIONS FOR BRAZILIAN JOURNAL PORTALS: COMPARISON OF DIFFERENT KNOWLEDGE AREAS

Érika Demachki^a
João de Melo Maricato^b

RESUMO

Objetivo: contribuir com as análises de coberturas de fontes de dados de indicadores altmétricos e de citação e das correlações entre eles, nas diferentes áreas de conhecimento. **Metodologia:** foram analisados indicadores de 51.200 artigos de quatro portais de periódicos, divididos em cinco áreas de conhecimento. **Resultados:** os acessos ao resumo e as leituras/capturas no *Mendeley* apresentaram maiores coberturas, embora haja grandes diferenças entre as áreas. A fonte de dados de citação de maior cobertura foi a *Crossref*, variando de 7,3 a 43,8%, dependendo da área. As correlações entre indicadores altmétricos e de citação também apresentam variações entre as áreas (entre fraca e moderada). A Ciências da Vida apresentou as correlações mais fortes, com destaque para as entre leitores no *Mendeley* e citações, variando de 0,432 a 0,491. **Conclusões:** conclui-se que o *Mendeley* (dados altmétricos) e a *Crossref* (dados de citação) podem ser consideradas fontes de dados alternativas e/ou complementares para a avaliação da ciência brasileira em todas as áreas do conhecimento, embora haja diferenças entre elas. Os dados sugerem, por fim, a importância das revistas científicas de acesso aberto e dos portais de periódicos de universidades para a divulgação da ciência produzida no País.

Descritores: Comunicação científica. Estudos bibliométricos. Estudos altmétricos.

^a Mestre em Ciências da Informação pela Universidade de Brasília (UnB). Técnico-administrativo em Educação da Universidade Federal de Goiás (UFG), Goiânia, Brasil. E-mail: erikademachki@gmail.com

^b Doutor em Ciência da Informação pela Universidade de São Paulo (USP). Docente da Faculdade de Ciência da Informação da Universidade de Brasília (UnB), Brasília, Brasil. E-mail: jmmaricato@gmail.com

Produção científica.

1 INTRODUÇÃO

A avaliação da atividade científica não é algo recente, sendo comumente realizada com o uso de indicadores obtidos a partir das produções científicas, sobretudo revistas e artigos científicos. Tradicionalmente, indicadores bibliométricos baseados em citação vêm sendo utilizados para sugerir qualidade a periódicos científicos e seus artigos, bem como para medir impactos da ciência. Com o surgimento da altmetria, que tem, como uma das vantagens, a de poder medir o impacto da produção científica em diversas fontes de dados – como as redes e mídias sociais –, muitos estudos desenvolvidos nos últimos anos apontam para um potencial de uso conjunto/complementar de indicadores baseados em citações e indicadores alométricos (MOHAMMADI *et al.*, 2015; MARICATO; MARTINS, 2017; HUANG; WANG; WU, 2018; SUN *et al.*, 2018; OUCHI *et al.*, 2019).

Analisar as correlações entre indicadores alométricos e de citação contribui para o debate sobre os níveis de complementaridade, de sobreposição e de relação causa-efeito entre os indicadores de citação e alométricos, bem como sobre a capacidade desses indicadores, baseados na web 2.0, em captar novas dimensões do impacto científico. Diversos são os estudos que correlacionam indicadores alométricos e bibliométricos de citação e que argumentam de maneira semelhante, como, por exemplo, os de Haustein *et al.*, (2014); Schlögl *et al.* (2014); Thelwall (2019); Chen *et al.* (2020); e Mohammadi e Thelwall (2014).

Indicadores alométricos – sejam eles de *downloads*, acessos, leitores, menções e outros – e indicadores bibliométricos de citação captam aspectos (diferentes e semelhantes) do impacto de publicações científicas, porém, em níveis ainda não totalmente compreendidos. Para Watson (2009) e Wang *et al.* (2014) os indicadores alométricos de acesso e de *download* podem ajudar a prever citações, a depender da força da correlação entre estes indicadores. É necessário ressaltar ainda a capacidade dos indicadores alométricos em captar

informações que os de citação não conseguem, a exemplo dos indicadores advindos das redes sociais, que ocorrem logo após a divulgação da produção nessas mídias (KEELE, 2016).

Já o estudo da cobertura das fontes de dados – bibliométricos e altmétricos – auxilia no entendimento acerca de quais delas reúnem mais informações quantitativas das publicações, que estão sendo mais utilizadas e/ou mencionadas pela comunidade científica e pela sociedade em geral. Diversos estudos analisaram as coberturas de indicadores altmétricos e bibliométricos de citação, entre eles: Alperin (2015); Haustein e Larivière (2014); Mohammadi *et al.* (2015); Haustein, Costas e Larivière (2015); Zahedi, Costas e Wouters (2017); Zahedi e Haustein (2018). Portanto, as análises propostas na pesquisa que ora se apresenta consideram a cobertura das fontes de dados e as correlações entre os indicadores altmétricos e de citação provindos destas fontes, o que permitiu observar duas dimensões relevantes de análise.

Desse modo, por um lado é possível compreender melhor sobre as complementariedades entre os indicadores altmétricos e bibliométricos e suas possibilidades de uso para avaliação em diferentes contextos. Por outro, a análise das coberturas propicia a possibilidade de conhecer as potencialidades das bases e fontes de dados para monitoramento da ciência em diferentes perspectivas – a exemplo de países e áreas do conhecimento. As tentativas de compreensão de coberturas e/ou correlações entre indicadores altmétricos e de citação, sob diferentes prismas, têm sido amplamente discutidas na literatura (LI; THELWALL; GIUSTINI, 2012; THELWALL *et al.*, 2013; HAUSTEIN; LARIVIÈRE, 2014; MOHAMMADI *et al.*, 2015).

A análise por área do conhecimento permite identificar tais dinâmicas (correlações e coberturas) dos indicadores altmétricos e de citação, fomentando discussões sobre possíveis relações com as práticas de comunicação científica de cada área. Ou seja, é relevante discutir as potencialidades e limitações destes indicadores e fontes em nível de área de conhecimento. Estudos como o de Mohammadi e Thelwall (2013), Haustein e Larivière (2014) e Mohammadi *et al.* (2015) descobriram correlações mais fortes para publicações nas áreas de Saúde, como Ciências Médicas e Biologia; e ciências duras, como Engenharia e

Bioquímica. Brigham (2014) salienta que algumas disciplinas são mais ativas *online* do que outras e que o engajamento pode variar de acordo com a plataforma de mídia social utilizada. Diante disso, ao analisar indicadores altmétricos, especialmente quando comparados com indicadores de citação, deve-se considerar a área de conhecimento das publicações (BRIGHAM, 2014).

Quanto aos indicadores altmétricos, estes podem ser extraídos de fontes como as plataformas agregadoras de dados altmétricos – por exemplo: *Altmetric* e *PlumX*; de redes sociais como *Twitter* e *Facebook*; e do gerenciador de referências *Mendeley*. Pesquisas anteriores demonstram que o *Mendeley* e o *Twitter* possuem as maiores coberturas altmétricas (COSTAS; ZAHEDI; WOUTERS, 2014; ROBINSON-GARCÍA *et al.*, 2014; HAUSTEIN; COSTAS; LARIVIÈRE, 2015). Indicadores altmétricos de uso, tais como os de acesso aos resumos e de *download* de artigos, podem ser extraídos diretamente das plataformas que hospedam as publicações científicas, como o *software* de editoração científica *Open Journal Systems* (OJS).

Os indicadores baseados em citações são, na maioria das vezes, extraídos de bases de dados bibliográficas como *Web of Science* (WoS) e *Scopus*, sendo estas as duas mais influentes (CÉSPEDES, 2021). Entretanto, estudos anteriores (MUGNAINI *et al.*, 2019; MELO; TRINCA; MARICATO, 2021) evidenciam uma baixa cobertura de indicadores de citação, nessas bases de dados, para artigos brasileiros em diferentes áreas de conhecimento, indicando a necessidade de se buscar fontes mais abrangentes para a produção de indicadores para países que não fazem parte do *mainstream* científico.

A ciência brasileira é fortemente publicada nas revistas nacionais (PACKER, 2011). Estas contam com uma forte política de publicação em acesso aberto, o que é considerado um fator de aumento da visibilidade científica (AHMED, 2007; DEMETER; ISTRATII, 2020). A gestão dos periódicos científicos brasileiros está majoritariamente vinculada a instituições de ensino superior e seus programas de pós-graduação, havendo importante papel destes no desenvolvimento da ciência brasileira. Assim, conhecer as coberturas dos artigos dessas revistas, em diferentes fontes de dados, e as correlações entre os indicadores altmétricos e de citação, proporciona maior conhecimento sobre

as possibilidades de se monitorar a ciência desenvolvida no País em relação às diferentes áreas do conhecimento. Além disso, tal abordagem estimula discussões sobre o papel dos portais de periódicos de acesso aberto na produção, organização e divulgação do conhecimento científico e a importância desse tipo de dispositivo para o desenvolvimento da ciência nacional.

Diante do exposto, a pesquisa aqui apresentada, teve, como objetivo, compreender as coberturas de fontes de dados e correlações entre indicadores altmétricos e bibliométricos dos artigos dos portais de periódicos brasileiros nas diferentes áreas do conhecimento com vistas a responder os seguintes questionamentos: os indicadores, altmétricos e bibliométricos, medem fenômenos diferentes ou são complementares? As fontes de dados analisadas têm potencial para monitorar a ciência brasileira? Quais as características de cobertura das áreas do conhecimento e quais áreas podem ser mais beneficiadas com os indicadores?

2 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa é exploratória-descritiva com abordagem metodológica quantitativa. Foram produzidos e analisados indicadores, altmétricos e de citação, de artigos indexados em portais de periódicos brasileiros de acesso aberto. Realizou-se análises de suas correlações e as coberturas em fontes e bases de dados. A pesquisa envolveu procedimentos de levantamento, classificação, tabulação e análise dos dados, conforme as etapas descritas a seguir.

A seleção dos portais de periódicos brasileiros para compor a pesquisa foi feita a partir do *Directory of Open Access Journals* (DOAJ). Foram enviados e-mails para os dez portais com maior número de revistas indexadas no diretório convidando-os a participar da pesquisa. Os portais de periódicos que aceitaram participar da pesquisa, disponibilizando-se a fornecer os dados de acessos e *downloads* solicitados, foram os das Universidades Federais de Santa Catarina (UFSC), de Goiás (UFG), do Rio Grande do Norte (UFRN) e de Brasília (UnB) (UNIVERSIDADE FEDERAL DE SANTA CATARINA, 2021; UNIVERSIDADE FEDERAL DE GOIÁS, 2021; UNIVERSIDADE FEDERAL DO RIO GRANDE DO

NORTE, 2021; UNIVERSIDADE DE BRASÍLIA, 2021).

Desses portais foram selecionadas apenas as revistas que têm ISSN e artigos com *Digital Object Identifier* (DOI). Para identificação das revistas foi consultada a REST API do *Crossref* (2020). Das 205 revistas com ISSN hospedadas nos portais de periódicos supracitados, 153 possuíam artigos com DOI registrados em ao menos um de seus números publicados.

A partir das 153 revistas, foram identificados o total de 52.354 publicações com DOI. Posteriormente foram excluídos textos que não são artigos, tais como Apresentação, Editorial, Expediente e Normas para publicação. Essa etapa de tratamento garantiu que o corpus seja composto exclusivamente por artigos científicos. A amostra final resultou em 51.200 artigos científicos dentre as 153 revistas que registram DOI para suas publicações.

As revistas foram categorizadas, manualmente, pelas suas áreas de conhecimento principais: Ciências da Vida (CV); Ciências Humanas (CH); Ciências Sociais Aplicadas (CSA); Linguística, Letras e Artes (LLA); e Ciências Exatas, Tecnológicas e Multidisciplinares (CETM). Para este agrupamento considerou-se a afinidade das revistas com base nos colégios e nas grandes áreas de conhecimento definidas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

A coleta dos dados altmétricos e bibliométricos, nas diferentes fontes de dados, foi feita entre os dias 21 de fevereiro e 2 de março de 2022. As fontes altmétricas selecionadas foram *Mendeley*, *Twitter* e *Facebook*. Os indicadores destas fontes foram coletados da plataforma *Altmetric* utilizando-se a ferramenta *Odisseia Metrics* (RAMOS; MARICATO, 2020), um *script* que consulta a *Application Programming Interface* (API) pública da plataforma a partir da busca pelos números de DOI das publicações.

As fontes de coleta de dados bibliométricos utilizadas foram *WoS*, *Scopus* e *Crossref*. A coleta dos dados de citação na *WoS* e na *Scopus* foi feita através da interface de busca avançada dos respectivos sites, utilizando-se as listas de números de DOI das publicações. Para a coleta de citações da *Crossref* utilizou-se a ferramenta *Odisseia Metrics* a partir da lista de DOIs dos artigos. A coleta dos indicadores altmétricos de acessos ao resumo e o número de *downloads*

dos artigos foi feita manualmente pelos gestores dos portais de periódicos a partir da interface do sistema OJS utilizando-se o gerador de relatórios nativo do sistema.

Os indicadores extraídos foram analisados a partir de duas perspectivas: coberturas e correlações. Como cobertura entende-se a ocorrência de ao menos uma menção ou citação dos artigos nas diferentes fontes, sendo produzidos indicadores em números absolutos (quantidade de artigos que foram mencionados ou citados ao menos uma vez) e sua respectiva porcentagem. A partir dos quantitativos de ocorrências de coberturas também foram calculadas as médias e medianas. Ou seja, foram calculadas médias e medianas dos artigos em relação aos quantitativos de menções ou citações em cada uma das fontes de dados.

Em relação aos indicadores de correlação, estes são calculados a partir do quantitativo de ocorrências de cada artigo (de citação, altmétricos e entre eles) utilizando-se a correlação de Spearman. Foi utilizado o *software* livre de testes estatísticos Jamovi¹ para executar os testes *Kurtosis* e *Skewness*, que indicaram uma distribuição anormal (alto desvio padrão) para todas as variáveis da amostra, confirmando que a utilização da correlação de Spearman era mais adequada do que a correlação de Pearson. Para a interpretação e análise do coeficiente de correlação de Spearman foram considerados os valores de r positivos ou negativos propostos por Silva Filho (2022): 0 – 0,2 (correlação nula); 0,21 – 0,4 (correlação fraca); 0,41 – 0,7 (correlação moderada); 0,71 – 0,9 (correlação forte); e 1 (correlação perfeita).

3 RESULTADOS E DISCUSSÕES

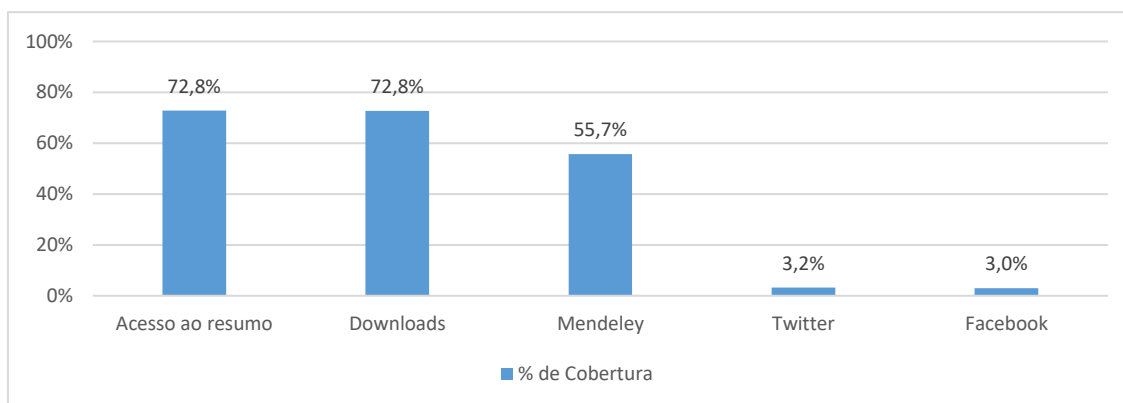
3.1 Coberturas de dados altmétricos

No Gráfico 1 são apresentadas as coberturas por fonte de dados altmétricas para todos os 51.200 artigos analisados. Em 72,8% dos artigos há algum acesso ao resumo e/ou *download*. Mais da metade dos artigos (55,7%),

¹ *Jamovi Stats Open Now*. 2021. Disponível em: <https://www.jamovi.org/>. Acesso em: 10 dez. 2021.

equivalente a 28.518, possuem algum leitor no *Mendeley*; enquanto 3% têm alguma menção no *Facebook* e 3,2% no *Twitter*. Estas coberturas indicam que, entre as mídias sociais, o *Mendeley* é a fonte altmétrica que possui a maior cobertura para a totalidade de artigos.

Gráfico 1 - Coberturas por fontes de dados altmétricas para a totalidade de artigos dos portais de periódicos



Fonte: Elaborado pelos autores (2022).

Na Tabela 1 são apresentadas as menções e coberturas dos artigos nas fontes de dados altmétricas por área de conhecimento. Observa-se que a área de Ciências Humanas possui a maior quantidade de artigos cobertos, seguida pelas de Ciências da Vida, Linguística, Letras e Artes, Ciências Sociais Aplicadas e Ciências Exatas, Tecnológicas e Multidisciplinares. Percebe-se que as coberturas variam entre as áreas de conhecimento. Isso ocorre, provavelmente, devido a fatores como a influência das práticas de comunicação em cada área de conhecimento (MUELLER, 2005); às características particulares das fontes de dados e plataformas de mídia social, bem como pelos diferentes perfis de usuários (MOHAMMADI *et al.*, 2015).

A Ciências Humanas é a área de conhecimento que apresentou maior número de menções totais em fontes de dados altmétricas (15.374.177), sendo a maior parte delas acessos ao resumo (81,8%) e *downloads* (80,4%). Isso sugere que há uma diferença nos padrões de *downloads* entre usuários de diferentes áreas de conhecimento, conforme destacado em outros estudos (SCHLÖGL *et al.*, 2014).

A Ciências Exatas, Tecnológicas e Multidisciplinar é a área de

conhecimento com o menor número de artigos, mas que possui a maior cobertura total de menções entre todas as áreas: 93,7% dos artigos têm ao menos uma menção em alguma das fontes de dados altmétricas. A cobertura total de menções indica que as fontes de dados altmétricas para a área de Ciências Exatas, Tecnológicas e Multidisciplinares têm a melhor complementaridade entre as áreas de conhecimento.

Dos 10.832 artigos de Ciências da Vida, 10.081 possuem ao menos uma menção em uma das fontes de dados altmétricas consideradas, equivalente a 93% do total.

Tabela 1 - Menções, coberturas, média e mediana de menções dos artigos por área de conhecimento dos portais de periódicos nas fontes altmétricas

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	82,4%	10	5
<i>Twitter</i>	458	300	2,8%	0	0
Acesso resumo	2.457.822	6.131	56,6%	401	209
<i>Downloads</i>	3.393.653	6.095	56,3%	554	243
<i>Facebook</i>	569	547	5%	0	0
Total**	5.959.032	10.081	93%	-	-
Ciências Humanas (CH) n = 19001					
<i>Mendeley</i>	56.197	9.122	48%	3	0
<i>Twitter</i>	2.174	729	3,8%	0,1	0
Acesso resumo	7.073.795	15.549	81,8%	455	205
<i>Downloads</i>	8.241.713	15.282	80,4%	530	214
<i>Facebook</i>	298	285	1,5%	0	0
Total**	15.374.177	17.540	92,3%	-	-
Ciências Sociais Aplicadas (CSA) n = 8316					
<i>Mendeley</i>	43.543	5.150	61,9%	5,2	2
<i>Twitter</i>	1.192	400	4,8%	0,1	0
Acesso resumo	2.876.059	5.963	71,7%	482	262
<i>Downloads</i>	3.225.760	5.907	71%	541	255
<i>Facebook</i>	371	300	3,6%	0	0
Total**	6.146.925	7.524	90,5%	-	-
Linguística, Letras e Artes (LLA) n = 9183					
<i>Mendeley</i>	14.735	3.117	33,9%	1,6	0
<i>Twitter</i>	382	173	1,9%	0	0
Acesso resumo	2.247.264	6.323	68,9%	355	207
<i>Downloads</i>	2.726.989	6.323	68,9%	431	210
<i>Facebook</i>	568	432	4,7%	0	0
Total**	4.989.938	7.522	81,9%	-	-
Ciências Exatas, Tecnológicas e Multidisciplinares (CETM) n = 3868					
<i>Mendeley</i>	15.096	2.221	57,4%	4	1
<i>Twitter</i>	70	47	1,2%	0	0
Acesso resumo	1.977.205	3.291	85,1%	601	370
<i>Downloads</i>	2.662.032	3.245	83,9%	809	376
<i>Facebook</i>	8	8	0,2%	0	0
Total	4.654.411	3.623	93,7%	-	-

Fonte: Elaborado pelos autores (2022).

O *Mendeley* se destaca na área de Ciências da Vida não apenas pela maior cobertura de leitores entre as áreas, de 82,4%, mas também pela média de 10 leitores por artigo. A cobertura de leitores no *Mendeley* para a área de Ciências da Vida está semelhante às encontradas por Haustein e Larivière (2014), variando de 62,8% a 81% para artigos de quatro subáreas, com uma média de 9,6 leitores por artigo. O fato de a área de Ciências da Vida 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, 2017) também identificaram uma maior cobertura para a área de Ciências da Vida e Medicina, com 73% e 93,1%, respectivamente. 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 Ciências Humanas, Ciências Sociais Aplicadas e Ciências Exatas, Tecnológicas e Multidisciplinar tiveram coberturas intermediárias de leitores no *Mendeley* – de 48%, 61,9% e 57,4%, respectivamente; similar às coberturas das áreas correspondentes analisadas por Hammarfelt (2014), Zahedi, Costas e Wouters (2014), que variam de 51% a 68%. A baixa cobertura de 33,9% de leitores no *Mendeley* para a área de Linguística, Letras e Artes também é próxima à cobertura de 39% encontrada por Zahedi, Costas e Wouters (2014).

Quanto ao *Twitter*, a área com a maior cobertura é a de Ciências Sociais Aplicadas, com 4,8%. E a de menor cobertura é a de Ciências Exatas, Tecnológicas e Multidisciplinar, com 1,2% dos artigos mencionados no *Twitter*. Essas coberturas são inferiores a todas as coberturas encontradas por Costas, Zahedi e Wouters (2014) e Haustein, Costas e Larivière (2015), ainda que, comparativamente com as demais áreas do conhecimento abordadas nestes estudos, as áreas relacionadas às Ciências Exatas e Engenharias tenham as menores coberturas de *tweets* entre as áreas.

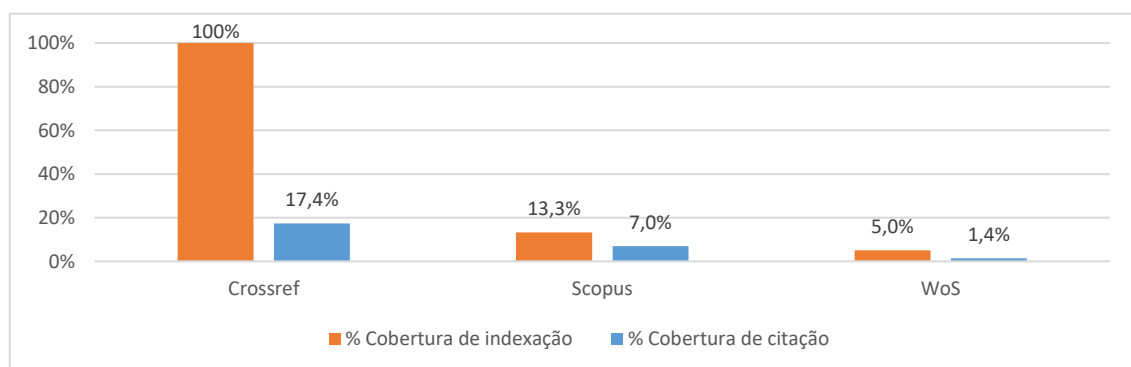
As coberturas no *Facebook* variam de 0,2% a 5%, sendo a maior

cobertura para a área de Ciências da Vida e a menor para as Ciências Exatas, Tecnológicas e Multidisciplinares. Haustein, Costas e Larivière (2015) também encontraram as menores coberturas no *Facebook* para as áreas equivalentes, 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. Isso pode ser explicado por fatores como: baixo uso dos artigos em geral, refletido no uso das mídias sociais; o uso dessas mídias sociais no Brasil ser baixo entre acadêmicos; ou, ainda, por haver uma cultura diferente no compartilhamento de pesquisas nas mídias sociais no Brasil (ALPERIN, 2015).

3.2 Coberturas de dados de citação

No Gráfico 2 são apresentadas as coberturas de citação e indexação nas fontes de dados bibliométricas para a totalidade de artigos analisados. Uma vez que todos os artigos estão registrados na *Crossref* – fonte que fornece o número de DOI, a cobertura de indexação nesta fonte é de 100% para os 51.200 artigos. As demais fontes, *Scopus* e *Web of Science* (WoS), indexam, respectivamente, apenas 13,3% e 5% dos artigos. Quanto à cobertura de citação, a *Crossref* também é a maior, com 17,4% dos artigos (equivalente a 8.908) com alguma citação nesta fonte. A menor cobertura de citação é para a WoS, com 1,4%.

Gráfico 2 - Coberturas de citação e indexação, por fontes de dados bibliométricas, para a totalidade de artigos dos portais de periódicos



Fonte: Elaborado pelos autores (2022).

Na Tabela 2 são apresentadas as coberturas de citação e indexação, assim como médias e medianas de citações para os artigos por área de conhecimento. A área de Ciências da Vida se destaca com o maior volume total de citações: 34.853 citações recebidas. Enquanto 49,1% dos artigos de Ciências da Vida possuem pelo menos uma citação nas três fontes de dados bibliométricas, as demais áreas possuem uma cobertura menor, que varia de 9,1% a 14,8%. O que indica que mais artigos da Ciências da Vida estão sendo citados do que de outras áreas do conhecimento.

Tabela 2 - Citações, coberturas de citação e de indexação, médias e medianas de citações para artigos dos portais por área de conhecimento nas 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	43,8%	10.832	100%	43,8%	1,6	0
<i>Scopus</i>	15.168	2.824	26%	3.838	35,4%	73,5%	3,9	2
WoS	2.200	431	34%	576	5,3%	74,8%	3,8	2
Total**	34.853	5.317	49,1%	10.832	100%	-	-	-
Ciências Humanas (CH) n = 19001								
<i>Crossref</i>	3.830	2.060	10,8%	19.001	100%	10,8%	0,2	0
<i>Scopus</i>	726	401	2,1%	1.538	8%	26%	0,4	0
WoS	126	75	0,4%	472	2,4%	15,8%	0,2	0
Total**	4.682	2.406	12,7%	19.001	100%	-	-	-
Ciências Sociais Aplicadas (CSA) n = 8316								
<i>Crossref</i>	2.158	1.053	12,6%	8.316	100%	12,6%	0,2	0
<i>Scopus</i>	267	133	1,6%	486	5,8%	27,3%	0,5	0
WoS	140	90	1%	520	6,2%	17,3%	0,2	0
Total**	2.565	1.228	14,8%	8.316	100%	-	-	-
Linguística, Letras e Artes (LLA) n = 9183								
<i>Crossref</i>	1.075	672	7,3%	9.183	100%	7,3%	0,1	0
<i>Scopus</i>	448	202	2,2%	906	9,8%	22,3%	0,5	0
WoS	219	142	1,5%	978	10,6%	14,5%	0,2	0
Total**	1.742	836	9,1%	9.183	100%	-	-	-
Ciências Exatas, Tecnológicas e Multidisciplinares (CETM) n = 3868								
<i>Crossref</i>	634	357	9,2%	3.868	100%	9,2%	1,7	1
<i>Scopus</i>	6	5	0,1%	19	0,5%	26,3%	0,3	0
WoS	0	0	0%	0	0%	0%	0	0
Total**	640	362	9,4%	3.868	100%	-	-	-

Legenda: **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.

Fonte: Elaborado pelos autores (2022).

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* (WoS). Na fonte bibliométrica WoS os autores Haustein, Costas e Larivière

(2015) identificaram coberturas de citação variando entre 45,8% e 74,3%, muito acima das coberturas obtidas no presente estudo, que variaram de 0,4% a 4% entre as áreas. As coberturas de indexação das publicações na WoS também são baixas, variando de 0% para a área de Ciências Exatas, Tecnológicas e Multidisciplinar, a 10,7% para Linguística, Letras e Artes. Este resultado confirma a sub-representatividade da produção científica brasileira em bases de dados internacionais apontada por Mugnaini *et al.* (2019) e também por Melo, Trinca e Maricato (2021).

Na área de Ciências da Vida, que possui as maiores coberturas em todas as fontes de dados bibliométricas, 43,8% dos artigos possuem alguma citação na *Crossref*. O que representa uma cobertura próxima à de 35,9% encontrada por Demachki e Maricato (2022), para artigos desta mesma área, em 32 revistas do Portal de Periódicos da UFG. Essa “[...] 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, 2022).

Ainda que a cobertura de indexação para a área de Ciências da Vida na WoS não seja a maior entre as áreas, a sua média de citações recebidas para artigos indexados é a maior, com 3,8 citações por artigo, e mediana 2. A segunda maior cobertura de citação entre artigos indexados de Ciências da Vida é na *Scopus*, com 73,6%; e, por fim, na *Crossref*, com 43,8%. A média de citações recebidas na *Scopus* para artigos de Ciências da Vida é 3,9, ligeiramente maior do que a média para a WoS. Isso indica que, apesar da cobertura de citação para artigos indexados na WoS ser um pouco maior do que na *Scopus*, estas duas bases estão captando algumas citações diferentes. Diferentemente da descoberta ora apresentada, Visser, Van Eck e Waltman (2021) identificaram que a sobreposição de artigos entre a WoS e a *Scopus* foi a menor entre as fontes analisadas pelos autores – entre WoS, *Crossref*, *Dimensions* e *Microsoft Academic*. Essa diferença é intrigante, sendo necessária a realização de pesquisas com outras abordagens metodológicas para entendê-las melhor.

A *Crossref* possui, naturalmente, uma cobertura de indexação total para toda a amostra de artigos estudados, superando a *Scopus* e a WoS. Essa

cobertura se deve ao fato de todos os artigos possuírem número de DOI, o que implica no registro de todos eles na *Crossref*. Em um comparativo entre produções científicas de 20 países nas fontes de dados bibliométricas *Scopus*, *WoS* e *Dimensions*, Singh *et al.* (2021) identificaram que, para 90% dos países, a fonte que indexa um maior volume de publicações é a *Dimensions*, fonte que depende fortemente dos dados da *Crossref*. Isso porque essas fontes cobrem dados muito semelhantes (VISSER; VAN ECK; WALTMAN, 2021). Portanto, é natural esperar que a *Crossref* supere, em cobertura de indexação, as fontes de dados *Scopus* e *WoS*, visto que estas bases possuem políticas de indexação restritivas.

Ao comparar as coberturas altmétricas e bibliométricas, a área de Ciências da Vida possui as maiores no *Mendeley* (82,3%) e citações em todas as fontes (49,1%). Observou-se que a maior diferença entre cobertura de leitores no *Mendeley* e citações totais é para a área de Linguística, Letras e Artes (24,8%), seguida pela Ciências da Vida (33,2%). Já a área de Ciências Exatas, Tecnológicas e Multidisciplinar possui a maior diferença (48%) entre as coberturas no *Mendeley* (57,4%) e de citações totais (9,3%). Zahedi, Costas e Wouters (2014) afirmam que áreas com mais leitores ou com equilíbrio proporcional entre leitores e citações podem se beneficiar do uso do *Mendeley*, em termos de ter mais impacto de leitura do que impacto de citação.

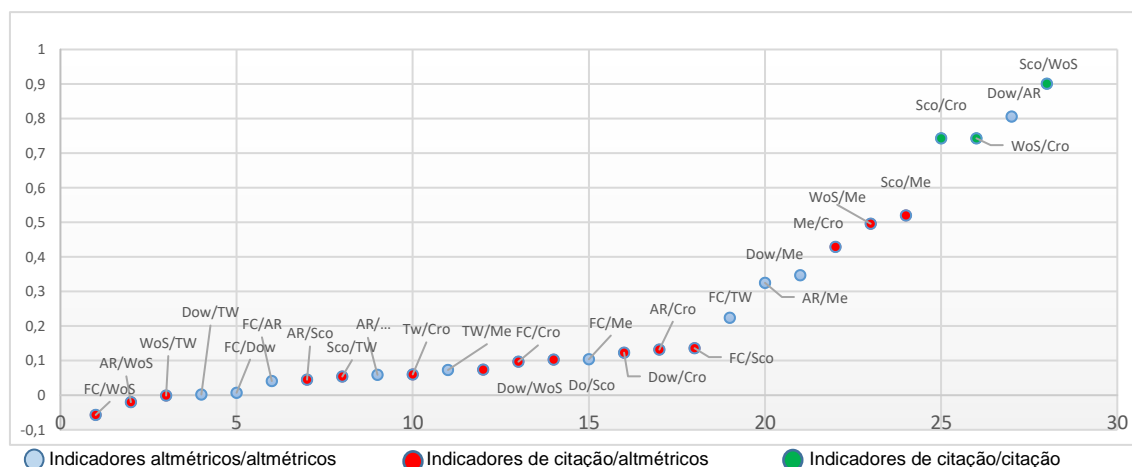
Tendo em vista as diferenças entre as coberturas nas áreas do conhecimento, fica evidente que a área de Ciências da Vida possui as maiores coberturas de leitores no *Mendeley*, menções no *Facebook* e citações nas três fontes bibliométricas. Enquanto isso, a área de Ciências Exatas, Tecnológicas e Multidisciplinar possui as menores coberturas de indicadores altmétricos do *Twitter* e *Facebook*, bem como de indicadores de citações na *WoS* e na *Scopus*, em contraste com a maior cobertura de acessos ao resumo e *downloads*.

3.3 Correlações entre os indicadores

O Gráfico 3 tem o objetivo de ilustrar os dados de correlação da totalidade de artigos de uma maneira mais amigável. Nesta se pode observar as correlações entre cada um dos indicadores de citação, altmétricos e entre eles.

Destaca-se, no Gráfico 3, tais indicadores com os tipos de dados correlacionados em diferentes cores: entre indicadores de citação e altmétricos – pontos em vermelho; entre indicadores de citação – pontos em verde; e, entre indicadores altmétricos – pontos em azul.

Gráfico 3 - Correlações entre indicadores altmétricos e bibliométricos para a totalidade de artigos do portal



Legenda: AR = Acesso ao resumo; Dow = Download; FC = Facebook; TW = Twitter; Me = Mendeley; WoS = Web of Science; Sco = Scopus; Cro = Crossref.

Fonte: Elaborado pelos autores (2022).

Entre indicadores altmétricos a maior correlação é entre acessos ao resumo e *downloads*; e a menor, entre *downloads* e *tweets*. Entre indicadores de citação e altmétricos as correlações de maior força são moderadas entre as citações nas três bases e os leitores no *Mendeley*, com destaque para as citações na *Scopus*. Estas correlações indicam o *Mendeley* como a mais promissora fonte de dados altmétrica, entre as analisadas, para complementar citações. As demais correlações entre citações e indicadores altmétricos variam de fracas a quase nulas. Entre os indicadores de citação a correlação mais forte de todas está entre citações na *Scopus* e na *WoS*. De fato, as correlações entre as citações são fortes para todas as fontes, o que indica que as citações nestas bases são, em algum nível, semelhantes.

Ao analisar os indicadores por área de conhecimento, dispostas na Tabela 3, pode-se observar correlações moderadas para a área de Ciências da Vida entre leitores no *Mendeley* e citações nas três fontes bibliométricas, com o coeficiente de Spearman variando entre 0,432, para citações na *Scopus*, até

0,491 para citações na WoS. Haustein *et al.* (2014) também obtiveram correlações moderadas entre leitores no *Mendeley* e citações na WoS de 0,530, 0,448 e 0,439 para áreas de Pesquisa Biomédica, Biologia e Medicina, respectivamente; assim como as correlações moderadas de 0,54 a 0,67 que Thelwall e Sud (2016) obtiveram para citações na *Scopus* para diversas subáreas de Ciências da Vida. Sendo assim, a força das correlações obtidas no presente estudo, para a área de Ciências da Vida está, em consonância com resultados anteriores. Correlações onde não houve indicadores em comum para um mesmo artigo não puderam ser calculadas (erro expresso na tabela como NaN), e correlações onde a amostra de artigos é pequena não foram analisadas (expressas na tabela como -).

Tabela 3 - Correlações entre indicadores altmétricos e bibliométricos para os artigos dos portais por área de conhecimento, dados calculados utilizando o software Jamovi

Ciências da Vida								
	<i>Crossref</i>	<i>Mendeley</i>	<i>Twitter</i>	<i>WoS</i>	<i>Scopus</i>	Acesso Resumo	Downloads	Facebook
<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	—	
Facebook	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	Facebook
<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	—	
Facebook	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	Facebook
<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	—	
Facebook	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>	<i>Acesso Resumo</i>	<i>Downloads</i>	<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	—			
<i>Acesso Resumo</i>	0.152	0.190	0.084	0.185	0.077	—		
<i>Downloads</i>	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>	<i>Acesso Resumo</i>	<i>Downloads</i>	<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	-	NaN	NaN	—			
<i>Acesso Resumo</i>	0.203	0.514	0.085	NaN	-	—		
<i>Downloads</i>	0.218	0.581	0.046	NaN	-	0.793	—	
<i>Facebook</i>	0.047	0.009	0.150	NaN	NaN	0.036	0.040	—

Fonte: Elaborado pelos autores (2022).

As correlações entre leitores no *Mendeley* e citações na *WoS* são fracas para as áreas de Linguística, Letras e Artes (0,250), Ciências Humanas (0,251) e Ciências Sociais Aplicadas (0,285), assim como em estudos anteriores, para estas mesmas fontes de dados (HAUSTEIN *et al.*, 2014; MOHAMMADI; THELWALL, 2014; THELWALL, 2019). Uma vez que a área de Ciências Exatas, Tecnológicas e Multidisciplinar não possui nenhuma revista indexada na *WoS*, a correlação para esta fonte não foi calculada.

Considerando que a *Crossref* possui as maiores coberturas de citação para as áreas de conhecimento, a análise das correlações entre o número de citações provenientes desta base com os leitores do *Mendeley* representa melhor o cenário da influência entre estes indicadores por área. Sendo assim, a única correlação de força moderada entre leitores no *Mendeley* e citações na *Crossref* é na área de Ciências da Vida (0,445), com as demais correlações fracas, variando entre 0,254 e 0,316. Assim como a correlação entre leitores no *Mendeley* e citações na *WoS*, a correlação entre leitores no *Mendeley* e citações na *Crossref* para a área de Linguística, Letras e Artes também é a mais fraca entre as áreas (0,254). Já entre leitores no *Mendeley* e citações na *Scopus*, a correlação se mantém fraca (0,391), mas é superior à de outras áreas, como Ciências Humanas (0,140) e Ciências Sociais Aplicadas (0,006). Áreas com

correlações mais fracas entre leitores e citações podem indicar que os leitores não são, necessariamente, pesquisadores com interesse em citar o artigo lido. 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.

Entre os indicadores altmétricos a correlação entre leitores no *Mendeley* e acessos ao resumo é moderada para as áreas de Ciências Sociais Aplicadas (0,527) e Ciências Exatas, Tecnológicas e Multidisciplinares (0,514), enquanto que nas demais áreas a correlação é fraca ou nula. Quanto aos *downloads*, a correlação com os leitores no *Mendeley* é moderada para Ciências Sociais Aplicadas (0,489) e Ciências Exatas, Tecnológicas e Multidisciplinares (0,581). Comparativamente, o estudo de Schlögl *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 Ciências Sociais Aplicadas; e correlações menores entre leitores e citações na *Scopus* (0,51 e 0,59).

O *Facebook* apresenta correlações nulas com todos os indicadores altmétricos e bibliométricos, exceto quando se trata do *Twitter*. Para a área de Ciências Sociais Aplicadas o *Facebook* e o *Twitter* se correlacionam moderadamente (0,411). Para as áreas de Ciências da Vida e Linguística, Letras e Artes as correlações entre estes indicadores altmétricos são fracas (0,313 e 0,219, respectivamente). A força destas correlações indica que as menções de artigos nestas duas redes sociais estão relacionadas, principalmente para artigos da área de Ciências da Vida, onde os usuários (pesquisadores ou não) compartilham os artigos que consideram relevantes em ambas as redes sociais.

Na análise entre indicadores de citação, a correlação entre citações na WoS e na *Scopus* é perfeita para a área de Ciências da Vida. Enquanto isso, para a área de Ciências Sociais Aplicadas esta correlação é fraca (0,330). Isso indica que para a área de Ciências da Vida há um número maior de citações sobrepostas entre as fontes *Scopus* e WoS, com o oposto ocorrendo para artigos da área de Ciências Sociais Aplicadas, onde a maioria das citações captadas pelas fontes WoS e *Scopus* para um mesmo artigo são diferentes (VISSER; VAN

ECK; WALTMAN, 2021).

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. Para a área de Ciências Exatas, Tecnológicas e Multidisciplinar essas correlações não foram calculadas, uma vez que cinco dos 3.868 artigos da área possuem alguma citação na *Scopus*, uma amostra menor do que a mínima necessária (50) para análises correlacionais (FRAENKEL; WALLEN; HYUN, 2012).

A correlação entre os indicadores de acesso ao resumo e citações da *Crossref* é nula para todas as áreas de conhecimento, variando de 0,085 a 0,203. Para a Ciências da Vida, apesar da correlação entre acessos ao resumo e citações na *WoS* ser nula, de 0,050, a correlação entre *downloads* e citações na *WoS* é moderada, de 0,449.

Entre as postagens no *Facebook* e as citações na *Crossref* a correlação é nula para todas as áreas, com a maior correlação, de 0,162, para as Ciências Sociais Aplicadas. Para todas as áreas fica evidente que as postagens no *Facebook* não exercem influência no número de citações que os artigos recebem nas fontes bibliométricas. As correlações entre o *Twitter* e as citações na *Crossref* são próximas às correlações entre o *Facebook* e estas citações, sendo nulas para todas as áreas. Estes resultados estão em concordância com os de Chen *et al.* (2020), que encontraram uma correlação fraca de 0,004 entre o indicador do Altmetric – composto, em sua maioria, por menções no *Twitter* e *Facebook* – e citações para artigos de Ciências Clínicas e Ciência Básica.

4 CONCLUSÃO

A análise dos dados evidenciou que, para artigos publicados em portais de periódicos brasileiros de acesso aberto, os artigos apresentam baixa cobertura nas bases de dados de citação internacionalmente consolidadas. Novas fontes de indicadores de citação devem ser levadas em consideração na avaliação e medição da ciência publicada no Brasil, visto que a *Crossref* apresenta maior cobertura que as demais bases analisadas.

Conforme as coberturas obtidas no estudo apresentado, a área de

Ciências da Vida demonstra ter mais leitores no *Mendeley* por artigo, bem como maior quantidade de artigos lidos/capturados pela plataforma. Isso indica que parece haver maior utilização da plataforma por usuários interessados em artigos científicos desta área. A cobertura de citação também indica que, para algumas áreas, como a Ciências da Vida, as fontes de dados bibliométricas captam mais citações iguais do que outras, indicando uma sobreposição entre as fontes.

Ainda que as coberturas de menções em postagens no *Facebook* e no *Twitter* sejam mais baixas que as de estudos anteriores, o comportamento de publicações das áreas de conhecimento relacionadas às Ciências Exatas reafirma que a prática de compartilhar ou debater conteúdo científico desta área de conhecimento no *Facebook* não é comum. De modo geral, as coberturas encontradas estão em concordância com resultados de pesquisas anteriores e indicam haver diferenças nas práticas acadêmicas entre as áreas do conhecimento.

As correlações mais fortes encontradas entre indicadores altmétricos e de citação também foram para a área de Ciências da Vida, e a mais fraca para a área de Linguística, Letras e Artes, indicando uma possível maior variedade de público para esta área – que nem sempre cita o artigo, lendo-o com outros objetivos que não apenas a pesquisa científica. O *Mendeley* se reafirma como a melhor fonte de indicadores altmétricos para medição do impacto acadêmico nas mídias sociais, bem como para complementar a avaliação de publicações científicas, ajudando a prever citações futuras. Estudos futuros devem continuar a explorar o potencial do *Mendeley*, utilizando amostras maiores e mais variadas, de modo a compreender as possibilidades de uso dos indicadores altmétricos de acordo com as práticas científicas de cada área de conhecimento, adequando este uso para complementar os indicadores de citação. Assim, o desenvolvimento e a implantação de novas formas de avaliação da qualidade de publicações científicas se adequarão às variações existentes dentro da comunicação científica em cada área de conhecimento, além de abranger mais fielmente a publicação científica de países sub-representados nos modelos atuais de avaliação de qualidade.

REFERÊNCIAS

- AHMED, A. Open access towards bridging the digital divide-policies and strategies for developing countries. **Information Technology for Development**, [S. l.], v. 13, n. 4, p. 337-361, 2007. Disponível em: <https://www.tandfonline.com/doi/abs/10.1002/itdj.20067>. Acesso em: 19 out. 2022.
- ALPERIN, J. P. Geographic variation in social media metrics: an analysis of Latin American journal articles. **Aslib Journal of Information Management**, [S. l.], 2015. Disponível em: https://www.researchgate.net/publication/277911952_Geographic_variation_in_social_media_metrics_An_analysis_of_Latin_American_journal_articles. Acesso em: 18 out. 2022.
- BRIGHAM, T. J. An introduction to altmetrics. **Medical Reference Services Quarterly**, [S. l.], v. 33, n. 4, p. 438-447, 2014. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/02763869.2014.957093>. Acesso em: 19 out. 2022.
- CÉSPEDES, L. Revistas latino-americanas e línguas hegemônicas para publicação acadêmica no Scopus e Web of Science. **Trabalhos em Linguística Aplicada**, [S. l.], v. 60, n. 1, p. 141-154, 2021. Disponível em: <https://www.scielo.br/j/tla/a/hfQ6XVh9MDJbj9ZNVqRzn3y/?lang=en#>. 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. Disponível em: https://www.researchgate.net/publication/333022684_The_relationship_between_citations_downloads_and_alternative_metrics_in_rheumatology_publications_a_bibliometric_study. Acesso em: 15 jan. 2022.
- 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, 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. 2022.
- CROSSREF. **Rest API**. [S. l.: s. n.], 2020. Disponível em: <https://www.crossref.org/documentation/retrieve-metadata/rest-api/>. Acesso em: 15 jan. 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**, [S. l.], p. 1-16, 2022. Disponível em:

<https://www.tandfonline.com/doi/full/10.1080/00987913.2022.2066967>. Acesso em: 20 jan. 2022.

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.

FRAENKEL, J. R.; WALLEN, N. E.; HYUN, H. H. **How to design and evaluate research in education**. 11ª edição. Nova York: McGraw-hill, 2012.

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://www.researchgate.net/publication/271919328_Using_altmetrics_for_assessing_research_impact_in_the_humanities. 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://www.researchgate.net/publication/273955648_Characterizing_Social_Media_Metrics_of_Scholarly_Papers_The_Effect_of_Document_Properties_and_Collaboration_Patterns. 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. **Anais [...]**. Helsinki: Finland, 2014. Disponível em: <https://docs.lib.purdue.edu/iatul/2014/altmetrics/2/>. Acesso em: 22 jul. 2022.

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. Disponível em: https://www.researchgate.net/publication/266396518_Tweets_vs_Mendeley_readers_How_do_these_two_social_media_metrics_differ. Acesso em: 15 jan. 2022.

HUANG, W.; WANG, P.; WU, Q. A correlation comparison between Altmetric Attention Scores and citations for six PLOS journals. **PLoS One**, [S. l.], v. 13, n. 4, Apr. 2018. Disponível em: https://www.researchgate.net/publication/324253303_A_correlation_comparison_between_Altmetric_Attention_Scores_and_citations_for_six_PLOS_journals. Acesso em: 25 jan. 2022.

KEELE, B. J. Law libraries as publishers: counting things, with altmetrics. **ALL-SIS Newsletter**, [S. l.], v. 35, n. 2, p. 15-16, 2016. Disponível em: <https://scholarworks.iupui.edu/handle/1805/8249>. Acesso em: 15 fev. 2022.

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://www.researchgate.net/publication/257662983_Validating_Online_Reference_Managers_for_Scholarly_Impact_Measurement. Acesso em: 12 dez. 2021.

MARICATO, J. M.; MARTINS, D. L. Almetria: complexidades, desafios e novas formas de mensuração e compreensão da comunicação científica na web social. **Biblios**, [S. l.], v. 68, n. 68, p. 48–68, 2017. Disponível em:

http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1562-47302017000300004. Acesso em: 3 mar. 2022.

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**, Campinas, v. 33, 2021. Disponível em:

<https://www.scielo.br/j/tinf/a/KmfGWMDK8zSvZszWfSX4VnP/>. Acesso em: 3 jan. 2022.

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.

Anais [...]. Vienna: Austrian Institute of Technology GmbH, 2013. p. 200-214. Disponível em:

https://www.researchgate.net/publication/279553954_Assessing_the_Mendeley_readership_of_social_sciences_and_humanities_research. Acesso em: 12 dez. 2021.

MOHAMMADI, E.; THELWALL, M. Mendeley readership altmetrics for the social sciences and humanities: research evaluation and knowledge flows.

JASIST, [S. l.], v. 65, n. 8, 2014. Disponível em:

<https://asistdl.onlinelibrary.wiley.com/doi/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. **JASIST**, [S. l.], v. 66, n. 9, p. 1832-1846, 2015. Disponível em:

https://www.researchgate.net/publication/260515768_Who_Reads_Research_Articles_An_Altmetrics_Analysis_of_Mendeley_User_Categories. Acesso em: 15 mar. 2022.

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, v. 31, 2019. Disponível em:

https://www.researchgate.net/publication/335612109_Panorama_da_producao_cientifica_do_Brasil_alem_da_indexacao_uma_analise_exploratoria_da_comunicacao_em_periodicos. Acesso em: 4 jan. 2022.

MUELLER, S. P. M. A publicação da Ciência: áreas científicas e seus canais preferenciais. **DataGramZero**, [S. l.], v. 6, n. 1, fev. 2005. Disponível em: <https://repositorio.unb.br/handle/10482/980>. Acesso em: 19 jul. 2022.

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.researchgate.net/publication/337144778_Do_altmetrics_correlate_with_citations_A_study_based_on_the_1000_most-cited_articles. Acesso em: 19 jul. 2022.

PACKER, A. L. Os periódicos brasileiros e a comunicação da pesquisa nacional. **Revista USP**, São Paulo, n. 89, p. 26-61, 2011. Disponível em: <https://www.revistas.usp.br/revusp/article/view/13868>. Acesso em: 15 jan. 2022.

RAMOS, T. B. C.; MARICATO, J. M. **Odisseia metrics**: ferramenta de extração de dados. Versão 1.0. [S. l.: s. n.], 2020. Disponível em: https://colab.research.google.com/drive/1L75mJbqteVE0kMoTMK1_kLVXWFZpQumi#scrollTo=ohRqrEb7RXIm. Acesso em: 10 dez. 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.

SCHLÖGL, C.; GORRAIZ, J.; GUMPENBERGER, C.; JACK, K.; KRAKE, P. Are downloads and readership data a substitute for citations? The case of a scholarly journal. *In*: BIENAL COURSE AND CONFERENCE LIBRARIES IN THE DIGITAL AGE, 2014, Zadar, Croatia. **Anais** [...]. Zadar, Croatia: Department of Library and Information Science, University of Zadar, 2014. p. 255-261. Disponível em: https://www.fhgr.ch/fileadmin/fhgr/angewandte_zukunftstechnologien/SII/publikationen/2014/sii-publikationen-Proceedings_der_LIDA_2014.pdf#page=254. Acesso em: 15 jan. 2022.

SILVA FILHO, R. C. **Influência das métricas de mídias sociais nas citações dos periódicos brasileiros de Enfermagem indexados na Scopus**. 2022. Tese (Doutorado em Comunicação e Informação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022. Disponível em: <https://www.lume.ufrgs.br/handle/10183/238845>. Acesso em: 15 maio 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**, [S. l.], v. 126, n. 6, p. 5113-5142, 2021. Acesso em: 25 jan. 2022.

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, 2018. **Anais** [...]. 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.

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. Disponível em: https://www.researchgate.net/publication/319937939_Do_Mendeley_reader_counts_indicate_the_value_of_arts_and_humanities_research. Acesso em: 15 mar. 2022.

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. Disponível em: https://www.researchgate.net/publication/236978654_Do_Altmetrics_Work_Twitter_and_Ten_Other_Social_Web_Services. Acesso em: 15 fev. 2022.

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. Disponível em: https://www.researchgate.net/publication/277724700_Mendeley_readership_counts_An_investigation_of_temporal_and_disciplinary_differences. Acesso em: 25 abr. 2022.

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.

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

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.

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.

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**, [S. l.], v. 2, n. 1, p. 20-41, 2021.

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://www.researchgate.net/publication/265687564_From_Attention_to_Citati

on_What_and_How_Does_Altmetrics_Work. 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.

ZAHEDI, Z.; COSTAS, R.; WOUTERS, P. How well developed are altmetrics? A cross disciplinary analysis of the presence of “alternative metrics” in scientific publications. **Scientometrics**, [S. l.], v. 101; n. 2, p. 1491-1513, 2014.

Disponível em: <https://link.springer.com/article/10.1007/s11192-014-1264-0>.

Acesso em: 11 dez. 2021.

ZAHEDI, Z.; COSTAS, R.; WOUTERS, P. Mendeley readership as a filtering tool to identify highly cited publications. **Journal of the Association for Information Science and Technology**, [S. l.], v. 68, n. 10, p. 2511-2521, 2017.

Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.23883>.

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.

COVERAGE OF DATA SOURCES AND CORRELATIONS BETWEEN ALTMETRIC INDICATORS AND CITATIONS FOR BRAZILIAN JOURNAL PORTALS: COMPARISON OF DIFFERENT KNOWLEDGE AREAS

ABSTRACT

Objective: to contribute to the analysis of altmetric and citation data sources coverage and the correlations between altmetric and bibliometric indicators. **Methodology:** it was analyzed indicators for 51,200 articles from four journal portals, divided into five areas of knowledge. **Results:** the accesses to the abstract and the Mendeley readers had the greatest altmetric coverage, although they show great differences between areas. The citation data source with the greatest coverage was Crossref, ranging from 7.3% to 43.8%, depending of the area. The correlations between altmetric indicators and citations also presents variations between the areas (from weak to moderate). Life Sciences showed the strongest correlations, with emphasis on the correlations between Mendeley readers and citations, ranging from 0.432 to 0.491. **Conclusions:** we conclude that Mendeley (altmetric data) and Crossref (citation data), can be considered alternative and/or complementary data sources for the evaluation of Brazilian science in all areas of knowledge, although there are differences between them. Finally, the data suggest the importance of open access scientific journals and university journal portals for the dissemination of science in the country.

Descriptors: Scientific Communication. Bibliometric studies. Altmetric studies. Scientific production.

COBERTURA DE FUENTES DE DATOS Y CORRELACIONES ENTRE INDICADORES ALTMÉTRICOS Y COTIZACIONES EN PORTALES DE REVISTAS BRASILEÑAS: COMPARACIONES EN DIFERENTES ÁREAS DEL CONOCIMIENTO

RESUMEN

Objetivo: contribuir al análisis de la cobertura de las fuentes de datos de los indicadores altmétricos y de citas y las correlaciones entre ellos, en diferentes áreas del conocimiento. **Metodología:** se analizaron indicadores de 51.200 artículos de cuatro portales de revistas, divididos en cinco áreas de conocimiento. **Resultados:** los accesos al resumen y las lecturas/capturas en Mendeley mostraron una mayor cobertura, aunque existen grandes diferencias entre áreas. La fuente de datos de citas con mayor cobertura fue Crossref, con un rango de 7.3 a 43.8% dependiendo del área. Las correlaciones entre los indicadores altmétricos y de citas también muestran variaciones entre áreas (entre débiles y moderadas). Ciencias de la vida tuvo las correlaciones más fuertes, con énfasis en aquellas entre los lectores de Mendeley y las citas, con un rango de 0,432 a 0,491. **Conclusiones:** Se concluye que Mendeley (datos altmétricos) y Crossref (datos de citas) pueden ser considerados fuentes de datos alternativas y/o complementarias para la evaluación de la ciencia brasileña en todas las áreas del conocimiento, aunque existen diferencias entre ellos. Finalmente, los datos sugieren la importancia de las revistas científicas de acceso abierto y los portales de revistas universitarias para la divulgación de la ciencia que se produce en el país.

Descriptores: Comunicación científica. Estudios bibliométricos. Estudios altmétricos.

Recebido em: 19.12.2022

Aceito em: 24.03.2023