



MAPEAMENTO DO DOMÍNIO CIÊNCIA ABERTA E SAÚDE NA BASE DE DADOS *DIMENSIONS*

Taliane de Assis Oliveira

 <https://orcid.org/0000-0003-2266-8573>.

✉ taliane.assis@hotmail.com.

🏢 Universidade Federal do Rio Grande do Norte (UFRN) |

ROR: <https://ror.org/04wn09761> | Natal, Brasil.

Nancy Sánchez-Tarragó

 <https://orcid.org/0000-0002-5114-6072>.

✉ nancy.sanchez@ufrn.br.

🏢 Universidade Federal do Rio Grande do Norte (UFRN) |

ROR: <https://ror.org/04wn09761> | Natal, Brasil.

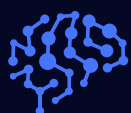
Eixo temático: Mapas e redes de ciência

Modalidade: Resumo expandido

DOI: 10.22477/ix.ebbc.344

Resumo: Objetiva realizar mapeamento bibliométrico do domínio ciência aberta e saúde, a partir da análise da produção científica na base de dados *Dimensions*, utilizando indicadores de produtividade de autores, instituições e periódicos, distribuição temporal, citação de artigos, redes de cocitação de documentos e redes de coocorrência de palavras. Como resultados se constata o crescimento dessa produção ao longo do tempo, alavancada também pela pandemia de Covid-19. Os documentos e temas mais representativos do domínio estão alinhados com as abordagens atuais na ciência aberta: uso de diretrizes para favorecer o relato transparente de métodos de pesquisa, de plataformas que fomentam a abertura de todo o ciclo de pesquisa, abertura de dados e implantação de comunidades colaborativas.

Palavras-Chave: Ciência aberta. Saúde. Análise de domínio. Bibliometria. Mapas de ciência.



1 INTRODUÇÃO

Conforme Delfanti e Pitrelli (2015), a ciência aberta é um conceito amplo que engloba diversas práticas e ferramentas relacionadas ao uso de tecnologias digitais colaborativas e alternativas de propriedade intelectual. Essa abordagem mobiliza diferentes níveis e escopos de abertura, promovendo maior dinamismo nas atividades de ciência, tecnologia e inovação, bem como maior diversidade de perspectivas no campo científico, incluindo diversidade de línguas, documentos e atores envolvidos (Unesco, 2022). O conceito abrange uma variedade de movimentos, iniciativas e práticas heterogêneas, tais como acesso aberto à literatura científica, dados abertos, ciência cidadã, avaliação aberta e software livre.

As ciências da saúde têm desempenhado um papel pioneiro em iniciativas voltadas para a abertura do conhecimento científico desde os primórdios do Movimento pelo Acesso Aberto à Informação. Isso se reflete na criação de periódicos de acesso aberto, repositórios especializados, políticas de depósito e financiamento da pesquisa, entre outras ações. A crise sanitária desencadeada pela pandemia de Covid-19 reiterou a importância e a urgência da colaboração internacional e do compartilhamento aberto e rápido de dados e informações, destacando ainda mais a relevância da ciência aberta. Segundo Besançon *et al.* (2021), mais de 80.000 preprints e artigos revisados por pares relacionados à Covid-19 e ao SARS-CoV-2 foram publicados desde o surgimento deste vírus em dezembro de 2019, grande parte disponibilizados gratuitamente (Gianola *et al.*, 2020). Na área da saúde, portanto, a adoção da ciência aberta traz potencialmente uma série de benefícios, incluindo a facilitação de colaborações e parcerias científicas, o aprimoramento das capacidades analíticas e investigativas, a detecção precoce de ameaças sanitárias e ambientais, o acompanhamento em tempo real da resposta a essas ameaças, a embasamento de decisões políticas, o aumento da participação pública, da transparência e da responsabilização (Strydom *et al.*, 2022).

Atualmente, os esforços em prol da ciência aberta não se limitam apenas aos produtos da atividade científica, mas também envolvem transformações em todo o ciclo da pesquisa. De acordo com Thibault *et al.* (2023), quatro grandes abordagens estão em desenvolvimento no contexto da ciência aberta: a) promoção da transparência nos métodos de pesquisa, por meio da adoção de diretrizes padronizadas de relatório, como as diretrizes *Consolidated Standards of Reporting Trials* (CONSORT) e *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA); b) incentivo ao compartilhamento de todos os componentes da pesquisa, além do artigo de periódico, incluindo o registro de protocolos de estudo, dados, códigos e preprints. Exemplos disso são o *ClinicalTrials.gov* e o *Open Science Framework*; c) promoção do trabalho em equipe, mediante o reconhecimento e a valorização de diversas especialidades e contribuições ao ciclo da pesquisa. Exemplo disso é a *Contributor Roles Taxonomy* (CRediT); e d) fomento de uma cultura de pesquisa que valorize a abertura e transforme as práticas de maneira efetiva, incorporando a ciência aberta nos currículos e fundindo iniciativas bottom-up e top-down.

Tais caminhos e iniciativas têm desenvolvimentos desiguais e dinâmicos em diferentes áreas disciplinares. Nesse sentido, esta pesquisa objetiva caracterizar a produção científica no domínio da ciência aberta e saúde, de maneira a compreender sua estrutura intelectual e social e como estão se refletindo as principais tendências e desenvolvimentos da ciência aberta na produção científica da área das ciências da saúde.

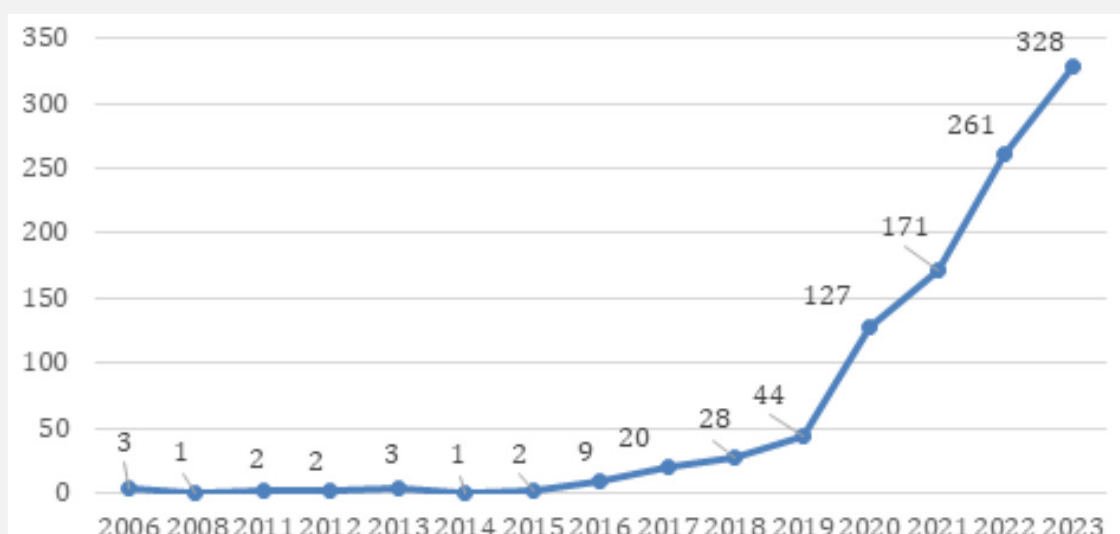
2 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa se baseia na análise de domínio proposta por Birger Hjørland e Hanner Albreschtsen (1995). O domínio vai além das áreas de conhecimento interrelacionadas e inclui a organização social refletida nas conexões entre autores, instituições, periódicos, documentos e vocabulários. As análises bibliométricas e os mapas de ciência ajudam a compreender esses domínios. Para isso, nesta pesquisa realizou-se uma busca na base de dados Dimensions em janeiro de 2024, com a seguinte estratégia: ("ciência aberta" AND saúde) OR ("open science" AND health) OR ("ciencia abierta" AND salud)), no campo *Title and abstract*, utilizando como filtro a tipologia de documento *Artigo* e o recorte temporal até 2023. Os registros recuperados foram exportados em formato *Comma-Separated Values* (CSV) para o *software* VOSviewer (versão 1.6.20). Foram analisados os indicadores de produtividade de autores, instituições e periódicos, a distribuição temporal e os artigos com maior número de citações. Também foram elaboradas redes de cocitação de documentos e coocorrência de palavras. A partir desses mapas, foram analisados os documentos e temas que constituem referências intelectuais para a pesquisa no domínio, bem como os temas que refletem as tendências contemporâneas da ciência aberta na saúde.

2 RESULTADOS

Foram recuperados 1.014 artigos, dos quais 97% estavam em língua inglesa. A análise diacrônica mostra um crescimento consistente na produção científica sobre ciência aberta na área da saúde (Figura 1). Os primeiros artigos registrados na Dimensions remontam a 2006, com um aumento significativo e constante a partir de 2019. Esse aumento nas publicações está relacionado ao período da pandemia de Covid-19, que demandou a rápida disponibilização e acesso a estudos científicos, destacando a urgência da abertura de periódicos comerciais por assinatura, bem como a implementação de repositórios de preprints, iniciativas de avaliação aberta, infraestruturas interoperáveis, políticas de gestão de dados abertos, entre outros.

Figura 1 - Distribuição de documentos no domínio da ciência aberta e saúde na Dimensions



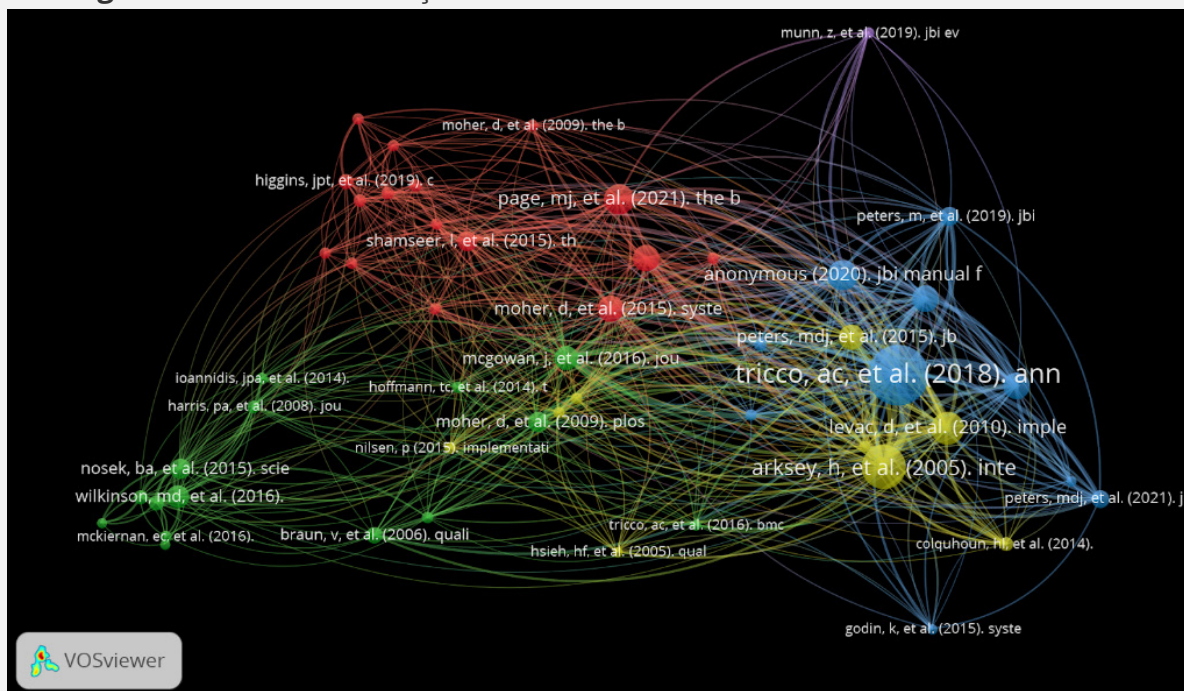
Fonte: Dados da pesquisa (2024).

Dos 5.380 autores identificados, apenas sete têm mais de 5 publicações no domínio, com destaque para David Moher, epidemiologista irlandês e cientista sênior do Programa de Epidemiologia Clínica do Instituto de Pesquisa do Hospital de Ottawa, Canadá, com 8 artigos. Entre as 1.828 instituições, 17 se destacam com mais de 15 artigos produzidos, sendo 10 da América do Norte (Canadá e Estados Unidos), 4 da Europa (com destaque para o Reino Unido), 2 da Oceania (com destaque para a Austrália) e apenas uma da América do Sul (Brasil-Universidade de São Paulo).

Entre os periódicos científicos com mais de dez publicações no domínio da ciência aberta e saúde, destacam-se os seis mais prolíficos: *BMJ Open* (134 artigos), *JBI Evidence Synthesis* (48), *Systematic Reviews* (47), *PLoS ONE* (28), *JMIR Research Protocols* (17) e *Medicine* (14). Todos os seis periódicos são de acesso aberto, na modalidade Ouro, pois cobram taxas de processamento (*Article Processing Charge* - APC).

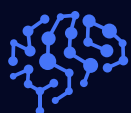
A Figura 2 apresenta a rede de cocitação de documentos citados, conformada mediante o cálculo da frequência com que pares de documentos são citados juntos nas referências dos artigos estudados. A análise de cocitação é fundamental para identificar as bases intelectuais de um domínio e seus principais referentes. Para criar o mapa, foram selecionados os 51 documentos com pelo menos 10 citações, a partir de um total de 34.335 documentos citados.

Figura 2 – Rede de cocitação de documentos. Domínio ciência aberta e saúde



Fonte: Dados da pesquisa (2024).

O mapa revela quatro *clusters* coloridos. A análise dos títulos e resumos dos trabalhos permitiu identificar dois temas principais: os *clusters* amarelo, azul e vermelho são compostos, fundamentalmente, por trabalhos que discutem diretrizes de revisão sistemática, com destaque para a metodologia *Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols* (PRISMA-P), amplamente utilizada nas ciências da saúde para garantir padronização e transparência nos métodos de pesquisa. Por outro lado, o cluster verde agrupa trabalhos que abordam a importância das práticas de ciência aberta para



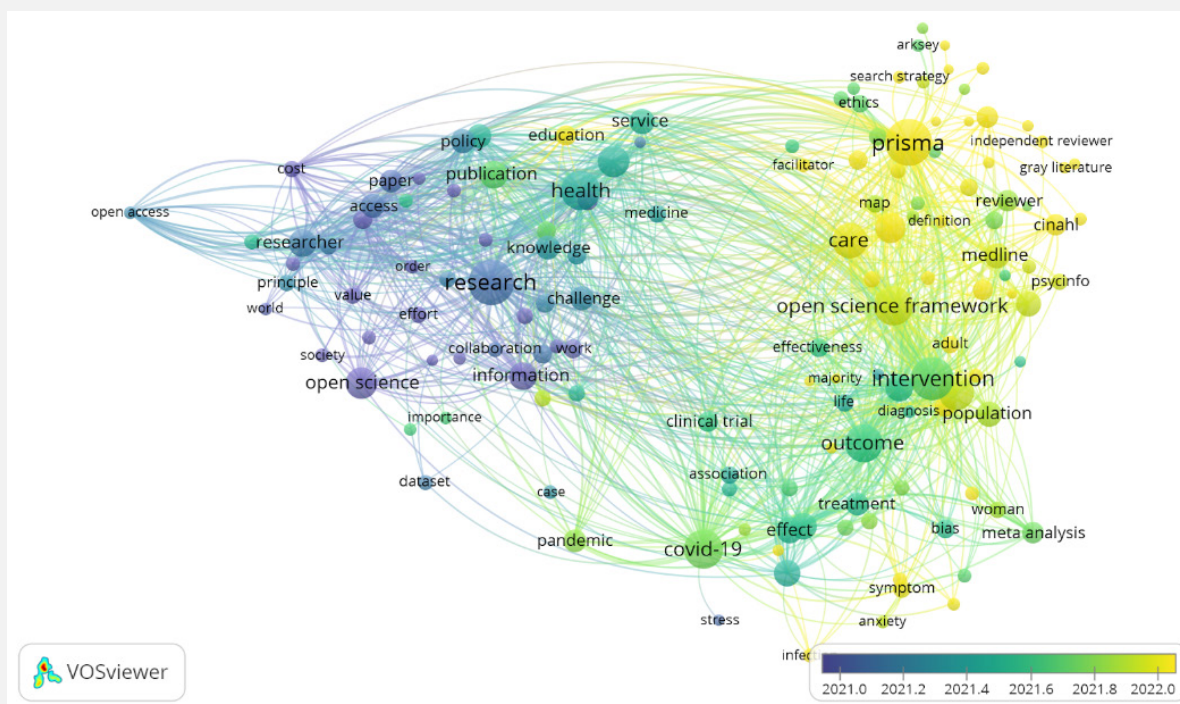
a pesquisa, assim como discussões relacionadas com as infraestruturas de compartilhamento e reuso de dados científicos, como a implementação dos princípios *Findable, Accessible, Interoperable, Reusable* (FAIR) para os dados científicos. Esses trabalhos, publicados entre 2008 e 2021, constituem referências importantes no domínio da ciência aberta e saúde.

Para entender melhor os principais temas discutidos no domínio, realizou-se uma análise de coocorrência de palavras, contabilizando a frequência com que pares de palavras extraídas do campo Title and abstract são usadas juntas. Alta coocorrência permite inferir alta relevância dos temas representados pelos termos. Foi criado um tesauro para padronizar alguns termos. Após a aplicação do tesauro, de um total de 29.843 palavras, foram selecionadas aquelas com pelo menos 40 ocorrências, totalizando 248 palavras; destas, apenas 149 atenderam ao limiar de 60% de relevância do VOSviewer. Uma filtragem manual adicional descartou 7 palavras de uso comum. Utilizou-se a visualização *Overlay* para representar o uso das palavras ao longo do tempo, conforme a data de publicação dos artigos (Figura 3).

Em consonância com a análise de cocitação de documentos que destacou os principais trabalhos e temas que fundamentam as pesquisas no domínio, o mapa de coocorrência de palavras revela três clusters com os temas mais relevantes pesquisados e publicados atualmente: o cluster violeta agrupa temas mais “antigos”, representados pelos termos *research, open science, open access, information, researcher, collaboration*, sugerindo discussões conceituais, políticas e técnicas sobre a implementação de iniciativas e infraestruturas de ciência aberta na pesquisa em saúde. No cluster verde, encontramos termos mais associados à pesquisa e prática em saúde, como *health, intervention, outcome, treatment, risk* e *Covid-19*. Já o cluster amarelo, aparentemente com publicações mais recentes, reúne os dois termos que aparecem com maior frequência no corpus: *PRISMA* e *Open Science Framework*.

Open Science Framework trata-se de uma plataforma on-line que promove fluxos de trabalho abertos abarcando todo o ciclo de pesquisa, desde o desenvolvimento de uma ideia de pesquisa, a concepção de um estudo, o armazenamento e a análise de dados coletados e a redação e publicação de relatórios ou artigos (Foster; Deardorff, 2017). Esta plataforma é um dos espaços recomendados pelas diretrizes PRISMA para depositar arquivos suplementares de revisões sistemáticas de forma aberta e transparente (Page *et al.*, 2020). De fato, desde 2016 o Comitê Internacional de Editores de Periódicos Médicos (ICMJE) também anunciou a exigência do compartilhamento aberto de dados de ensaios clínicos publicados. Isto tem implicações para a concepção, condução e relato de revisões sistemáticas e meta-análises, melhorando as possibilidades para as sínteses de evidências, como relata o artigo de Ross (2016), um dos mais influentes no corpus analisado.

Figura 3 - Rede de coocorrência de palavras no domínio da ciência aberta e saúde



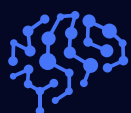
Fonte: Dados da pesquisa (2024).

Além disso, é possível observar na cor verde-amarela o termo Covid-19 articulando-se com os três clusters. A pandemia evidenciou a importância da ciência aberta na elaboração e comunicação de revisões sistemáticas, fundamentais para o desenvolvimento de pesquisas sobre tratamentos, vacinas e outras intervenções de saúde. A Covid-19 também impulsionou o desenvolvimento de softwares, aplicações e infraestruturas de ciência aberta, conforme demonstrado no Quadro 1.

Neste Quadro 1, estão listados os artigos com mais de 130 citações no domínio. Seus temas ilustram as vinculações temáticas esboçadas nos mapas anteriores, com ênfase nos artigos que abordam a criação de infraestruturas abertas e interoperáveis para compartilhamento de artigos e conjuntos de dados.

Quadro 1 - Artigos com mais de 130 citações no domínio ciência aberta e saúde

Artigo	Citações
Haendel, M. A. <i>et al.</i> (2021). <i>The National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment</i> . J Am Med Inform Assoc. v. 28, n. 3, p. 427-443. https://doi.org/10.1093/jamia/ocaa196	351
Bullock, J. <i>et al.</i> (2020). <i>Mapping the landscape of Artificial Intelligence applications against COVID-19</i> . Journal of Artificial Intelligence Research. v. 69. https://doi.org/10.1613/jair.1.12162	329
Honorato, R.V. <i>et al.</i> (2021). <i>Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem</i> . Frontiers in Molecular Biosciences. v. 28, n. 8, p. 729513. https://doi.org/10.3389/fmolb.2021.729513	297
Woelfle <i>et al.</i> (2011). <i>Open science is a research accelerator</i> . Nature Chemistry. n. 3, p. 745–748. https://doi.org/10.1038/nchem.1149	186



Artigo	Citações
Agley <i>et al.</i> (2021). <i>Misinformation about COVID-19: evidence for differential latent profiles and a strong association with trust in science</i> . BMC Public Health. v. 21, n. 89. https://doi.org/10.1186/s12889-020-10103-x	160
Noroni, N (2021). <i>Addressing bias in big data and AI for health care: A call for open science</i> . Patterns (NY). v. 8; n. 2, p. 100347. https://doi.org/10.1016/j.patter.2021.100347	131

Fonte: Dados da pesquisa (2024).

Por exemplo, o artigo de Haendel *et al.* (2021) relatam a criação da *National COVID Cohort Collaborative* (N3C), uma comunidade científica aberta que construiu rapidamente uma infraestrutura escalável incorporando dados clínicos multicêntricos para análise da Covid-19, possibilitando a colaboração entre médicos, pesquisadores e cientistas de dados. Honorato *et al.* (2021) abordam a criação do projeto WeNMR, que oferece serviços para o portal *European Open Science Cloud*. Por outro lado, Noroni *et al.* (2021) destacam a importância de dados públicos abertos, interoperáveis e diversos para treinar inteligências artificiais e evitar vieses e discriminações em diagnósticos e cuidados de saúde.

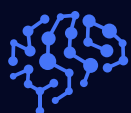
4 CONCLUSÕES

O mapeamento do domínio da ciência aberta e saúde, a partir da análise dos artigos registrados na base de dados Dimensions, revelou um crescimento significativo dessa produção ao longo do tempo, impulsionado também pela pandemia de Covid-19. Em termos de afiliação dos autores, instituições da América do Norte e Europa predominam, destacando-se o Brasil, em 17º lugar em produtividade, com a Universidade de São Paulo como principal contribuinte. Foi observada uma conexão importante entre o uso de metodologias de revisões sistemáticas, fundamentais nas ciências da saúde, e as exigências de publicação aberta de ensaios clínicos e arquivos suplementares das revisões sistemáticas. Entre os artigos mais relevantes, destacam-se relatos de criação de infraestruturas abertas para compartilhamento de dados e artigos. Duas tendências contemporâneas da ciência aberta citadas na literatura foram identificadas como relevantes neste *corpus*: a promoção da transparência nos métodos de pesquisa por meio da adoção de diretrizes padronizadas de relatório, como as diretrizes PRISMA e o incentivo ao compartilhamento e à abertura de todo o ciclo de pesquisa por meio do *Open Science Framework*.

REFERÊNCIAS

BESANÇON, Lonni *et al.* Open science saves lives: lessons from the COVID-19 pandemic. **BMC Medical Research Methodology**, [S. l.], v. 21, article number 117, p. 1-18, 2021. DOI: <https://doi.org/10.1186/s12874-021-01304-y>. Disponível em: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-021-01304-y>. Acesso em: 8 abr. 2024.

DELFANTI, Alessandro; PITRELLI, Nico. Ciência aberta: revolução ou continuidade. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (org.). **Ciência aberta, questões abertas**. Brasília:



Ibict; Rio de Janeiro: UNIRIO, 2015. Cap. 3, p. 59-69. Disponível em: <http://livroaberto.ibict.br/handle/1/1060>. Acesso em: 14 jan. 2024.

FOSTER, Erin D.; DEARDORFF, Ariel. Open Science Framework (OSF). **Journal of Medical Library Association**, Chicago, v. 105, n. 2, p. 203-206, Apr. 2017. DOI: <https://doi.org/10.5195/jmla.2017.88>. Disponível em: <https://jmla.pitt.edu/ojs/jmla/article/view/88>. Acesso em: 14 abr. 2024.

GIANOLA, Silvia *et al.* Characteristics of academic publications, preprints, and registered clinical trials on the COVID-19 pandemic. **PLoS ONE**, [S. l.], v. 15, n. 10, e0240123, 2020. DOI: <https://doi.org/10.1371/journal.pone.0240123>. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0240123>. Acesso em: 14 jan. 2024.

HAENDEL, Melissa A. *et al.* The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. **Journal of the American Medical Informatics Association**, [S. l.], v. 28, n. 3, p. 427-443, 2021. DOI: <https://doi.org/10.1093/jamia/ocaa196>. Disponível em: <https://academic.oup.com/jamia/article/28/3/427/5893482>. Acesso em: 2 maio 2024.

HJØRLAND, Birger, ALBRECHTSEN, Hanne. Toward a new horizon in information science: Domain-analysis. **Journal of the American Society for Information Science**, [S. l.], v. 46, n. 6, p. 400-425, 1995. DOI: [https://doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-ASIJ2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASIJ2>3.0.CO;2-Y). Disponível em: [https://asistdl.onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-4571\(199507\)46:6%3C400::AID-ASIJ2%3E3.0.CO;2-Y](https://asistdl.onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4571(199507)46:6%3C400::AID-ASIJ2%3E3.0.CO;2-Y). Acesso em: 14 abr. 2024.

HONORATO, Rodrigo V. *et al.* Structural biology in the clouds: the WeNMR-EOSC Ecosystem. **Frontiers in Molecular Biosciences**, [S. l.], v. 8, article 729513, July 2021. DOI: <https://doi.org/10.3389/fmolb.2021.729513>. Disponível em: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.729513/full>. Acesso em: 2 maio 2024.

NORORI, Natalia *et al.* Addressing bias in big data and AI for health care: a call for open science. **Patterns**, [S. l.], v. 2, n. 10, article 100347, 8 Oct. 2021. DOI: <https://doi.org/10.1016/j.patter.2021.100347>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666389921002026?via%3Dihub>. Acesso em: 2 maio 2024.

PAGE, Matthew J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, [S. l.], v. 372, n. 71, 2021. Published 29 March 2021. Research Methods & Reporting. DOI: <https://doi.org/10.1136/bmj.n71>. Disponível em: <https://www.bmj.com/content/372/bmj.n71>. Acesso em: 2 maio 2024.

ROSS, Joseph S. Clinical research data sharing: what an open science world means for researchers involved in evidence synthesis. **Systematic Reviews**, [S. l.], v. 5, article number 159, 2016. DOI: <https://doi.org/10.1186/s13643-016-0334-1>. Disponível em: <https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-016-0334-1>. Acesso em: 2 maio 2024.

STRYDOM, Adèle *et al.* Open access and its potential impact on public health: a South African perspective. **Frontiers in Research Metrics and Analytics**, [S. l.], v. 7, article 975109, 2022. DOI: <https://doi.org/10.3389/frma.2022.975109>. Disponível em: <https://www.frontiersin.org/articles/10.3389/frma.2022.975109/full>. Acesso em: 8 abr. 2024.



THIBAUT, Robert T. *et al.* Open Science 2.0: towards a truly collaborative research ecosystem.

PLoS Biology, [S. l.], v. 21, n. 10, e3002362, Oct. 2023. Version 2. DOI: <https://doi.org/10.1371/journal.pbio.3002362>. Disponível em: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3002362>. Acesso em: 14 abr. 2024.

UNESCO. **Recomendação da UNESCO sobre Ciência Aberta**. [S. l.]: Unesco, c2022. 34 p. DOI: <https://doi.org/10.54677/XFFX3334>. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por. Acesso em: 14 de abr. 2024.