



**GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação**

**ISSN 2177-3688**

**CITAÇÕES ABERTAS COMO ALTERNATIVA PARA ESTUDOS BIBLIOMÉTRICOS: O CASO DE DUAS REVISTAS DA COLEÇÃO SCIELO**

***OPEN CITATIONS AS AN ALTERNATIVE FOR BIBLIOMETRIC STUDIES: THE CASE OF TWO JOURNALS FROM THE SCIELO COLLECTION***

**Sônia Elisa Caregnato** - Universidade Federal do Rio Grande do Sul  
**Rene Faustino Gabriel Junior** - Universidade Federal do Rio Grande do Sul  
**Ivan Heibi** - Universidade de Bologna, Itália  
**Sílvio Peroni** - Universidade de Bologna, Itália

**Modalidade: Trabalho Completo**

**Resumo:** Os estudos sobre produção e citação da literatura científica brasileira, assim como de outros países periféricos, enfrentam obstáculos relacionados à utilização de índices de citações tradicionais como fonte de dados. O surgimento da Initiative for Open Citation (I4OC), que promove a disponibilização aberta e estruturada de citações por meio da Crossref, impactou a realização desses estudos. A OpenCitations, membro da I4OC, criou uma fonte alternativa e aberta para estudos de citação, a OpenCitations Index of Crossref open DOI-to-DOI citations (COCI), que coleta metadados e citações como Linked Data. Nesse contexto, o objetivo deste trabalho é discutir o potencial do COCI para estudos bibliométricos, avaliando a proporção de artigos de duas revistas brasileiras com DOI publicadas em acesso aberto pelo Scielo que está presente no COCI, comparativamente aos dados nas bases Scielo Citation Index, Web of Science e Scopus. Apesar de promissor, ainda são necessários estudos para avaliar o potencial do COCI em estudos bibliométricos, comparando sua cobertura com outras bases e analisando o volume de citações recuperadas.

**Palavras-chave:** citações abertas; índices de citações; OpenCitations/COCI.

**Abstract:** Studies on the production and citation of Brazilian scientific literature, as well as that of other peripheral countries, face obstacles related to the use of traditional citation indices as data source. The emergence of the Initiative for Open Citation (I4OC), which promotes the open and structured availability of citations through Crossref, has impacted the conduct of these studies. OpenCitations, a member of I4OC, has created an alternative and open source for citation studies, the OpenCitations Index of Crossref open DOI-to-DOI citations (COCI), which collects metadata and citations as Linked Data. The aim of this study is to discuss the potential of COCI for bibliometric studies, evaluating the proportion of articles from two Brazilian journals with DOIs published in open access by Scielo that are present in COCI, in comparison to data from Scielo Citation Index, Web of Science, and Scopus. Despite its promise, further research is still needed to assess the potential of COCI in bibliometric studies, comparing its coverage with other databases and analyzing the volume of citations retrieved.

**Keywords:** open citations; citation indexes; OpenCitations/COCI.



## 1 INTRODUÇÃO

Os estudos sobre produção e citação da literatura científica brasileira, bem como a de outros países periféricos, enfrentam grandes obstáculos ao utilizar os índices de citações tradicionais. A utilização do Google Scholar, que se pensava poder ser uma alternativa a essas ferramentas, não se mostrou como uma opção viável, devido ao volume de problemas na indexação (HALEVI; MOED; BAR-ILAN, 2017), à dificuldade de extração e download de dados (MARTÍN-MARTÍN *et al.*, 2017), às limitadas opções de busca e à falta de transparência do algoritmo (MARTÍN-MARTÍN *et al.*, 2021).

Considerando a agenda atual da Ciência Aberta, que busca ampliar a qualidade, o acesso e a transparência da pesquisa, inclusive da pesquisa bibliométrica (BORNMANN *et al.*, 2021), a busca por fontes de dados alternativas torna-se ainda mais relevante. A chamada pela “abertura da caixa preta” da avaliação da produção e citação científica pelo Manifesto de Leiden (HICKS *et al.*, 2015) e pela Declaration on Research Assessment (DORA) (2012) enfatiza a necessidade de aprimorar as formas como a bibliometria/cientometria é aplicada.

Um movimento recente e importante neste sentido é o da Initiative for Open Citation (I4OC), lançado oficialmente em 2017. Trata-se do esforço colaborativo de organizações e indivíduos para promover a disponibilização de citações à literatura científica de forma aberta e estruturada (SCHIERMEIER, 2017). O instrumento para isso é a Crossref, entidade que fornece DOI para publicações acadêmicas e armazena metadados dessas publicações juntamente com as listas de referências. Até recentemente, essas referências só eram disponibilizadas no formato aberto se o editor concordasse formal e explicitamente. No entanto, como resultado da defesa do acesso aberto às citações, a partir de junho de 2022 a Crossref liberou a distribuição de todas as citações depositadas por seus membros. Estima-se que hoje haja mais de 61 milhões de publicações com citações abertas (I4OC, 2023).

Um dos membros fundadores da I4OC é a OpenCitations, uma organização independente, liderada pela comunidade, sem fins lucrativos e voltada à infraestrutura da Ciência Aberta, cujo objetivo é coletar e publicar de forma aberta metadados precisos e abrangentes que descrevem as publicações científicas do mundo e as citações que as conectam. Atualmente gerida a partir do Research Centre for Open Scholarly Metadata, da University of Bologna, a OpenCitations foi estabelecida com o objetivo específico de romper com o *status quo* ao fornecer uma alternativa totalmente gratuita e aberta para acessar dados de citação em formato de Linked Data (PERONI; SHOTTON, 2020).

Dentre os diversos serviços, atividades e ferramentas desenvolvidas pela OpenCitations, destaca-se o COCI, que é um índice de citações abertas de DOI a DOI da Crossref (OpenCitations Index of Crossref open DOI-to-DOI citations). Ou seja, o COCI coleta metadados de publicações que receberam DOI e cujos metadados foram depositados na Crossref para criar um índice de citações em que as citações são tratadas e disponibilizadas como Linked Data, em vez de serem definidas simplesmente como relações entre dois recursos bibliográficos. (HEIBI; PERONI; SHOTTON, 2019).

O COCI foi projetado para ser de livre acesso e permitir que pesquisadores, instituições e organizações analisem e utilizem os dados de citação para diversos fins de pesquisa e avaliação. Atualmente, o índice conta com quase um bilhão e meio de citações, relativas 77 milhões de registros bibliográficos (<https://opencitations.net/index/coci>). É importante frisar, no entanto, que o COCI somente adiciona DOIs validados no Crossref, ou seja, se os editores adicionaram metadados com problemas, os DOIS não serão incorporados ao índice.

Pelas suas características, o COCI se configura como um candidato a suprir a necessidade de fontes mais inclusivas para a realização de estudos bibliométricos (tanto para estudos que procuram entender a composição e as práticas de produção científicas, como aqueles que buscam auxiliar nos processos relacionados à avaliação da atividade científica) acerca da literatura científica que se encontra fora do *mainstream*, como é o caso de grande parte dos artigos publicados em revistas científicas nacionais.

Mesmo apresentando este potencial, que tende a se expandir na medida em que mais editores registrem artigos no Crossref para obtenção de DOI, ainda não há muitos estudos que se proponham a avaliar o potencial do COCI como fonte de dados para estudos bibliométricos. Também não há estudos que comparem sua cobertura com aquela de outros índices de países periféricos, como o Scielo, nem o volume de citações que recupera.

Assim, o objetivo deste trabalho é discutir o potencial do COCI para estudos bibliométricos, avaliando a proporção de artigos de revistas brasileiras com DOI publicadas em acesso aberto pelo Scielo que está presente no COCI, assim como comparando os dados recuperados nele com aquele nas bases Scielo, Web of Science e Scopus.

Trata-se de buscar fontes de dados bibliométricos mais inclusivas, mas que, ao mesmo tempo, disponibilizem informações confiáveis e que permitam a extração de grandes quantidades de dados. Além disso, é necessário que essas fontes deem autonomia ao

pesquisador para calcular os indicadores que julgar apropriado para aquele campo/situação específica, ou seja, que ofereçam os dados não agradados, em oposição a indicadores prontos de revistas, como FI, JSR e h5, por exemplo.

Para esta avaliação inicial, foram escolhidas duas revistas da coleção Scielo: Revista da Sociedade Brasileira de Medicina Tropical (RSBMT) e Anais da Academia Brasileira de Ciências (AABC). A primeira foi selecionada pela visibilidade e por apresentar DOI em todos os seus artigos, desde 1967. Os Anais integram este estudo inicial pelo seu prestígio e pela sua antiguidade, que garante mais tempo para receber citações, mesmo que tenha sido incorporada na coleção Scielo somente em 2000, o que faz com que os artigos anteriores não tenham recebido DOI.

Além de estarem ambas em acesso aberto e terem suas citações também abertas no Crossref, são revistas bastante citadas na coleção Scielo. Por estarem indexadas na WoS e Scopus, é possível traçar comparações entre o COCI e essas duas bases, além de que o fato de serem publicada pelo Scielo permite que se tenha os números da produção totais mapeados previamente.

## **2 METODOLOGIA**

A primeira etapa consistiu na extração de dados por meio de script na linguagem Python utilizando a técnica de Web Crawling. O script foi executado em agosto de 2022 nas páginas iniciais das duas revistas no Scielo para extrair as URL dos fascículos e de cada artigo com seu respectivo DOI e ano de publicação. As informações coletadas foram armazenadas em um banco de dados MySQL e resultaram em 5302 registros DOI da Revista da Sociedade Brasileira de Medicina Tropical e 3713 da Anais da Academia Brasileira de Ciências.

Na segunda etapa, realizada em outubro de 2022, os DOIs dos artigos foram buscados no COCI e extraídas as relações entre artigos citados e citantes (quando havia pelo menos uma citação). O resultado foi apresentado com um arquivo CSV, que consiste numa lista de todas as citações a cada um dos DOIs das publicações no formato ["oci"] ["citing"] [field "cited"] [field "creation"] [field "timespan"] [field "journal\_sc"] [field "author\_sc"], exemplificado a seguir.

Para o DOI 10.1590/0037-8682-0275-2021 em RSBMT foi encontrada a seguinte citação no COCI, explicitada como uma linha no arquivo CSV:

02001050900360001000263030101330000020303020201-  
020010509003600000307630806080263000207056302000201, 10.1590/0102-  
311x00233221, 10.1590/0037-8682-0275-2021, 2022, POY, no, yes

Onde:

[02001050900360001000263030101330000020303020201-  
020010509003600000307630806080263000207056302000201] representa o  
Identificador de Citação Aberta (OCI);

[10.1590/0102-311x00233221] representa o DOI do documento citante;

[10.1590/0037-8682-0275-2021] representa o DOI do documento citado;

[2022] representa a data de publicação do documento citado;

[POY] representa o intervalo de tempo entre a publicação e a citação recebida;

[no] indica que não se trata de autocitação de periódico;

[yes] indica que se trata de uma autocitação de autor.

Da Coleção Principal da Web of Science, que inclui os índices Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, Arts & Humanities Citation Index, Conference Proceedings Citation Index e Emerging Sources Citation Index, e o Scielo Citation Index, também na Plataforma Web of Science, foram recuperados os números de produção e citação. Nos resultados essas duas fontes são apresentadas como WoS e SCI. Por fim, a base de dados Scopus também forneceu os dados de produção e citação indexados nela. As coletas nessas bases de dados foram realizadas no mês de abril de 2023.

Considerou-se o número total de artigos apresentado no site do Scielo como o padrão para avaliação da cobertura nas demais fontes, já que se trata do *publisher* das revistas. Assim, o índice de cobertura da RSBMT foi derivado da fórmula “Cobertura = (Itens Relevantes Recuperados no COCI/WoS/SCI/Scopus) / (Total de Itens no Site da Scielo)”. Como os Anais da Academia Brasileira de Ciências entraram na Coleção Scielo em 2000 e, portanto, somente passaram a receber DOI naquele ano, considerou-se como padrão o total da base de dados WoS, que apresentou o resultado mais elevado. Neste caso a fórmula foi “Cobertura = (Itens Relevantes Recuperados no COCI/SCI/Scopus) / (Total de Itens WoS)”.

Em relação às citações recebidas, no entanto, não é possível prever o número exato, pois os resultados são dependentes da base de dados utilizada para coleta. Por isso, a cobertura das citações não foi calculada, mas sim a média de citações recebidas por publicação em cada base individualmente.

### 3 RESULTADOS E DISCUSSÃO

A Revista da Sociedade Brasileira de Medicina Tropical (RSBMT) é uma publicação oficial da Sociedade Brasileira de Medicina Tropical, uma organização que reúne profissionais de diversas áreas da saúde interessados na área especificada no título do periódico. Trata-se de uma revista multidisciplinar e aberta, que publica bimestralmente textos em inglês e que é indexada em várias bases de dados, incluindo, além do Scielo, PubMed, Lilacs, Web of Science, Scopus e Latindex. Fundada em 1967, a revista tem como objetivo promover a pesquisa, o conhecimento e a disseminação de informações sobre doenças tropicais e infecto-parasitárias (Sociedade Brasileira de Medicina Tropical, 2023).

Na coleta de DOIs por meio do Web Crawling, foram encontrados 5302 registros relativos aos 5454 artigos da revista publicados na Scielo (todas os artigos receberam DOI desde 1967), ou seja, 2,8% dos DOIs não foram encontrados na busca feita pelo robô na web.

**Tabela 1** - Comparação dos números de publicações e de citações recuperados da Revista da Sociedade Brasileira de Medicina Tropical entre diferentes fontes de dados

	Início	Nº Publicações	Cobertura (%)	Nº Cit. recebidas	Citações/Publ
COCI	1967	4150	76,1	38019	9,2
SCI	2000	3406	62,4	9817	2,9
Scopus	1972	4811	88,2	48679	10,1
WoS	2000	3402	62,4	27513	8,1

**Fonte:** Elaborado pelos autores

Dentre os 5302 DOIs relativos aos documentos publicados pela revista, COCI encontrou 4150, além de 38019 DOIs de documentos citantes, como mostra a Tabela 1. Portanto, ele conseguiu recuperar 76,1% dos artigos existentes. Há que se considerar, no entanto, que o COCI recupera citações a publicações, o que exclui as publicações que não foram citadas nenhuma vez. Esse resultado no COCI, se for confirmado em etapas subsequentes da pesquisa, difere de estudo anterior que encontrou pouco mais de 50% das citações encontradas no WoS estavam presentes no COCI (CHUDLARSKÝ; DVOŘÁK, 2020). A WoS e o SCI foram as fontes de menor cobertura, mas isso se deve também ao fato de indexarem somente as publicações de 2000 em diante. Em relação às citações, o SCI

recuperou apenas uma média de 2,9 citações por artigo, bem menos que a Coleção Principal da WoS.

A Scopus recuperou o maior número de publicações, incluindo citadas e não-citadas, e também registrou o maior número de citações a essas publicações. A média de citações por publicação no COCI e na Scopus não difere significativamente.

O número bem maior de citações na Scopus provavelmente se deve a incorporação nela de dados de citação das publicações da Elsevier, que até recentemente não estavam disponíveis em formato aberto e, portanto, não podem ser recuperadas pelo COCI. Por outro lado, não se pode descartar a duplicação de registros; como observado por van Eck e Waltman (2019) há sérios problemas de duplicação de registros, entre outros problemas, na Scopus e no WoS.

No geral, o COCI demonstrou ter uma boa cobertura dos artigos publicados na Revista da Sociedade Brasileira de Medicina Tropical.

Os Anais da Academia Brasileira de Ciências (AABC) configuram-se como a revista científica mais antiga do Brasil. Fundada em 1929, ela vem sendo publicada de forma ininterrupta desde então, tendo como objetivo principal promover a disseminação do conhecimento científico nas diversas áreas da ciência abrangidas pelas Seções da Academia Brasileira de Ciências. Trata-se, portanto, de uma revista multidisciplinar, que é publicada bimestralmente e é indexada no Chemical Abstracts Service, na Medline, na Lilacs, na BIOSIS, na Web of Science e na Scopus, entre outras (Academia Brasileira de Ciências, 2023).

**Tabela 2** - Comparação dos números de publicações e de citações recuperados dos Anais da Academia Brasileira de Ciências entre diferentes fontes de dados

	Início	Nº Publicações	Cobertura (%)	Nº Cit. recebidas	Citações/Publici.
COCI	2000	2708	39,1	27975	10,3
SCI	2002	3732	53,8	3980	1,1
Scopus	1945	4280	61,7	42146	9,8
WoS	1945	6933	100,0	41238	5,9

**Fonte:** Elaborado pelos autores

Na coleta de DOIs na web foram encontrados 3713 registros. Não se sabe o número total de artigos da revista, pois a coleção Scielo não incorpora os fascículos publicados entre 1929 e 1999. A WoS e Scopus, por outro lado, indexam artigos publicados desde 1945 pela

revista. Como a WoS apresenta o maior número de registros entre todas as fontes consultadas, assumiu-se esse número com sendo o total, para fins de comparação.

Dentre os 3713 DOIs relativos aos itens publicados pela revista desde 2000 (os que estão na Coleção Scielo), COCI encontrou 2708, além de 27975 DOIs de documentos citantes, como mostra a Tabela 2. Há duas explicações possíveis pela baixa cobertura do COCI (39,1%): uma grande parcela dos trabalhos da revista não foi citada, e, portanto, não estabeleceu relações de citação para serem expostas, ou muitos dos metadados e listas de citação dos Anais apresentam problemas de registro, o que faria com que não fossem recuperados do Crossref.

Diferente de vários estudos sobre a cobertura das bases de dados de citação, em que a Scopus geralmente aparece com tendo a maior cobertura do que a WoS (por exemplo, HARZING, 2019; MARTÍN-MARTÍN *et al.*, 2021; VISSER; VAN ECK; WALTMAN, 2021), para os dados de produção citação da AABC isso não aconteceu. Para os dados de citações, no entanto, a Scopus supera a WoS, com uma média de 9,8 citações por artigo e a COCI supera ambas, com uma média de 10,3 citações por artigo. A peculiaridade do caso da revista indica que será preciso investigar mais profundamente as razões para essas discrepâncias.

Em relação aos dados apresentados, pode haver indícios de que a liberação de todas as citações depositadas no Crossref para acesso aberto em junho de 2022 (I4OC, 2023) já esteja fazendo efeito, na medida em que permite uma recuperação maior de citações pelo COCI. Isso é particularmente relevante quando comparado a 2019, quando Nishioka *et al.* (2019) encontraram abertas as listas de citações de somente 24,2% das publicações com DOI no Crossref.

## **CONSIDERAÇÕES FINAIS**

O COCI demonstrou ter uma boa cobertura dos artigos publicados na RSBMT, tendo conseguido recuperar 76,1% dos artigos existentes. Ele também apresentou uma média de citações por publicação semelhante à Scopus, uma das bases de dados de citação mais abrangente, dentre as fechadas.

A cobertura do COCI para os artigos publicados na AABC foi bastante menor, o que pode ser explicado pela baixa taxa de citação dos trabalhos da revista ou por possíveis problemas nos metadados e nas listas de citação dos Anais. Em relação às citações, o COCI

superou tanto WoS como Scopus. A particularidade dos dados da revista indica a necessidade de investigação mais aprofundada.

Os dados do Scielo Citation Index se mostraram bastante modestos em relação às duas revistas, o que aponta para uma relevância maior para o Scielo como *publisher* do que como fornecedor de um índice de citações.

Provavelmente a liberação de todas as citações depositadas no Crossref para acesso aberto possibilitará que se criem fontes de dados para estudos bibliométricos mais inclusivas.

A OpenCitations tem planos de estabelecer um índice compreensivo de citações, a partir de publicações de todo o mundo. Este índice abrangerá não apenas as citações fornecidas pela Crossref (ou seja, conjunto de dados COCI), mas também citações de outras fontes primárias proeminentes, como Datacite (<https://datacite.org/>) e Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>). Ao incorporar essas fontes adicionais, a OpenCitations tem como objetivo aumentar significativamente a abrangência das citações, potencialmente fortalecendo a robustez dos resultados da pesquisa aqui apresentada.

## REFERÊNCIAS

ACADEMIA BRASILEIRA DE CIÊNCIAS. **Anais da Academia Brasileira de Ciências**: sobre o periódico. 2023. Disponível em: <https://www.scielo.br/journal/aabc/about/#about>

BORNMANN, L., GUNS, R., THELWALL, M., WOLFRAM, D. Which aspects of the Open Science agenda are most relevant to scientometric research and publishing? An opinion paper. **Quantitative Science Studies**, v. 2, n.2, p. 438–453, 2021. Disponível em: [https://doi.org/10.1162/qss\\_e\\_00121](https://doi.org/10.1162/qss_e_00121)

CHUDLARSKÝ, T.; & DVOŘÁK, J. Can crossref citations replace web of science for research evaluation? The share of open citations. **Journal of Data and Information Science**, v. 5, n. 4, 2020, p. 35–42. Disponível em: <https://doi.org/10.2478/jdis-2020-0037>.

DECLARATION ON RESEARCH ASSESSMENT (DORA). **San Francisco Declaration on Research Assessment**, 2012. Disponível em: <https://sfdora.org/read/>

HALEVI, G, MOED, H., BAR-ILAN, J. Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature. **Journal of Informetrics**, v. 11, n. 3, p. 823-834, 2017. Disponível em: <https://doi.org/10.1016/j.joi.2017.06.005>.

HARZING, AW. Two new kids on the block: How do Crossref and Dimensions compare with Google Scholar, Microsoft Academic, Scopus and the Web of Science? **Scientometrics**, v.120, p. 341–349, 2019. Disponível em: <https://doi.org/10.1007/s11192-019-03114-y>

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>

HICKS, D., WOUTERS, P., WALTMAN, L. *et al.* Bibliometrics: The Leiden Manifesto for research metrics. **Nature**, v. 520, p. 429–431, 2015. Disponível em <https://doi.org/10.1038/520429a>

I4OC. **Initiative for Open Citations**. 2023. Disponível em: <https://i4oc.org/>. Acesso em: 17 jun. 2023.

MARTIN-MARTIN, A., ORDUNA-MALEA, E., HARZING, A.-W., DELGADO LÓPEZ-CÓZAR, E. Can we use Google Scholar to identify highly-cited documents? **Journal of Informetrics**, v. 11, n. 1, p. 152–163, 2017. Disponível em: <https://doi.org/10.1016/j.joi.2016.11.008>.

MARTÍN-MARTÍN, A., THELWALL, M., ORDUNA-MALEA, E. *et al.* Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. **Scientometrics**, v. 126, p. 871–906, 2021. Disponível em: <https://doi.org/10.1007/s11192-020-03690-4>

NISHIOKA, C. *et al.* Evaluating the Availability of Open Citation Data. BIRNDL 2019: Bibliometric-enhanced Information Retrieval and Natural Language Processing for Digital Libraries, 2019. **Anais...** 2414, p. 123-129, 2019. Disponível em: [https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/244871/1/BIRNDL2019\\_123.pdf](https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/244871/1/BIRNDL2019_123.pdf)

PERONI, S.; SHOTTON, D. OpenCitations, an infrastructure organization for open scholarship. **Quantitative Science Studies**, v.1, n. 1, p. 428–444, 2020 Disponível em: [https://doi.org/10.1162/qss\\_a\\_00023](https://doi.org/10.1162/qss_a_00023)

SCHIERMEIER, Q. Initiative aims to break science's citation paywall. **Nature**, 2017. Disponível em: <https://doi.org/10.1038/nature.2017.21800>

SOCIEDADE BRASILEIRA DE MEDICINA TROPICAL. **Revista da Sociedade Brasileira de Medicina Tropical**: sobre o periódico. 2023. Disponível em: <https://www.scielo.br/journal/rsbmt/about/#about>

VAN ECK, N.J.; WALTMAN, L. Accuracy of citation data in Web of Science and Scopus. **Anais...** Proceedings of the 16th International Conference of the International Society for Scientometrics and Informetrics, p. 1087-1092, 2019. Disponível em: <https://doi.org/10.48550/arXiv.1906.07011>

VISSER, M.; VAN ECK, N.; 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](https://doi.org/10.1162/qss_a_00112)