



# 10° EBBC

Encontro Brasileiro de  
Bibliometria e Cientometria



# A UFRJ E O ACESSO ABERTO NO SÉCULO XXI: uma análise com base em dados abertos da *OpenAlex*

**Gabriel Alves Vieira**

<https://orcid.org/0000-0002-5529-7628> | [gabriel.vieira@bioqmed.ufrj.br](mailto:gabriel.vieira@bioqmed.ufrj.br)  
Universidade Federal do Rio de Janeiro (UFRJ)

**Jacqueline Leta**

<https://orcid.org/0000-0002-3271-7749> | [jleta@bioqmed.ufrj.br](mailto:jleta@bioqmed.ufrj.br)  
Universidade Federal do Rio de Janeiro (UFRJ)

## Resumo:

Este trabalho analisa o volume e impacto da produção científica da UFRJ (2001-2025) na OpenAlex. Dos 129.213 artigos recuperados, 62,8% são de acesso aberto (AA), superando o modelo fechado desde 2007. Apesar das publicações de acesso restrito apresentarem medianas de Field-Weighted Citation Impact (FWCI) superiores, o AA domina a elite de publicações com mais citações, sugerindo um efeito de auto-seleção em vez de causalidade direta. O estudo reforça a viabilidade de fontes bibliográficas abertas e demonstra que, apesar da maior visibilidade do AA, documentos sem nenhuma citação ainda são frequentes, desafiando a ideia de vantagem de citação garantida.

**Palavras-Chave:** Acesso aberto. Ciência aberta. OpenAlex. Linguagem R. UFRJ.

## EIXO TEMÁTICO:

Acesso Aberto, Ciência Aberta e Dados Abertos

## MODALIDADE:

Resumo expandido

**DOI:** 10.22477/x.ebbc.1109

## COMO CITAR ESTE ARTIGO:

VIEIRA, Gabriel Alves; LETA, Jacqueline. A UFRJ e o acesso aberto no século XXI: uma análise com base em dados abertos da OpenAlex. In: EBBC - Encontro Brasileiro de Bibliometria e Cientometria. 10. **Anais [...]**. Curitiba, 2026. DOI: 10.22477/x.ebbc.1109. Disponível em: <https://doi.org/10.22477/x.ebbc.1109>.



## 1 INTRODUÇÃO

A ciência aberta pode ser definida como “os esforços de pesquisadores, governos, agências de financiamento à pesquisa ou da própria comunidade científica para tornar os principais produtos dos resultados de pesquisas financiadas com recursos públicos publicamente acessíveis em formato digital” (OECD, 2015). Tal tema é muito amplo e pode ser abordado em virtualmente qualquer faceta do processo de produção científica, como dados, código, protocolos e revisão por pares. Dentro deste contexto, uma dimensão da ciência aberta que vem ganhando cada vez mais destaque na literatura é a do acesso aberto às publicações científicas (TENNANT *et al.*, 2016; PIWOWAR *et al.*, 2018).

Com o advento e evolução da Internet, a publicação de artigos científicos em meio digital passou a se popularizar. Isso leva à substituição paulatina do modelo tradicional de publicação em meio impresso, caracterizado pelo acesso restrito por assinatura e alto custo material, o que, por sua vez, resulta no movimento de acesso aberto das produções acadêmicas (PIWOWAR *et al.*, 2018). Para além das óbvias vantagens em termos de alcance, se argumenta que a prática da publicação em acesso aberto causa um aumento no número de citações recebidas, ou seja, proporciona uma “vantagem de citação” a documentos publicados seguindo essa estratégia (EYSENBACH, 2006).

A despeito da quantidade expressiva de trabalhos sobre acesso aberto, raramente os dados analisados são disponibilizados publicamente. Em grande parte, isso se deve à natureza comercial de várias fontes secundárias de informação bibliográfica utilizadas nestas pesquisas. Scopus, Web of Science e Dimensions são exemplos de serviços que não permitem a distribuição pública de seus dados (WALTMAN & LARIVIÈRE, 2020). Entretanto, há fontes secundárias não comerciais

disponíveis ao público, que permitem o compartilhamento e redistribuição pública dos dados com pouca ou nenhuma restrição. Dentre as opções disponíveis, a OpenAlex (<https://openalex.org>) se destaca por ser completamente aberta (PRIEM et al., 2022).

Tendo isso em vista, o objetivo principal deste trabalho é identificar a adesão de uma instituição acadêmica ao modelo de publicação em acesso aberto, focando na relação entre modelo de publicação e as citações recebidas. A Universidade Federal do Rio de Janeiro (UFRJ) foi escolhida, seja porque é a universidade de vínculo dos autores deste trabalho, seja porque é a maior e mais antiga universidade pública mantida por recursos do governo federal no país (OLIVEIRA, 2019). Além disso, vale também destacar que a escolha se deve por motivos inerentes à própria pesquisa realizada nesta universidade, que é um dos fatores que a coloca em posição de prestígio nos mais diversos rankings. Por exemplo, ao se considerar ranqueamentos de instituições da América Latina, a UFRJ é a 6ª colocada na avaliação de 2024 da Quacquarelli Symonds e ocupa a 11ª posição no Times Higher Education de 2023.

## 2 METODOLOGIA

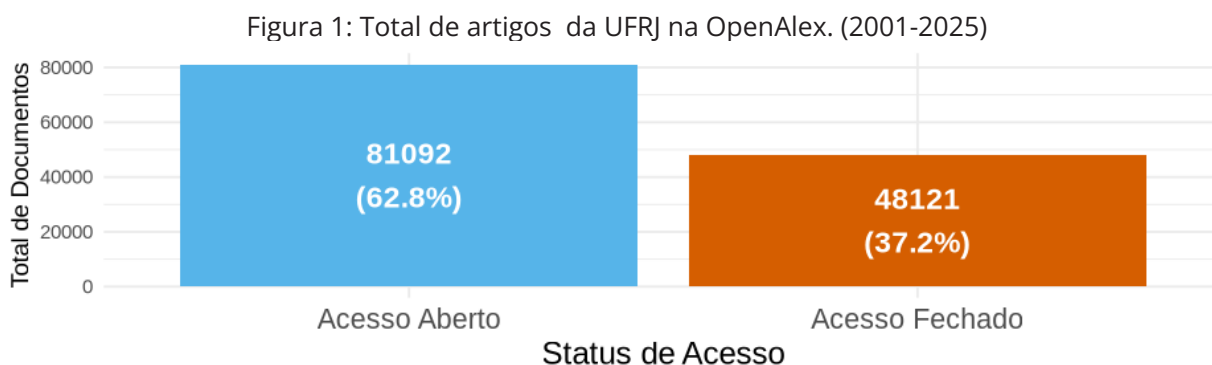
Os dados foram recuperados através da interface de programação de aplicações (API) da OpenAlex. O escopo do trabalho se restringe a artigos publicados em periódicos (i.e. artigos de pesquisa original e revisões) entre 2001 e 2025. A estratégia de busca identificou artigos com ao menos uma autoria filiada à UFRJ ou unidades vinculadas (e.g., Hospital Universitário Clementino Fraga Filho) por meio do parâmetro de filtragem *authorships.institutions.lineage*. Para garantir a integridade da coleta e mitigar riscos de interrupção no fluxo de transferência de grandes volumes de dados, as requisições foram segmentadas por ano de publicação. O

download dos conjuntos de dados foi realizado em 04/03/2026.

A recuperação, limpeza e visualização dos dados foi realizada através da linguagem R (R CORE TEAM, 2025).<sup>1</sup> Para além dos campos sobre acesso, recuperamos e analisamos o FWCI, um indicador que compara o número real de citações recebidas por um documento com o número esperado de citações para documentos do mesmo tipo, ano de publicação e disciplina (PURKAYASTHA *et al.*, 2019).

### 3 RESULTADOS E DISCUSSÃO

Um total de 129.213 documentos foram recuperados, dos quais 81.092 (62.8 %) são de acesso aberto e 48.121 (37.2%) de acesso fechado (Figura 1). Este resultado indica uma grande adesão aos veículos de acesso aberto na UFRJ durante o primeiro quarto do século XXI.

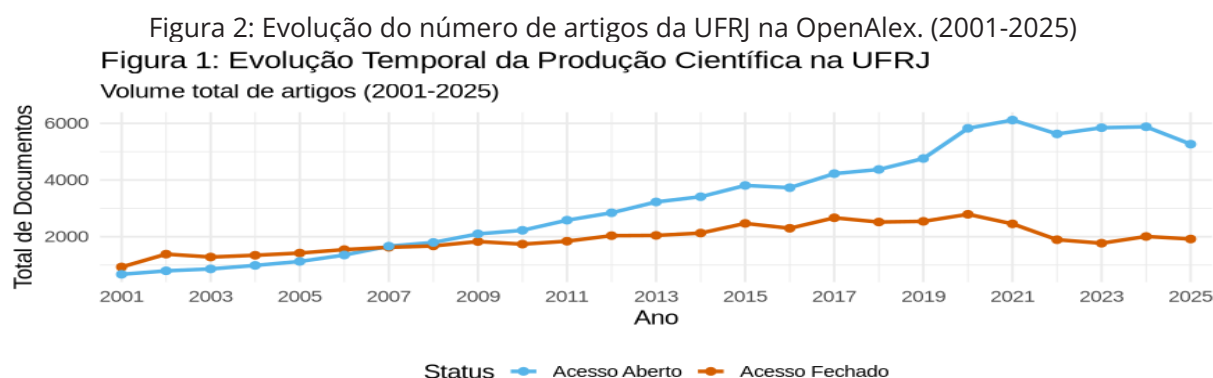


Fonte: Elaborado pelos autores

Ao avaliar a evolução anual do número de artigos (Figura 2), é possível observar que o volume de publicações em acesso fechado é consistentemente superior ao de acesso aberto no início do século. Entretanto, o número de publicações anuais abertas superou as fechadas a partir de 2007. Este padrão não é exclusivo da produção da UFRJ, sendo também observado, por exemplo, no PubMed Central, repositório no qual a adesão à publicação em acesso aberto cresceu vertiginosamente no início do século (TENNANT *et al.*, 2016).

<sup>1</sup> Todos os scripts utilizados estão disponíveis publicamente em [https://github.com/gavieira/ufrj\\_oa](https://github.com/gavieira/ufrj_oa)

A partir de 2010, o volume de publicações em acesso aberto passa por crescimentos anuais muito expressivos quando comparado ao de acesso restrito. Isso persiste até 2021, quando o número de publicações abertas oscila e mostra reduções relevantes. Entretanto, o volume de publicações de acesso restrito também é reduzido entre 2020 e 2022, o que, provavelmente, está relacionado à pandemia de COVID-19, durante a qual houve queda substancial da produtividade científica (RAYNAUD *et al.*, 2021).



Fonte: Elaborado pelos autores

Com relação ao impacto das publicações, o total de artigos publicados em acesso aberto da UFRJ apresenta um FWCI médio de 5,88 enquanto os de acesso restrito somam 4,68. Este resultado corrobora uma correlação entre publicações em acesso aberto e um maior número de citações recebidas (EYSENBACH, 2006).

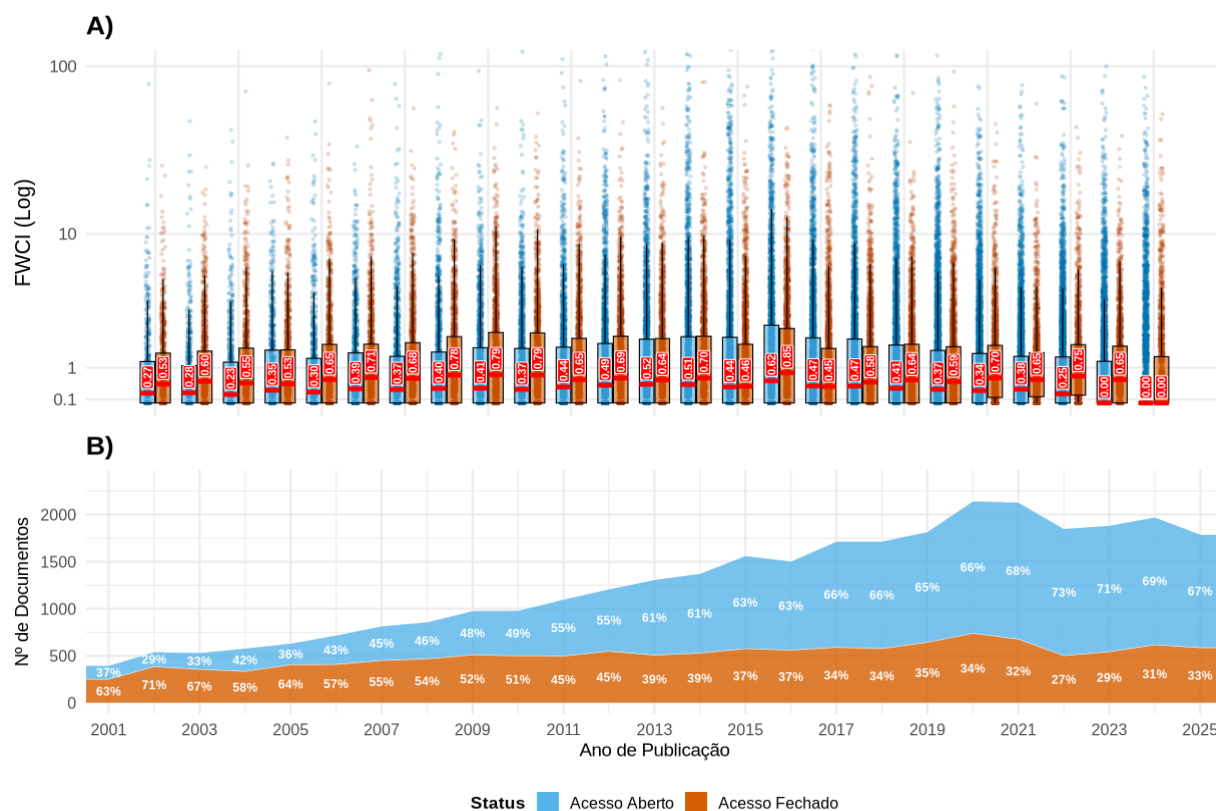
A Figura 3A mostra a distribuição dos valores de FWCI por ano e tipo de acesso nas publicações analisadas. Nela, temos tanto os pontos correspondentes aos artigos individuais quanto boxplots representando a distribuição dos dados, onde o valor central do conjunto de dados (mediana) é representado por uma linha vermelha, acima da qual temos seu valor numérico. Há uma quantidade considerável de *outliers* (i.e. artigos altamente citados) no conjunto de dados estudado, os quais podem facilmente distorcer análises baseadas em média e dificultar a visualização

. O conjunto de dados utilizado neste trabalho, assim como uma breve descrição de todos os campos recuperados, pode ser acessado em <https://zenodo.org/records/18883350>

dos dados por conta de escala. A escolha da mediana como medida de centralidade, assim como o uso da escala logarítmica, se deram para minimizar o impacto negativo desses *outliers* na interpretação e visualização dos dados. A mediana das distribuições, assim como suas variâncias, são predominantemente maiores em artigos publicados em acesso restrito. Após uma análise exploratória dos dados, confirmou-se que esse resultado é, em grande parte, oriundo dos documentos sem citações: artigos com FWCI igual a zero apresentam maior proporção no modelo de publicação aberto do que no restrito.

Por outro lado, a Figura 3B analisa apenas os 25% dos artigos com maior valor de FWCI de cada ano. Em outras palavras, foram analisadas apenas as publicações situadas no percentil 75 ou superior em termos de FWCI, considerando artigos publicados no mesmo ano. Nesta fração de documentos altamente citados, a proporção de artigos provenientes da via de publicação aberta é expressivamente maior nos anos mais recentes, o que explica o maior valor médio de FWCI para estes documentos e pode ser uma distorção causada por *outliers*. Logo, no conjunto de artigos da UFRJ, ao mesmo tempo em que a publicação em acesso aberto não garantiu o sucesso em termos de impacto, é ela que se destaca dentre a maior parte dos artigos. Isso vai de encontro à idéia de uma relação causal entre acesso aberto e número de citações, sendo mais compatível com um efeito de auto-seleção, no qual autores de artigos de alta qualidade apresentam preferência pela estratégia de publicação aberta (GAULÉ & MAYSTRE, 2011).

Figura 3: Artigos e estratégia de publicação da UFRJ na OpenAlex quanto ao (A) log da distribuição dos valores de FWCI e (B) artigos com maior impacto. (2001-2025)



Fonte: Elaborado pelos autores.

## 4 CONSIDERAÇÕES FINAIS

Este estudo piloto foca no perfil de publicação de artigos da UFRJ com relação à estratégia de publicação escolhida, ao longo do século XXI. Os resultados mostram que a instituição apresentou um crescimento substancial em sua adesão à publicação em acesso aberto. Além disso, temos que, enquanto a mediana do FWCI é superior para publicações em acesso restrito, a fração de artigos com maior impacto é majoritariamente composta por documentos publicados em acesso aberto, o que pode indicar um viés autoral de publicação em vez de uma relação causal.

Em suma, este é o primeiro de uma série de estudos focados nas estratégias de

publicação de universidades brasileiras, nos quais pretendemos utilizar dados da OpenAlex na busca de compreender melhor, especialmente, como se dá a relação entre o tipo de estratégia e o impacto. Importante destacar que tanto a fonte secundária quanto as ferramentas do ecossistema R utilizadas neste projeto são completamente abertas, o que permite compartilhar não apenas os resultados em forma de publicações, mas subprodutos da pesquisa como os dados processados e scripts gerados. Assim, esperamos não apenas abordar com profundidade o tema do acesso aberto, como também contribuir para uma ciência mais transparente, aberta e reprodutível.

## AGRADECIMENTOS

Este estudo conta com recursos do Edital Universal CNPq 2025 (Processo nº 403050/2025-1). O autor agradece à Fundação COPPETEC pelo apoio através da concessão de bolsa de Pós-Doutorado.

## REFERÊNCIAS

EYSENBACH, Gunther. Citation Advantage of Open Access Articles. **PLOS Biology**, [s. l.], v. 4, n. 5, p. e157, 16 maio 2006. ISSN 1545-7885. DOI [10.1371/journal.pbio.0040157](https://doi.org/10.1371/journal.pbio.0040157). Acesso em: 02 jun. 2026.

GAULÉ, Patrick; MAYSTRE, Nicolas. Getting cited: Does open access help? **Research Policy**, [s. l.], v. 40, n. 10, p. 1332-1338, 1 dez. 2011. ISSN 0048-7333. DOI [10.1016/j.respol.2011.05.025](https://doi.org/10.1016/j.respol.2011.05.025). Acesso em: 02 jun. 2026.

OECD. Making Open Science a Reality. Paris: OECD Publishing, 2015. (OECD Science, Technology and Industry Policy Papers, n. 25). DOI: <http://dx.doi.org/10.1787/5jrs2f963zs1-en>. Acesso em: 02 jun. 2026.

OLIVEIRA, Antonio José Barbosa de. **Uma breve história da UFRJ**. Conexão UFRJ, 27 set. 2019. Disponível em: <https://conexao.ufrj.br/2019/09/uma-breve-historia-da-ufrj/>. Acesso em: 15 mar. 2026.

PRIEM, Jason; PIWOWAR, Heather; ORR, Richard. **OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts**. [S. l.]: arXiv, 17 jun. 2022. arXiv:2205.01833 [cs]. DOI: [10.48550/arXiv.2205.01833](https://doi.org/10.48550/arXiv.2205.01833). Disponível em: <http://arxiv.org/abs/2205.01833>. Acesso em: 12 mar. 2026.

PURKAYASTHA, Amrita; PALMARO, Eleonora; FALK-KRZESINSKI, Holly J.; BAAS, Jeroen. Comparison of two article-level, field-independent citation metrics: Field-Weighted Citation Impact (FWCI) and Relative Citation Ratio (RCR). **Journal of Informetrics**, [s. l.], v. 13, n. 2, p. 635–642, 1 maio 2019. ISSN 1751-1577. DOI: [10.1016/j.joi.2019.03.012](https://doi.org/10.1016/j.joi.2019.03.012). Acesso em: 02 jun. 2026.

PIWOWAR, Heather et al. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. **PeerJ**, [s. l.], v. 6, p. e4375, 13 fev. 2018. ISSN 2167-8359. DOI: [10.7717/peerj.4375](https://doi.org/10.7717/peerj.4375). Acesso em: 02 jun. 2026.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2025. Disponível em: <https://www.R-project.org/>. Acesso em: 02 jun. 2026.

RAYNAUD, Marc et al. Impact of the COVID-19 pandemic on publication dynamics and non-COVID-19 research production. **BMC Medical Research Methodology**, [s. l.], v. 21, n. 1, p. 255, 22 nov. 2021. ISSN 1471-2288. DOI [10.1186/s12874-021-01404-9](https://doi.org/10.1186/s12874-021-01404-9). Acesso em: 02 jun. 2026.

TENNANT, Jonathan P.; WALDNER, François; JACQUES, Damien C.; MASUZZO, Paola; COLLISTER, Lauren B.; HARTGERINK, Chris. H. J. The academic, economic and societal impacts of Open Access: an evidence-based review. **F1000Research**, [s. l.], v. 5, p. 632, 21 set. 2016. ISSN 2046-1402. DOI [10.12688/f1000research.8460.3](https://doi.org/10.12688/f1000research.8460.3). Acesso em: 02 jun. 2026.

WALTMAN, Ludo; LARIVIÈRE, Vincent. Special issue on bibliographic data sources. **Quantitative Science Studies**, [s. l.], v. 1, n. 1, p. 360–362, 1 fev. 2020. ISSN 2641-3337. DOI [10.1162/qss\\_e\\_00026](https://doi.org/10.1162/qss_e_00026). Acesso em: 02 jun. 2026.