

DADOS ABERTOS E BIBLIOTECAS: ANÁLISE BIBLIOMÉTRICA DA PRODUÇÃO CIENTÍFICA (2015–2025)

OPEN DATA AND LIBRARIES: A BIBLIOMETRIC ANALYSIS OF SCIENTIFIC PRODUCTION (2015–2025)

JULIANA ARETZ CUNHA DE QUEIROZ AFONSO DETONI

Universidade de Brasília (UnB)

ORCID: <https://orcid.org/0000-0003-3916-7548>

ROR: <https://ror.org/02xjp8v59>

FELIPE AUGUSTO ARAKAKI

Universidade de Brasília (UnB)

ORCID: <https://orcid.org/0000-0002-3983-2563>

ROR: <https://ror.org/02xjp8v59>

MODALIDADE: RESUMO EXPANDIDO

Resumo: este estudo apresenta uma análise bibliométrica da produção científica sobre dados abertos em bibliotecas, no período de 2015 a 2025. O objetivo é mapear e analisar o volume e a distribuição temporal da produção científica sobre o tema, identificando lacunas e tendências que subsidiem etapas investigativas futuras. A coleta foi realizada em bases nacionais e internacionais, com buscas em português e inglês, com refinamento temporal. Os resultados evidenciam diferenças expressivas entre a produção recuperada nos dois idiomas: há produção expressiva em inglês e produção reduzida em português. Os resultados sugerem a necessidade de aprofundar a investigação sobre a abertura de dados institucionais das bibliotecas.

Palavras-chave: dados abertos; bibliotecas; bibliometria; produção científica.



Abstract: this study presents a bibliometric analysis of scientific production on open data in libraries from 2015 to 2025. The objective is to map and analyze the volume and temporal distribution of scientific production on the topic, identifying gaps and trends to inform future research stages. Data were collected from national and international databases using searches in Portuguese and English, with temporal refinement. Results reveal expressive differences between the production recovered in both languages: there is substantial production in English and reduced production in Portuguese. The findings suggest the need to deepen the investigation on the opening of libraries' institutional data.

Keywords: open data; libraries; bibliometrics; scientific production.

1 INTRODUÇÃO

A gestão de dados abertos pressupõe sua disponibilização em formatos acessíveis, interoperáveis e reutilizáveis. Nesse contexto, as bibliotecas, historicamente vinculadas ao movimento de acesso aberto, são apontadas como potenciais protagonistas na promoção de práticas de abertura de dados. No entanto, observa-se que, embora haja avanços no campo da gestão de dados de pesquisa, a disponibilização dos dados institucionais das próprias bibliotecas permanece limitada.

Dados como estatísticas de uso, circulação, acervo e informações administrativas são frequentemente produzidos pelas bibliotecas, mas raramente disponibilizados sob os princípios dos dados abertos. Essa lacuna compromete a transparência institucional, dificulta comparações entre bibliotecas e limita o uso desses dados para fins de gestão, pesquisa e tomada de decisão.

Diante desse cenário, este estudo apresenta um

a análise bibliométrica da produção científica sobre dados abertos em bibliotecas, com o objetivo de mapear o volume e a distribuição temporal dessa produção nas principais bases nacionais e internacionais. Busca-se, com isso, identificar lacunas e tendências que subsidiem etapas investigativas futuras, em especial a análise qualitativa da literatura voltada especificamente à abertura de dados institucionais das bibliotecas.



2 REFERENCIAL TEÓRICO

Dados abertos são definidos como aqueles que podem ser livremente utilizados, reutilizados e redistribuídos, desde que observadas condições mínimas como atribuição de autoria. No contexto científico, os princípios FAIR — encontrabilidade, acessibilidade, interoperabilidade e reutilização — consolidaram-se como referência para boas práticas em gestão e compartilhamento de dados (Wilkinson *et al.*, 2016).

No contexto das bibliotecas, a discussão sobre dados abertos tem se concentrado na gestão de dados de pesquisa, destacando o papel dessas instituições como suporte à ciência aberta. Há, contudo, uma segunda dimensão menos explorada: a disponibilização dos próprios dados operacionais das bibliotecas. Sobre esse ponto, Ayre e Craner (2017, p. 181, tradução nossa) defendem que:

Se uma biblioteca coletasse e analisasse seus dados internos e os integrasse com dados disponíveis publicamente, ela poderia melhorar a eficiência dos fluxos de trabalho e fornecer suporte baseado em evidências para o desenvolvimento de programas. O compartilhamento de dados da biblioteca, como tecnologia interna, uso, estatísticas anônimas de circulação e metadados de catálogos, melhora a transparência da organização e pode fornecer aos cidadãos informações sobre o valor da biblioteca.

Os mesmos autores (Ayre; Craner, 2017, p. 176, tradução nossa) argumentam que as bibliotecas já possuem dados úteis prontos para publicação, como estatísticas sobre visitantes, uso de computadores e materiais emprestados; dados orçamentários acessíveis a jornalistas, acadêmicos e cidadãos interessados na destinação de recursos públicos; e informações já publicadas em seus sítios institucionais que, oferecidas em formato legível por máquina, tornam-se reutilizáveis em aplicações locais.

Na mesma direção, Jackson (2015, p. 5, tradução nossa) defende que as bibliotecas têm muito a ganhar ao tornarem seus dados abertos. O autor pondera que as bibliotecas não podem “continuar a citar os benefícios dos dados abertos e reclamar quando os órgãos públicos não são totalmente transparentes com seus dados se elas mesmas se recusam a ser transparentes com os próprios dados”.



3 METODOLOGIA

A pesquisa caracteriza-se como quantitativa, de natureza exploratória e tipo bibliográfico, baseada em análise bibliométrica descritiva da produção científica sobre dados abertos no contexto das bibliotecas.

As buscas foram realizadas nas bases Scopus, Web of Science, LISTA, BRAPCI, Portal de Periódicos da CAPES e BDTD, com o objetivo de contemplar tanto a literatura internacional quanto a produção nacional. Foram utilizados descritores em português e inglês, combinando termos relacionados a “dados abertos” e “bibliotecas” por meio de operadores booleanos: “dados abertos” AND “bibliotecas” para as buscas em português, e “open data” AND “libraries” para as buscas em inglês.

O recorte temporal compreende o período de 2015 a 2025, definido a partir das primeiras publicações identificadas sobre o tema, com vistas a contemplar a produção da última década. Não houve restrição quanto ao tipo de documento, com a finalidade de ampliar a abrangência da recuperação.

Os resultados apresentados correspondem aos dados brutos recuperados nas buscas, com refinamento temporal. Não foram aplicados filtros de exclusão por tipo de documento, remoção de duplicatas entre bases ou exclusão de subcampos temáticos correlatos — como *research data management* ou *linked open data*. Os resultados foram organizados por base de dados, idioma e ano de publicação, permitindo a análise temporal da produção.

4 RESULTADOS

As buscas em português retornaram volume reduzido: 35 registros na BRAPCI, 43 no Portal de Periódicos da CAPES, 9 na LISTA, 3 na Scopus, 2 na BDTD e zero na Web of Science. As buscas em inglês apresentaram volume expressivo: 1.525 registros no Portal de Periódicos da CAPES, 821 na Scopus, 316 na Web of Science e 94 na LISTA, conforme apresentado na Tabela 1.

Tabela 1 – Distribuição dos registros recuperados por base de dados e idioma (2015–2025)

Idioma	Base de dados	Estratégia de busca	Nº de resultados
Português	BDTD	"dados abertos" AND "bibliotecas"	2
Português	BRAPCI	"dados abertos" AND "bibliotecas"	35
Português	CAPES	"dados abertos" AND "bibliotecas"	43
Português	LISTA	"dados abertos" AND "bibliotecas"	9
Português	Scopus	"dados abertos" AND "bibliotecas"	3
Português	Web of Science	"dados abertos" AND "bibliotecas"	0
Inglês	CAPES	"open data" AND "libraries"	1525
Inglês	LISTA	"open data" AND "libraries"	94
Inglês	Scopus	"open data" AND "libraries"	821
Inglês	Web of Science	"open data" AND "libraries"	316

Fonte: Elaborado pelos autores (2026).

A distribuição temporal desses registros, considerando as bases *Scopus*, *Web of Science* e *LISTA* para a produção em inglês, e *Scopus*, *BRAPCI* e *LISTA* para a produção em português, é apresentada no Gráfico 1.

Gráfico 1 – Distribuição anual dos registros recuperados por idioma (2015–2025)



Fonte: Elaborado pelos autores (2026).

A distribuição temporal revela comportamentos distintos entre os dois idiomas. Em inglês, a produção mantém presença contínua ao longo do período, com variação entre 86 registros em 2015 e 132 em 2020, sem tendência de crescimento expressivo. Em português, observa-se crescimento gradual a partir de 2019 — de 1 registro em 2015 para 7 em 2021, mantendo-se entre 5 e 6 registros anuais até 2025.

O contraste entre os volumes nos dois idiomas sugere que o tema possui maior presença na literatura internacional do que na produção científica brasileira em Ciência da Informação, embora em nenhum dos casos se observe expansão expressiva ao longo da última década.

Ressalta-se, contudo, que os dados apresentados correspondem a registros brutos recuperados pelos descritores utilizados, sem triagem temática. É esperado que parte

da produção em inglês esteja associada a abordagens correlatas — como *linked open data* e *web semântica* —, que compartilham o vocabulário de busca, mas não tratam especificamente da abertura de dados institucionais e operacionais das bibliotecas.



5 CONSIDERAÇÕES FINAIS E DESDOBRAMENTOS DA PESQUISA

Os resultados indicam que, apesar da crescente relevância dos dados abertos em diferentes campos da ciência aberta, a produção científica sobre dados abertos e bibliotecas não apresenta tendência de crescimento expressivo ao longo da última década. Na literatura em português, o volume permanece reduzido, enquanto a produção internacional, embora mais numerosa, mantém comportamento relativamente estável no período analisado. Esse cenário reforça a necessidade de aprofundar a investigação sobre como a abertura de dados institucionais das bibliotecas tem sido abordada na literatura científica.

A análise suscita uma questão relevante para pesquisas futuras: por que a discussão sobre dados abertos nas bibliotecas permanece limitada? Uma hipótese possível é que práticas relacionadas à divulgação de dados estejam ocorrendo de forma dispersa e pouco reconhecida como iniciativas de dados abertos. Outra possibilidade, apontada pela própria literatura, é que as bibliotecas reconheçam o valor dos dados abertos sem necessariamente aplicar esse princípio aos seus próprios dados institucionais — o que configuraria uma contradição entre discurso e prática que merece investigação sistemática. (Ayre; Craner, 2017; Jackson, 2015).

Como desdobramentos, propõem-se duas frentes complementares. A primeira consiste em revisão de escopo formal conduzida segundo o protocolo PRISMA-ScR (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses — Scoping Reviews*), com análise qualitativa da literatura recuperada, categorizando-a por foco temático, tipo de biblioteca e abordagem metodológica, com remoção de duplicatas e exclusão de estudos centrados em *research data management* e *linked open data* como objeto principal.

A segunda consiste em investigação empírica das práticas de divulgação de dados pelas bibliotecas, mediante análise de relatórios, boletins e portais institucionais, com vistas à proposição de modelos e diretrizes para a abertura, padronização e descrição de conjuntos de dados que permitam comparações interinstitucionais, facilitem a tomada de decisão baseada em evidências e contribuam para o uso e reuso desses dados pela sociedade.



REFERÊNCIAS

- AYRE, Lori Bowen; CRANER, Jim. Open data: what it is and why you should care. **Public Library Quarterly**, v. 36, n. 2, p. 173-184, 2017. Disponível em: <https://doi.org/10.1080/01616846.2017.1313045>. Acesso em: 29 abr. 2026.
- JACKSON, Brian. The State of Canadian Library Data. **Partnership**: Canadian Journal of Library and Information Practice and Research, v. 10, n. 1, 2015. Disponível em: <https://journal.lib.uoguelph.ca/index.php/perj/article/view/3505>. Acesso em: 29 abr. 2026.
- WILKINSON, Mark D. et al. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific data**, v. 3, n. 1, p. 1-9, 2016. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso em: 1 maio 2026.