



ATRAÇÃO ENTRE PERIÓDICOS BRASILEIROS DE MEDICINA:

análise a partir de dados de citação do *OpenAlex*

Washington Luís Ribeiro de Carvalho Segundo

 <https://orcid.org/0000-0003-3635-9384>.

✉ washingtonsegundo@ibict.br.

🏢 Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) |

ROR:: <https://ror.org/006c42y96> | Brasília, Brasil.

Fabio Lorensi do Canto

 <https://orcid.org/0000-0002-8338-1931>.

✉ fabio.lc@ufsc.br.

🏢 Universidade Federal de Santa Catarina (UFSC) |

ROR:: <https://ror.org/041akq887> | Florianópolis, Brasil.

Adilson Luiz Pinto

 <https://orcid.org/0000-0002-4142-2061>.

✉ adilson.pinto@ufsc.br.

🏢 Universidade Federal de Santa Catarina (UFSC) |

ROR:: <https://ror.org/041akq887> | Florianópolis, Brasil.

Daniel Sunfeld

 <https://orcid.org/0000-0002-5147-3698>.

✉ danielsunfeld@gmail.com.

🏢 Universidade de Brasília (UnB) |

ROR:: <https://ror.org/02xftp8v59> | Brasília, Brasil.

Eixo temático: Mapas e Redes na Ciência

Modalidade: Resumo expandido

DOI: 10.22477/ix.ebbc.378

Resumo: Apresenta uma rede de citação entre periódicos construída com base no índice de atração. Foram extraídos do *OpenAlex* dados de artigos publicados entre 2018 e 2022 em periódicos brasileiros de medicina com índice-h maior que 10. Os dados de citação desse conjunto foram processados em uma infraestrutura em nuvem, calculando-se a frequência de citação entre dois periódicos e o total de citações recebidas individualmente. Foi construída uma rede baseada no índice de atração entre as publicações. Periódicos de maior impacto e com escopo temático mais amplo possuem maior centralidade; periódicos especializados se atraem mutuamente, formando agrupamentos temáticos.

Palavras-Chave: Análise de citação. Redes de citação. Índice de atração. Periódicos brasileiros - medicina.



1 INTRODUÇÃO

A análise de citação ocupa lugar de destaque no contexto dos estudos métricos (Hou, 2017). Envolve a análise dos padrões de citação em conjuntos de documentos, como artigos, periódicos, teses e dissertações. Possibilita uma visão mais precisa do nível de influência e de impacto entre publicações científicas em diferentes contextos, tais como áreas do conhecimento, instituições e países (Waltman, 2016).

Os resultados de estudos dessa natureza podem orientar a construção de indicadores, rankings acadêmicos, redes de citação e de co-citação, entre outros elementos relevantes no contexto da avaliação científica em larga escala. Dados de citação também são utilizados em processos de avaliação institucional, como concursos e processos seletivos, ingresso em programas de pós-graduação e concessão de bolsas e financiamentos de pesquisa (Ding *et al.*, 2020).

A maioria dos estudos baseados em análise de citação utilizam dados da *Web of Science* ou da *Scopus*, bases de dados comerciais com cobertura internacional e multidisciplinar (Waltman, 2016). Nos últimos anos, contudo, o lançamento da *OpenAlex* vem possibilitando uma perspectiva aberta de estudos bibliométricos e cientométricos. Entre as inúmeras vantagens desta fonte inovadora se destacam a ampla cobertura, o acesso aberto e a possibilidade de extração ilimitada de dados por meio de data dumps públicos com atualização mensal e de uma interface web (Priem *et al.*, 2022).

Diante desse novo cenário firmado na análise de dados científicos *FAIR* (acrônimo em inglês para dados localizáveis, acessíveis, interoperáveis e reutilizáveis) (Wilkinson *et al.*, 2016), este estudo apresenta os resultados preliminares da construção de uma rede de citação entre periódicos realizada no âmbito do projeto *Laguna Cloud* do IBICT (Carvalho Segundo *et al.*, 2023).

Foi selecionada uma amostra de periódicos brasileiros de medicina indexados na OpenAlex. Com base nos dados de citação entre essas publicações, surge a seguinte pergunta de pesquisa: como o índice de atração influencia a construção de uma rede de citação entre periódicos brasileiros de medicina?

O objetivo deste estudo, portanto, é o de analisar a influência do índice de atração na construção de uma rede de citações entre periódicos brasileiros de medicina. Partindo de uma primeira noção do efeito de atração entre publicações apresentada por Gorbea-Portal (2005), o índice proposto é obtido pelo quadrado da frequência com que dois periódicos se citam mutuamente em relação ao quadrado do total de citações recebidas por cada um, multiplicado pela medida de proporção entre o total de citações dos periódicos comparados.

Pretende-se, a partir desse experimento inicial, construir um modelo mais amplo, que permita mapear padrões de citação entre publicações e identificar discrepâncias, ou seja, comportamentos outliers que possam indicar más práticas de citação. Entre as más práticas mais conhecidas incluem-se a citação cruzada e o empilhamento de citações (citações deliberadas entre duas ou mais publicações) e o excesso de autocitações (um percentual de autocitação significativamente acima da média para a respectiva área) (Ioannidis, 2015; Taşkin *et al.*, 2021).

2 PROCEDIMENTOS METODOLÓGICOS

Foi realizada a carga do conjunto de dados de trabalhos (*works*) do *OpenAlex* em uma infraestrutura em nuvem da *Amazon Web Services (AWS)*, com gerenciamento por meio de uma arquitetura serverless e x86_64 e processamento de dados em *cluster Spark* com ferramenta EMR.

Foram processados dados de mais de 200 milhões de trabalhos de aproximadamente 150 mil periódicos, identificando os *OpenAlex_Source_ID* (identificador persistente de fontes), as ligações entre periódicos citantes e citados e o número de citações entre eles. Foram aplicados filtros por tipo de documento artigo (*article*) e por ano (*publication_year*) entre 2018 e 2022.

Considerando as limitações de espaço e o caráter experimental deste trabalho, foi aplicado um filtro para periódicos brasileiros indexados com o conceito "*Medicine*" do *Wikidata*¹ e com índice-h da *OpenAlex* superior a 10. A escolha da área de medicina se deu em razão de seu alto impacto e relevância no cenário da editoração científica nacional.

Os dados foram tabulados em dois arquivos .csv, sendo um referente às arestas (registros de citações entre periódicos A e B) e o outro representando os nós (dados dos periódicos).

Primeiramente foi calculada a distância d entre dois periódicos aplicando-se a Fórmula 1:

$$d_{AB} = \frac{C_{\overleftarrow{A}} + C_{\overleftarrow{B}}}{C_{\overrightarrow{AB}} + C_{\overrightarrow{BA}}}$$

Fórmula 1: distância d entre os periódicos A e B

na qual a distância d entre os periódicos A e B é igual ao total de citações recebidas por A mais o total de citações recebidas por B, dividido pelo número de citações de A para B somado ao número de citações de B para A.

Posteriormente foi calculado o índice de atração G entre os periódicos A e B, aplicando-se a Fórmula 2:

$$G_{AB} = \frac{C_{\overleftarrow{B}}/C_{\overleftarrow{A}}}{d_{AB}^2} = \frac{C_{\overleftarrow{B}}(C_{\overrightarrow{AB}} + C_{\overrightarrow{BA}})^2}{C_{\overleftarrow{A}}(C_{\overleftarrow{A}} + C_{\overleftarrow{B}})^2}$$

Fórmula 2: atração G entre os periódicos A e B

onde se assume que o total de citações de A é maior ou igual ao total de citações recebido por B. Tem-se que o índice de atração G entre os periódicos A e B é igual ao total de citações recebidas por B dividido pelo total de citações recebidas por A, dividido pela distância d entre A e B ao quadrado. Resulta que, a atração G entre A e B é dada pela razão do total de citações de B sobre o total de citações de A, vezes o

¹ <https://www.wikidata.org/wiki/Q111190>.



quadrado da razão da soma das citações de *A* para *B*, e *B* para *A*, e a soma do total de citações de *A* e o total de citações de *B*.

A partir do índice de atração entre todos os periódicos da amostra foi construída uma rede não direcional no *software Gephi*. Foi aplicada a distribuição *Force Atlas 2* e atribuído tamanho aos nós com base na soma de todas as citações recebidas por um periódico dos demais presentes na rede.

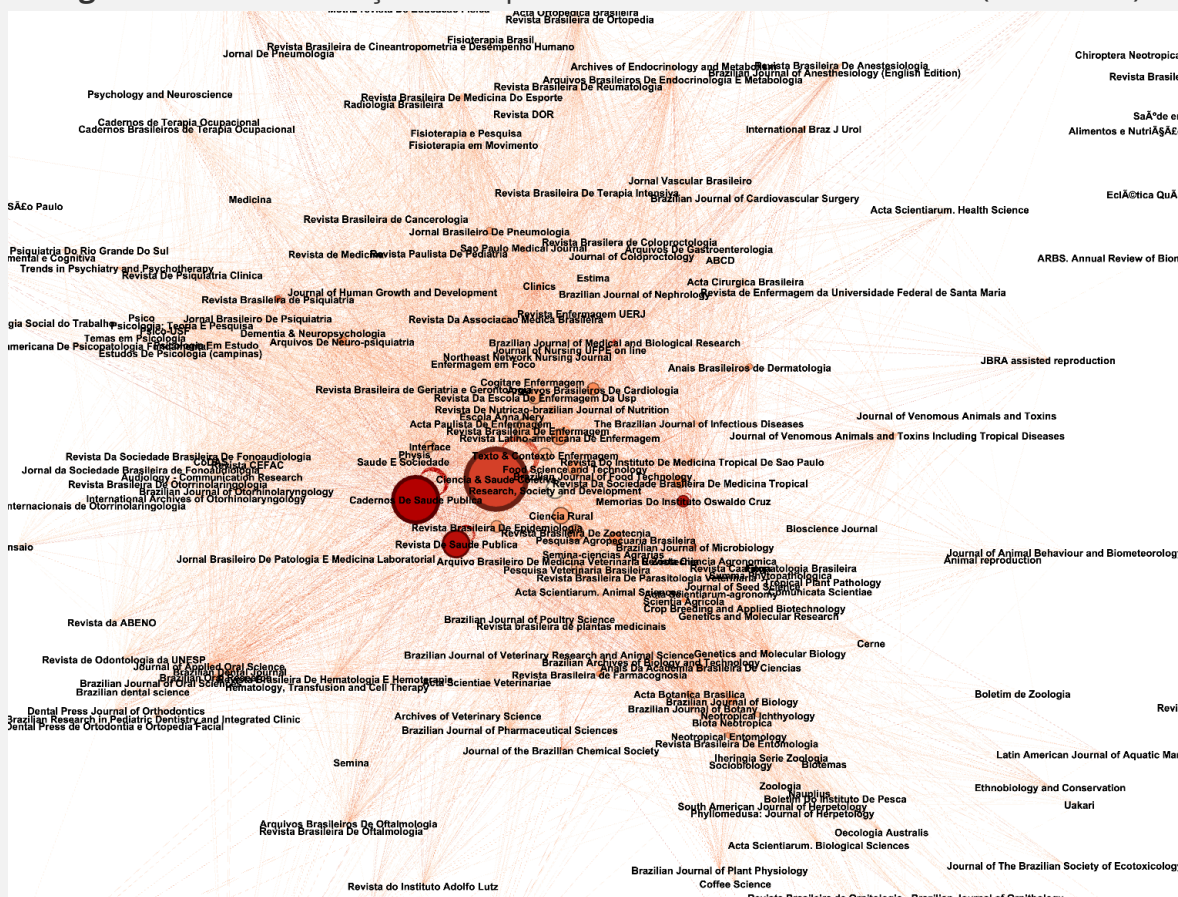
3 RESULTADOS

A Figura 1 apresenta a visão central da rede de atração entre os 273 periódicos de medicina com índice *h* superior a 10 indexados no *OpenAlex*. Em posição de centralidade localizam-se os periódicos com mais citações recebidas e de escopo temático mais abrangente, tais como saúde pública e coletiva e medicina geral e tropical.

Observa-se que em razão da maior frequência de citações recebidas, as publicações de maior impacto exercem força de atração em um maior conjunto mais amplo de publicações, tanto de áreas e especialidades de Ciências da Saúde, tanto de outras áreas do colégio de Ciências da Vida, como Ciências Agrárias e Biológicas.

Por outro lado, é possível identificar forte atração entre periódicos de mesmo escopo, representados na rede por clusters temáticos.

Figura 1 - Rede de atração entre periódicos brasileiros de medicina (2018-2022)



Fonte: Elaborado pelos autores (2024).

Nesse sentido, a Figura 2 mostra em destaque o cluster de periódicos de enfermagem, que se atraem mutuamente em meio a publicações de outras temáticas.

Figura 2 - Visão central da rede e de um agrupamento de periódicos em enfermagem

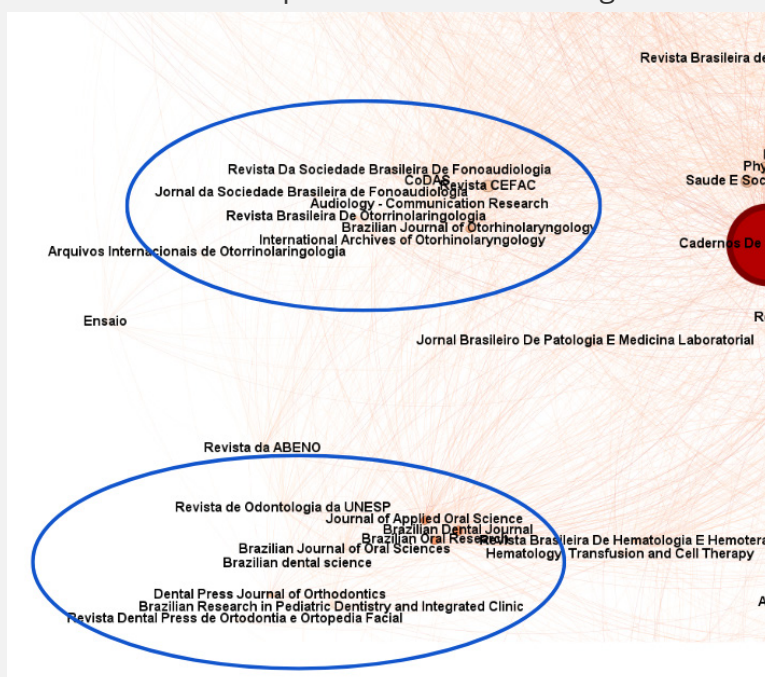


Fonte: Elaborado pelos autores (2024).

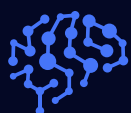
Nota-se, contudo, que os periódicos de enfermagem também exercem atração em títulos de outras especialidades da saúde, como cardiologia, geriatria e nutrição, sugerindo algum nível de conexão temática entre artigos publicados.

A Figura 3 mostra em detalhe outros dois clusters especializados, sendo um formado por periódicos de fonoaudiologia e otorrinolaringologia e outro, formado a partir da atração de periódicos de odontologia e de medicina oral.

Figura 3 - Visão clusters de periódicos de odontologia e fonoaudiologia



Fonte: Elaborado pelos autores (2024).



A proximidade entre os nós que compõem esses agrupamentos representa uma expressiva força de atração entre os periódicos. Por outro lado, a localização periférica e o relativo distanciamento de outros nós na rede sugerem menor força de atração em relação às publicações de outras áreas e especialidades.

Foram identificados ainda outros dois *clusters* na rede, sendo um formado pela aproximação de periódicos com aderência às áreas de Ciências Agrárias e outro à área de Ciências Biológicas. Esse fenômeno sugere que a vertente interdisciplinar das Ciências Médicas atraí publicações de outras áreas, que neste caso encontram características em comum a partir do colégio de Ciências da Vida da classificação da CAPES.

4 CONSIDERAÇÕES FINAIS

O cálculo do índice de atração oferece uma visão geral da variação do padrão de citação entre periódicos de uma mesma área ou áreas correlacionadas. Neste estudo, essa noção foi obtida por meio da análise de publicações de escopo geral e de especialidades dentro das Ciências Médicas.

Enquanto periódicos de amplo escopo tendem a atrair citações de variadas subáreas e especialidades médicas, periódicos altamente especializados exercem atração de citações majoritariamente de periódicos do mesmo escopo. Além disso, o maior nível de impacto reflete na posição centralidade da rede, considerando a força de atração exercida em publicações de diferentes áreas.

A rede tem caráter experimental e deve ser ampliada, tanto em relação ao universo de periódicos analisados como em relação ao número de métricas de rede calculadas. Como possibilidade de estudos futuros, pretende-se ampliar a análise de rede para todos os periódicos listados na última avaliação Qualis da CAPES, relevantes no contexto nacional. Pensando em uma perspectiva regional, pretende-se também analisar periódicos de países da América Latina.

Além do universo de publicações, a ampliação da pesquisa deverá abranger também o cálculo de outras métricas de rede, incluindo além do impacto de citação, variáveis como modelo de acesso, origem das publicações, idiomas, vinculação institucional e fontes de financiamento.

AGRADECIMENTOS

A FUNDEP pelo financiamento.

REFERÊNCIAS

CARVALHO SEGUNDO, Washington Luís Ribeiro de. *et al.* Projeto Laguna: infraestrutura de um lago de dados científicos em acesso aberto. **BiblioCanto**, v. 9, n. 2, p. 133–138, 2023. DOI: <https://doi.org/10.21680/2447-7842.2023v9n2ID33825>. Disponível em: <https://periodicos.ufrn.br/bibliocanto/article/view/33825>. Acesso em: 02 fev. 2024.



DING, Jingda; LIU, Chao.; KANDONGA, Goodluck Asobenie. Exploring the limitations of the h-index and htype indexes in measuring the research performance of authors. **Scientometrics**, v. 122, n. 3, p. 1303–1322, 2020. DOI: <https://doi.org/10.1007/s11192-020-03364-1>. Disponível em: <https://link.springer.com/article/10.1007/s11192-020-03364-1>. Acesso em: 02 fev. 2024.

GORBEA-PORTAL, Salvador. GORBEA PORTAL, SALVADOR. **Modelo teórico para el estudio métrico de la información documental**. Madrid: TREA, 2005.

HOU, Jianhua. Exploration into the evolution and historical roots of citation analysis by referenced publication year spectroscopy. **Scientometrics**, Dordrecht, v. 110, n. 3, p. 1437-1452, 2017. DOI: <https://doi.org/10.1007/s11192-016-2206-9>. Disponível em: <https://link.springer.com/article/10.1007/s11192-016-2206-9>. Acesso em: 02 fev. 2024.

IOANNIDIS, John. A generalized view of self-citation: Direct, co-author, collaborative, and coercive induced self-citation. **Journal of Psychosomatic Research**, v. 78, n. 1, p. 7-11, 2015. DOI: <https://doi.org/10.1016/j.jpsychores.2014.11.008>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0022399914003882?via%3Dihub>. Acesso em: 02 fev. 2024.

PRIEM, Jason. *et al.* OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. **ArXiv**. DOI: <https://doi.org/10.48550/arXiv.2205.01833>. Disponível em: <https://arxiv.org/abs/2205.01833>. Acesso em: 10 mar. 2024.

TAŞKIN, Zehra. *et al.* Self-Citation Patterns of Journals Indexed in the Journal Citation Reports. **Journal of Informetrics**, v. 15, n. 4, 2021. DOI: <https://doi.org/10.1016/j.joi.2021.101221>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157721000924>. Acesso em: 10 mar. 2024.

WALTMAN, Ludo. A review of the literature on citation impact indicators. **Journal of Informetrics**, v. 10, n. 2, p. 365–391, maio 2016. DOI: <http://dx.doi.org/10.1016/j.joi.2016.02.007>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1751157715300900?via%3Dihub>. Acesso em: 10 mar. 2024.

WILKINSON, Mark. D. *et al.* The fair guiding principles for scientific data management and stewardship. **Scientific data**, Nature Publishing Group, v. 3, n. 1, p. 1–9, 2016. DOI: <https://doi.org/10.1038/sdata.2016.18>. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso em: 08 mar. 2024.