



RASTREANDO A PRODUÇÃO CIENTÍFICA DO IBICT:

uma investigação bibliométrica em múltiplas fontes

Rayanne Thaynara Souza e Silva

 <https://orcid.org/0000-0002-2483-2317>.

 rayannesilva@ibict.br.

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

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

Giovana Cavalcanti de Mesquita

 <https://orcid.org/0009-0001-7478-2127>.

 giovanamesquita@ibict.br.

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

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

Maria Vitória Alves Barbosa

 <https://orcid.org/0009-0003-1783-9558>.

 mariabarbosa@ibict.br.

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

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

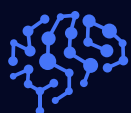
Eixo temático: Bases e Fontes de Dados

Modalidade: Resumo expandido

DOI: 10.22477/ix.ebbc.375

Resumo: Este estudo investiga a representatividade da produção científica do Ibict de 2013 a 2022 em diferentes bases de dados, como *Web of Science*, *Scopus* e *Open Alex*, analisando as características e diferenças entre elas. Embora a produção do Ibict seja mais notável em plataformas abertas, como a *Open Alex*, foram encontradas inconsistências e problemas de desambiguação nas métricas. Por fim, o estudo propõe melhorias na qualidade da desambiguação para aumentar a visibilidade e impacto das publicações do Instituto, além de sugerir estratégias para análises futuras.

Palavras-Chave: Bibliometria. Representação científica. Bases de dados científicas.



1 INTRODUÇÃO

Ao longo dos anos, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) tem se destacado pela criação de redes de colaboração com outras instituições, pela prospecção e transferência de tecnologias de código aberto para o Brasil, como *OJS*, *DSpace* e *Koha*, além de sua atuação na criação, manutenção e disseminação de serviços fundamentais como a BDTD, COMUT, CCN, Bibliodata, entre muitos outros. Essas iniciativas são algumas das que consolidam sua posição como referência nacional e internacional. Não obstante, é relevante ressaltar que o Instituto abriga um corpo destacado de pesquisadores, professores e profissionais técnicos, cuja produção científica é notável. Diante desse contexto, o presente trabalho busca avaliar a produção científica do Instituto e seu impacto por meio de uma análise bibliométrica.

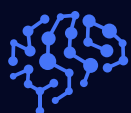
Originalmente, a bibliometria concentrava-se na análise métrica de livros. Com o tempo, ela se expandiu para incluir o estudo de outros tipos de produção bibliográfica, como artigos de periódicos e diversos documentos, e posteriormente passou a examinar a produtividade dos autores e a análise de citações (Araújo, 2006, p. 13). Assim, a bibliometria é uma disciplina da Ciência da Informação que emprega métodos estatísticos para examinar a produção científica e quantificar diversos aspectos da atividade acadêmica (Medeiros; Vitoriano, 2015, p. 492).

No contexto acadêmico, os pesquisadores frequentemente empregam análise bibliométrica por diversas razões. Isso inclui a identificação de tendências emergentes no desempenho de artigos e periódicos, padrões de colaboração e elementos essenciais da pesquisa, além de demonstrar a estrutura intelectual de um domínio específico na literatura existente. Na maioria dos casos, portanto, os dados utilizados costumam ser de grande volume quantitativo e de natureza objetiva. Entretanto, apesar das interpretações desses dados frequentemente refletirem avaliações objetivas, também podem ser influenciadas por interpretações subjetivas, que são estabelecidas por meio de técnicas e procedimentos (Donthu *et al.*, 2021).

Durante a presente pesquisa, foram identificados estudos relevantes para o Ibict, tais como “Estudos métricos da informação em ciência e tecnologia no Brasil realizados sobre a unidade de análise artigos de periódicos” de 2013 (Mueller, 2013), “Presença da bibliometria como processo metodológico em teses indexadas no Ibict (2005-2015)” de 2018 (Villanova; Silva, 2018) e “Mapeamento da produção científica na Biblioteca Digital de Teses e Dissertações do Ibict sobre a Pedagogia da Alternância de 2011 a 2018” de 2019 (Costa; Rodrigues, 2019). Contudo, o diferencial deste trabalho é identificar as publicações científicas realizadas por membros afiliados ao Ibict, visando oferecer uma compreensão ampla da sua representatividade em diferentes bases de dados.

2 CONTEXTO

Para conduzir este estudo, foram efetuados diversos testes em ferramentas que possibilitam a recuperação de dados sobre produção científica. Logo no início da coleta de dados, tornou-se evidente a presença de diferenças significativas nos resultados provenientes de diferentes bases. A seguir, destacam-se os pontos positivos e negativos observados em cada uma delas.



2.1 WEB OF SCIENCE

Os principais aspectos positivos da *Web of Science* (WoS) são a cobertura temporal extensa, iniciando em 1900, a alta precisão e rigor técnico. Entretanto, trata-se de uma base proprietária, de acesso restrito a todas suas funcionalidades, que tem um monitoramento limitado a um pequeno nicho de revistas, considerando apenas citações presentes em outros periódicos também indexados pela WoS. Além disso, apresenta limitações na busca de títulos em idiomas diferentes do inglês, pois estes são sempre traduzidos. Possui baixa cobertura de periódicos nas áreas de tecnologia, engenharias, ciências sociais, humanas, linguística e artes.

2.2 SCOPUS

A base *Scopus* apresenta como alguns dos aspectos positivos: a atualização diária, ampla cobertura em comparação com outras bases proprietárias (especialmente em termos de áreas do conhecimento, tipos de documentos e revistas de acesso aberto), inclusão de artigos “*in press*” que já foram aprovados para publicação mas aguardam a edição específica da revista, alta precisão e rigor técnico, revisão das entradas de publicações por especialistas de cada área e identificação das diversas variações de nomes de autores e instituições.

Porém, devido à sua natureza proprietária, o acesso integral às informações está restrito a assinantes da base. Além disso, não oferece uma visão completa de todas as publicações e citações associadas a uma instituição ou a um autor, uma vez que só inclui resultados de revistas indexadas pela *Scopus*, e tem baixa cobertura nas áreas de humanidades, artes e linguística. Outra questão relevante no meio acadêmico é compreender a imparcialidade da *Scopus* na indexação de periódicos, uma vez que a *Elsevier* é também uma editora científica.

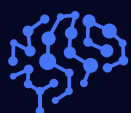
2.3 GOOGLE ACADÊMICO

O *Google Acadêmico* é um agregador de conteúdo de acesso gratuito, com ampla cobertura de fontes de informação, tipos de documentos, idiomas, áreas de conhecimento, níveis de acesso e editoras. No entanto, apresenta algumas limitações, como baixa precisão na recuperação dos resultados, pois pode recuperar e contabilizar várias versões do mesmo documento. Além disso, incorpora documentos que não são científicos, como apontamentos de aula ou sites, e oferece poucas funcionalidades, como a impossibilidade de filtrar documentos ou autores por afiliação¹.

2.4 OPEN ALEX

O *Open Alex* é um catálogo aberto que indexa aproximadamente 250 milhões de recursos de di-

¹ Existe um “atalho” para verificar autores afiliados, através de confirmação de e-mail institucional: https://scholar.google.com/citations?mauthors=ibict.br&hl=en&view_op=search_authors.



versas fontes. As suas principais características positivas incluem a relação entre informações, dados e código abertos aos objetivos de desenvolvimento sustentável estabelecidos pelas Nações Unidas, além de reunir informações de diversas fontes (fechadas e abertas), coletar uma maior diversidade de tipos de documentos e possuir uma boa cobertura de conteúdos em diversos idiomas e regiões geográficas. No entanto, apresenta problemas de acurácia em geral, demonstrando menor qualidade nos metadados, especialmente no que diz respeito ao tipo de documento (Jiao; Li; Fang, 2023) e na desambiguação de autores e instituições.

3 METODOLOGIA

Este estudo adota uma abordagem de pesquisa conhecida como métodos mistos ou quali-quantitativa. Essa metodologia combina abordagens qualitativas e quantitativas, integrando-as em um único estudo e seguindo a linha de pensamento pragmático, conforme definido por Creswell (2010). Com o objetivo de analisar a representatividade da produção científica do Ibict nas principais bases de dados utilizadas para este fim, realizou-se uma análise bibliométrica, método que permite avaliar a tipologia, a quantidade e a qualidade das fontes de informação citadas em pesquisas (Silva; Hayashi; Hayashi, 2011).

Considerando o método empregado, que engloba a análise de citações em dados bibliométricos, abrangendo unidades de publicação e citação (Jiao; Li; Fang, 2023), foi decidido não incluir o *Google Acadêmico*. Isso se deve à sua baixa precisão na entrega de resultados e à inclusão de diversos documentos não científicos. O *Google Acadêmico* tende a recuperar o mesmo documento múltiplas vezes, contando-o como distintos, e não oferece filtros de pesquisa para refinar os resultados, dificultando a recuperação exclusiva das produções científicas do Ibict, mesmo com o uso da ferramenta *Publish or Perish*².

Portanto, os resultados apresentados neste trabalho foram obtidos por meio da utilização de três fontes principais: *Web of Science*, *Scopus* e *Open Alex*. A coleta de dados ocorreu nessas três bases, onde as buscas foram realizadas utilizando a sigla do instituto (Ibict) e o nome completo (Instituto Brasileiro de Informação em Ciência e Tecnologia). Na *Scopus*, a busca foi feita pelo campo "*Organizations*", na *Web of Science* utilizou-se o filtro "*Affiliation*" e na *Open Alex* foi utilizado o filtro "*Institution*". Para facilitar a análise e permitir uma comparação mais precisa, foi estabelecido um recorte temporal de 10 anos (2013-2022) sem filtro por tipo de documento.

Após a coleta dos dados, foi elaborado um quadro para uma análise descritiva dos dados recuperados, contextualizando o impacto das publicações e citações ao longo do período investigado (2013 a 2022). Este processo nos possibilitou identificar quais trabalhos com filiação Ibict tiveram maior relevância e entender sua influência dentro de cada base de dados. Destacamos também as limitações de cada base e seu impacto na precisão dos documentos recuperados, garantindo assim a fidedignidade dos resultados acerca da produção dos pesquisadores vinculados ao Ibict.

² <https://harzing.com/resources/publish-or-perish>.



4 RESULTADOS

A coleta de dados aconteceu durante o período de janeiro/fevereiro de 2024, onde foi elaborado um quadro com informações pertinentes a cada autor (nome, período de permanência no Ibict, setor, área de interesse e o currículo Lattes) para a melhor descrição e visualização dos resultados. Até o presente momento foram recuperadas informações de aproximadamente setenta autores identificados com afiliação Ibict em suas publicações, para além das informações básicas, também foi acrescentado ao quadro os perfis dos autores nas bases estudadas (*WoS*, *Scopus* e *Open Alex*).

Os resultados deste estudo confirmam a hipótese de que a produção científica do Ibict possui maior representatividade em plataformas de acesso aberto, como a *Open Alex* (ver Tabela 1). Isso é evidenciado pelo fato de que uma parcela substancial das publicações vinculadas ao Instituto está em revistas de acesso aberto e/ou em língua portuguesa, que não foram indexadas nas duas plataformas proprietárias examinadas neste estudo. Além disso, a *Open Alex* também recupera outros tipos de produções acadêmico-científicas, não se limitando apenas a artigos de revistas, mas também livros, teses, entre outros. Adicionalmente, apresenta uma cobertura mais ampla nas áreas de humanidades e ciências sociais, incluindo fontes não publicadas em inglês, e destaca-se por uma representação mais abrangente do sul global.

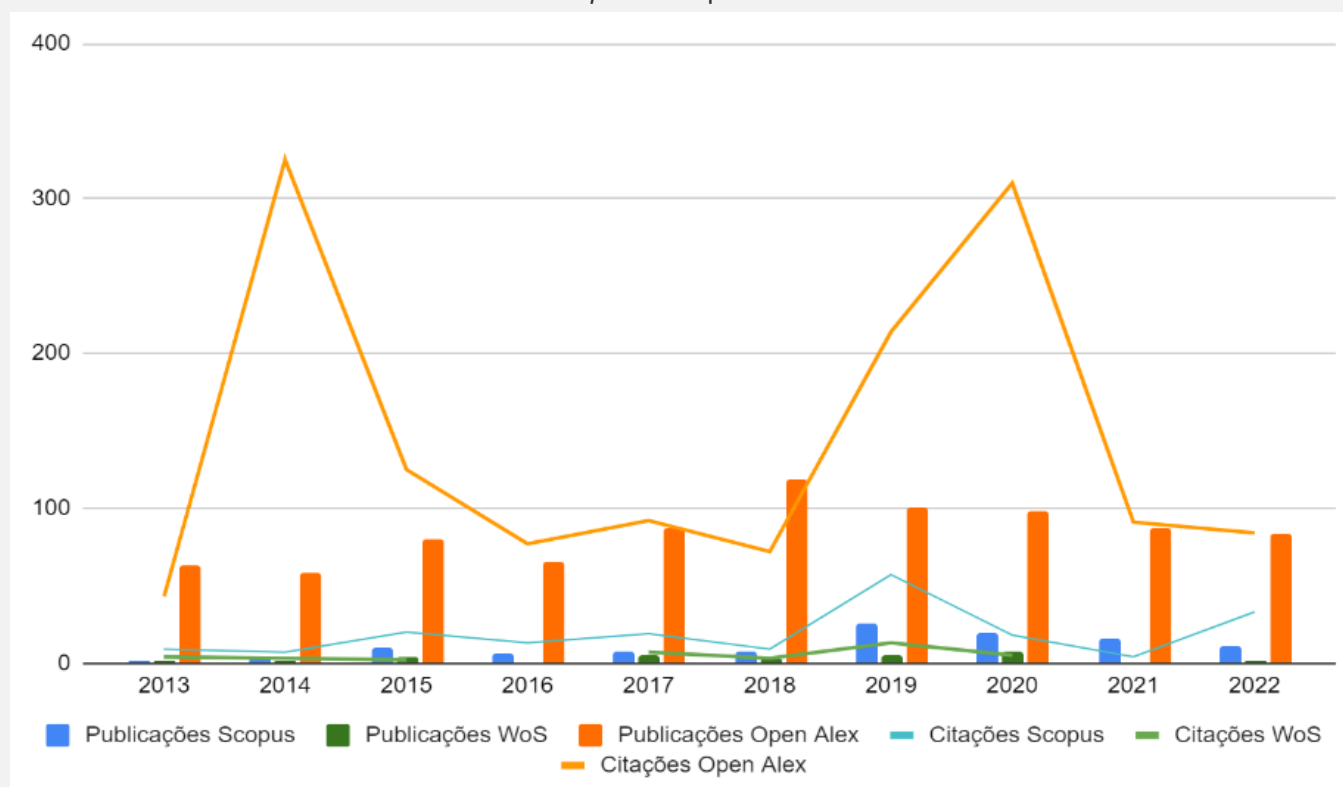
Tabela 1 - Total de publicações e citações com afiliação Ibict nas bases de dados *Web of Science*, *Scopus* e *Open Alex*.

2013-2022	<i>Scopus</i>	<i>WoS</i>	<i>Open Alex</i>
Total de documentos	108	28	841
Total de citações	189	37	1433
Índice H	7	3	16

Fonte: elaborado pelo autor (2024).

Embora, de maneira geral, a quantidade de citações exceda a quantidade de documentos produzidos e/ou indexados, conforme demonstrado no Gráfico 1, foi observado em alguns anos específicos nas três bases que o número de publicações foi maior que o número de citações. Estes anos incluem: 2015 e 2020 na *Web of Science*, 2018 na *Open Alex* e 2021 na *Scopus*.

Gráfico 1 - Total de publicações e citações com afiliação Ibict nas bases de dados *Web of Science*, *Scopus* e *Open Alex* por ano.



Fonte: Elaborado pelo autor (2024).

4.1 LIMITAÇÕES DA PESQUISA

Durante parte da pesquisa deste estudo, foi utilizada a ferramenta *Publish or Perish* com o objetivo de padronizar os dados. No entanto, identificaram-se várias inconsistências, principalmente na discrepância entre os números apresentados e os resultados obtidos por meio de busca direta em cada base de dados. Ao utilizar essa ferramenta pôde-se notar maior fidelidade nos dados da *Open Alex*, por exemplo, e uma discrepância significativa nos dados da *Scopus*, por exemplo.

Outra contingência identificada diz respeito à recuperação de documentos na *Open Alex*. Durante a análise dos artigos mais citados, observou-se um problema de desambiguação que levou à atribuição equivocada de alguns artigos do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict/Fiocruz) como sendo produções do Ibict. Além disso, foi constatado um equívoco na desambiguação de alguns autores, como por exemplo as duas autoras Maristela Martins e Maria Martins, que não possuem afiliação com o Ibict em suas produções, mas entram no quantitativo de recuperação do instituto, o que impacta diretamente na precisão do total de citações atribuídas ao Instituto.

4 CONCLUSÕES

Durante a condução do presente estudo, concluiu-se que, embora a *Open Alex* tenha sido aponta-

da como uma base com maior cobertura e representatividade da produção científica do Ibict, nenhuma delas reflete de maneira totalmente precisa a realidade, por diversos motivos. Isso ocorre tanto pela omissão, quanto pela recuperação incorreta de documentos vinculados ao instituto. No entanto, é possível identificar estratégias para aprimorar a qualidade da desambiguação na plataforma *Open Alex*, tanto em relação aos pesquisadores, por meio do correto preenchimento das informações no *ORCID*, quanto em relação às instituições, ao associarem-se ao *Crossref*, por exemplo. Essas ações podem contribuir para a correção dos metadados das publicações filiadas, uma vez que ambas as bases são coletadas pela *Open Alex*.

Ao preencher a tabela resumo com informações dos autores, identificaram-se problemas, como perfis duplicados, que dificultam a contagem precisa de publicações e citações de um mesmo autor. Essas duplicações podem ocorrer devido à omissão de um dos sobrenomes ou a diferentes tipos de abreviações, resultando em múltiplos perfis para alguns autores nas três bases de dados utilizadas. Recomenda-se, portanto, que os autores tomem medidas proativas, como solicitar a fusão de perfis nessas bases, evitar a criação de mais de um *ID ORCID* e verificarem seus perfis no *Google Acadêmico*.

Para estudos subsequentes, recomenda-se considerar a inclusão da plataforma *Dimensions*, desenvolvida pela *Digital Science*. Embora não tenha sido viável para este estudo, a plataforma oferece uma versão de avaliação do produto. Além disso, sugere-se verificar a disponibilidade da plataforma *Laguna*, atualmente em desenvolvimento pelo Ibict, para comparação na recuperação dos dados. Essas adições podem enriquecer a análise e oferecer uma compreensão mais abrangente da representatividade da produção científica do Ibict. Além disso, propõe-se que se discutam metodologias que possibilitem avaliar outros tipos de documentos produzidos pelo Ibict, que não são indexados por bases especializadas em publicações científicas, como as publicações técnicas (análises e relatórios), bem como materiais de divulgação ou recursos educacionais.

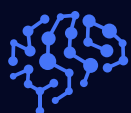
REFERÊNCIAS

ARAÚJO, Carlos A. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 11-32, 2006. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/16>. Acesso em: 24 jan. 2024.

COSTA, Odaleia Alves da; RODRIGUES, Anny Camila Lima. Mapeamento da produção científica na BDTD do IBICT sobre a Pedagogia da Alternância de 2011 a 2018. **Revista Brasileira de Educação do Campo**, Palmas, v. 4, p. e7257, 2019. DOI: <https://doi.org/10.20873/uft.rbec.e7257>. Disponível em: <https://sistemas.uft.edu.br/periodicos/index.php/campo/article/view/7257>. Acesso em: 19 abr. 2024.

CRESWELL, John W. **Projeto de pesquisa: método qualitativo, quantitativo e misto**. 3 ed. Porto Alegre: Penso, 2010.

DONTHU, Noveen. *et al.* How to conduct a bibliometric analysis: An overview and guidelines. **Journal of Business Research**, Athens, v. 133, p. 285-296, Sept. 2021. Disponível em: <https://doi.org/10.1016/j.jbusres.2021.04.070>. Acesso em: 22 fev. 2024.



JIAO, Chenyue; LI, Kai; FANG, Zhichao. How are exclusively data journals indexed in major scholarly databases? An examination of four databases. **Scientific Data**, London, v. 10, n. 737, p. 1-9, Oct. 2023. DOI: <https://doi.org/10.1038/s41597-023-02625-x>. Disponível em: <https://www.nature.com/articles/s41597-023-02625-x>. Acesso em: 2 fev. 2024.

MEDEIROS, José Mauro Gouveia de; VITORIANO, Maria Albeti Vieira. A evolução da bibliometria e sua interdisciplinaridade na produção científica brasileira. **RDBC**: Revista Digital de Biblioteconomia e Ciência da Informação, Campinas, v. 13, n. 3, p. 491-503, set. 2015. Disponível em: <https://doi.org/10.20396/rdbci.v13i3.8635791>. Acesso em: 24 jan. 2024.

MUELLER, Suzana. Estudos métricos da informação em ciência e tecnologia no Brasil realizados sobre a unidade de análise artigos de periódicos. **Liinc em Revista**, Brasília, v. 9, n. 1, 2013. DOI: <https://doi.org/10.18617/liinc.v9i1.558>. Disponível em: <https://revista.ibict.br/liinc/article/view/3429>. Acesso em: 19 abr. 2024.

SILVA, Márcia Regina; HAYASHI, Carlos Roberto Massao; HAYASHI, Maria Cristina Piumbato Innocentini. Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. **InCID**: Revista de Ciência da Informação e Documentação, São Paulo, v. 2, n. 1, jan./jun. 2011. DOI: <https://doi.org/10.11606/issn.2178-2075.v2i1p110-129>. Disponível em: <https://www.revistas.usp.br/incid/article/view/42337>. Acesso em: 6 abr. 2024.

VILLANOVA, Andre Philippe; SILVA, Marcia Regina. Presença da bibliometria como processo metodológico em teses indexadas no IBICT (2005-2015). **Biblionline**, João Pessoa, v. 14, n. 2, p. 11-24, 2018. DOI: <https://doi.org/10.22478/ufpb.1809-4775.2018v14n2.41376>. Disponível em: <https://periodicos.ufpb.br/index.php/biblio/article/view/41376>. Acesso em: 19 abr. 2024.