



# COBERTURA DE CITAÇÕES DA OPENALEX, DA SCOPUS E DA WEB OF SCIENCE:

análise comparativa a partir da produção científica da Universidade Federal de Santa Catarina

## Fabio Lorensi do Canto

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

 fabio.lc@ufsc.br.


 Universidade Federal de Santa Catarina (UFSC) |

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

## 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) |

 <https://ror.org/02263ky35> | Brasília, Brasil.

## Adilson Luiz Pinto

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


 adilson.pinto@ufsc.br.


 Universidade Federal de Santa Catarina (UFSC) |

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

## Patricia da Silva Neubert

 <https://orcid.org/0000-0002-8909-1898>.

 patricia.neubert@ufsc.br.

 Universidade Federal de Santa Catarina (UFSC) |

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

---

**Eixo temático:** Métodos Bibliométricos e de citação

**Modalidade:** Pecha Kucha

**DOI:** 10.22477/ix.ebbc.321

**Resumo:** Analisa a cobertura de citações da *OpenAlex*, da *Scopus* e da *Web of Science* a partir dos artigos da UFSC (2018-2022). Foram analisados: a) todos os artigos; b) artigos com DOI e; c) artigos com DOI recuperados em mais de uma fonte. A *Web of Science* registra mais citações, tem a maior média por artigo e o menor percentual de não citados. Entre artigos com DOI em mais de uma fonte, os resultados foram similares. O *OpenAlex* é a fonte mais ampla, mas por sua característica de repositório, inclui também publicações de menor impacto.

**Palavras-Chave:** Análise de citação. Bases de dados científicas. Bibliometria.



## 1 INTRODUÇÃO

O lançamento do *OpenAlex* representou uma nova perspectiva para a bibliometria e a cientometria internacional (Priem; Piwowar; Orr, 2022). A ampla cobertura, o acesso aberto e a disponibilização de dados e de indicadores de diversas entidades científicas vêm possibilitando a realização de estudos antes possíveis apenas por meio de bases de dados fechadas, notadamente *Scopus* e *Web of Science* (WoS) (Waltman, 2016).

No contexto da análise de citação (Hou, 2017), o *OpenAlex* tem maior cobertura global de citações em comparação a outras fontes científicas. Segundo dados divulgados em sua documentação, seriam 248M de trabalhos e 1.9B de citações diante de 90M/1.8B da WoS e 89M/1.8B da *Scopus* 140M/1.7B da *Dimensions* e 145M/1.45B da *CrossRef* (Priem; Piwowar; Orr, 2022).

Como esses números oferecem apenas uma visão geral, é relevante analisar a questão a partir de um conjunto delimitado de publicações, que permitam uma comparação direta entre as principais fontes bibliométricas. Este estudo propõe uma análise da cobertura de citações a partir de um conjunto de artigos publicados em um período de cinco anos por pesquisadores com vínculo institucional com a Universidade Federal de Santa Catarina (UFSC).

A escolha da UFSC se deu em razão de sua significativa produção científica, que lhe confere a 7ª posição no *Leiden Ranking 2023*<sup>1</sup> entre universidades brasileiras. Além disso, mantém cursos de graduação e programas de pós-graduação em diversas áreas do conhecimento, o que garante a variabilidade temática da amostra de publicações (Universidade Federal de Santa Catarina, 2022).

Essa pesquisa adquire relevância diante da necessidade de mudanças no modelo de avaliação da atividade científica nacional, atualmente um modelo baseado majoritariamente em indicadores de produção científica produzidos por plataformas comerciais, para que, gradativamente incorpore dados e indicadores obtidos em sistemas alinhados à Ciência Aberta, com recursos de software livre e subsidiados com conjuntos de dados FAIR (*Findable, Accessible, Interoperable and Reusable*) (Wilkinson et al., 2016).

## 2 METODOLOGIA

Foram recuperados artigos científicos e de revisão publicados entre 2018 e 2022 com registro de vínculo institucional dos autores com a UFSC na *OpenAlex*, na *Scopus* e na *Web of Science Core Collection*. No *OpenAlex* e na WoS há tratamento de entrada por nome da instituição, tendo sido utilizado o recurso de autopreenchimento do termo de busca. Na *Scopus* foi utilizada a *string* “universidade federal de santa catarina” OR UFSC.

Os dados foram extraídos em formato ‘csv’, incluindo os campos DOI e número de citações recebidas. Foi realizado o cruzamento de todos os registros de DOI, identificando os artigos recuperados em mais de uma base de dados. Os dados foram analisados a partir de três perspectivas: a) todos os artigos

---

1 <https://www.leidenranking.com/ranking/2023/list>

recuperados; b) artigos com DOI e; c) artigos com DOI registrados em mais de uma fonte.

### 3 RESULTADOS

A Tabela 1 mostra o resultado da análise de cobertura de citações nas três fontes considerando todos os artigos recuperados (com e sem DOI). A cobertura da *OpenAlex* em relação à produção da UFSC é significativamente mais ampla no período, com mais que o dobro de artigos recuperados em comparação com a *Scopus* e a *WoS*. O percentual mais elevado de artigos não citados e a menor média de citação por artigo reflete a sua característica de repositório e sua política inclusiva, que indexa publicações de diferentes níveis de impacto, áreas do conhecimento, idiomas e regiões.

**Tabela 1** - Citações de artigos da UFSC por fonte (2018-2022)

	OpenAlex	Scopus	WoS
Total de artigos	26.910	9.405	12.476
Total de citações	145.206	85.150	149.056
Média citação/artigo	5,4	9,1	12
Artigos não citados	12.384 (46%)	1.868 (19,8%)	2.312 (18,5%)

**Fonte:** Dados da pesquisa (2024).

Os dados mostram, por outro lado, a vantagem da *WoS* considerando o volume total de citações, a média de citação por artigo e o menor percentual de artigos não citados. Esse resultado possivelmente se deve em razão da política de indexação mais restrita da *WoS Core Collection*, que tem as publicações de alto impacto como prioridade da coleção.

A Tabela 2 mostra os resultados restritos aos artigos com DOI, que são mais de 90% do total de artigos recuperados na *OpenAlex* e mais de 95% na *Scopus* e na *WoS*. Os dados apresentam, portanto, pouca variação em relação à Tabela 1, com aumento sutil na média de citações por artigo e diminuição do percentual de não citados.

**Tabela 2** - Citações de artigos da UFSC com DOI por fonte (2018-2022)

	OpenAlex	Scopus	WoS
Total de artigos	24.262 (90,16%)	9.151 (97,3%)	12.029 (96,42%)
Total de citações	144.832	84.810	145.420
Média citação/artigo	5,97	9,26	12,9
Artigos não citados	10.181 (41,9%)	1.723 (18,8%)	1.999 (16,6%)

**Fonte:** Dados da pesquisa (2024).

A Tabela 3 mostra o total de citações de artigos com DOI recuperados em mais de uma fonte. Essa comparação a partir de conjuntos idênticos mostra uma vantagem geral da *Scopus* e da *OpenAlex* em relação à *WoS*.



**Tabela 3** - Citações de artigos da UFSC com DOI recuperados em mais de uma fonte

nº artigos	OpenAlex	Scopus	WoS
7704	73.504	74.282	X
10.361	117.668	X	104.502
6142	X	67.869	60.117
<b>5374</b>	<b>58.962</b>	<b>59.354</b>	<b>52.585</b>

**Fonte:** Dados da pesquisa (2024).

Por outro lado, a análise de correlação de Pearson entre os conjuntos de citações recebidas pelos artigos com DOI presentes em mais de uma fonte foi de 0.99 nas três combinações (*OpenAlex* x *Scopus*, *OpenAlex* x *WoS* e *Scopus* x *WoS*). Esse resultado sugere uma variação insignificante no número de citações registradas na análise comparativa artigo por artigo.

## 4 CONSIDERAÇÕES FINAIS

A cobertura de citações aos artigos analisados reflete a política de indexação de cada fonte. A *WoS*, com seu enfoque em periódicos de prestígio apresenta números gerais expressivos. No entanto, quando comparada com a *OpenAlex* e a *Scopus* a partir de artigos em comum, as diferenças entre as fontes são estatisticamente insignificantes. Estes dados sugerem que os indicadores de citação de *OpenAlex* podem ser utilizados para obtenção de indicadores da produção científica, contribuindo para minimizar a dependência de indicadores produzidos por editores e bases de dados comerciais, contribuindo com um modelo de avaliação mais aberto.

Esse resultado pode corroborar uma hipótese mais ampla, de que o *OpenAlex* pode substituir bases de dados científicas mantidas por assinatura, pois oferece recursos de avaliação de impacto equivalentes, com uma cobertura mais ampla, aberta e acesso ilimitado à dados FAIR.

## REFERÊNCIAS

HOU, J. 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.

PRIEM, J.; PIWOWAR, H.; ORR, R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. In: INTERNATIONAL CONFERENCE ON SCIENCE AND TECHNOLOGY INDICATORS, 26., 2022, Granada. **Anais [...]**. Granada: arXiv, 2022. DOI: <https://doi.org/10.48550/arXiv.2205.01833>. Disponível em: <https://arxiv.org/abs/2205.01833>. Acesso em: 08 mar. 2024.

UNIVERSIDADE FEDERAL DE SANTA CATARINA. Pró-reitoria de Pesquisa e Inovação. **Indicadores de produção científica**. Florianópolis, 2022. Disponível em: <https://propesq.ufsc.br/indicadores-de-producao-cientifica>. Acesso em: 08 mar. 2024.



WALTMAN, L. 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: 08 mar. 2024.

WILKINSON, M. D. *et al.* The fair guiding principles for scientific data management and stewardship. **Scientific data**, 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.