



CAPÍTULO 4

ALÉM DO ACESSO ABERTO: DESAFIOS E PERSPECTIVAS NA ERA DOS DADOS CIENTÍFICOS ABERTOS

Fabiano Couto Corrêa da Silva¹
Laura Vilela Rodrigues Rezende²
Elaine Rosangela de Oliveira Lucas³



¹Universidade Federal do Rio Grande do Sul (UFRGS). ORCID: <https://orcid.org/0000-0001-5014-8853>, E-mail: fabianocc@gmail.com

²Universidade Federal de Goiás (UNIR). ORCID: <http://orcid.org/0000-0002-8891-3263>. E-mail: laura_rezende@ufg.br

³Universidade do Estado de Santa Catarina (UFSC). ORCID: <https://orcid.org/0000-0002-2796-3566>. E-mail: lani@udesc.br

4.1 INTRODUÇÃO

Em menos de uma década, a ciência experimentou transformações significativas no seu cenário institucional e regulatório. A adoção do *General Data Protection Regulation* (GDPR) (European Union, 2016), implementado em 2018, promoveu mudanças substanciais na coleta e gestão de dados científicos. Esse regulamento ampliou o controle dos indivíduos sobre seus dados pessoais, exigindo transparência no uso e armazenamento, bem como a implementação de medidas de segurança robustas para sua proteção. Além disso, o GDPR afetou a transferência internacional de dados, exigindo conformidade com regras específicas para garantir a proteção de dados fora da União Europeia (UE), e estabeleceu responsabilidades às instituições de pesquisa, incluindo a realização de Avaliações de Impacto na Proteção de Dados e a designação de Encarregados de Proteção de Dados.

A entrada em vigor da Lei Geral de Proteção de Dados (LGPD) no Brasil em 2020 (Brasil, [2020], art. 65, inc. II), representa avanços significativos na regulamentação da privacidade e proteção de dados. Ambas as legislações compartilham objetivos comuns, como fortalecer a proteção dos dados pessoais e aumentar o controle dos indivíduos sobre suas informações. Enquanto o GDPR serviu de modelo para muitos aspectos da LGPD, existem particularidades na legislação brasileira, como a autoridade nacional responsável pela aplicação da lei e nuances nas definições de dados sensíveis. Ambas as leis impõem obrigações semelhantes às organizações, como a necessidade de consentimento para coleta e uso de dados, proteção e anonimização de dados sensíveis, e a nomeação de um responsável pela proteção de dados. A principal diferença reside nos contextos regionais específicos: enquanto o GDPR tem um impacto abrangente nas transferências de dados internacionais para países da UE, a LGPD molda o cenário de proteção de dados no Brasil, influenciando também as relações comerciais e de dados com outros países, incluindo aqueles da UE.

Paralelamente, a emergência de iniciativas como a Declaração de São Francisco sobre Avaliação de Pesquisa (DORA, 2012)¹, os princípios FAIR do programa H2020², o Plano S e a Coalizão S³, e a *Recomendação da UNESCO sobre ciência aberta* (Unesco, 2022)⁴, juntamente com experimentos em modelos alternativos de publicação, tais como *Preprints*, *Data Papers* e *Open Peer Review*, entre outros, provoca-

¹ Disponível em: <https://sf.dora.org/read/>. Acesso em: 13 jun. 2025.

² H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020. Disponível em: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf. Acesso em: 13 jun. 2025.

³ Disponível em: <https://www.coalition-s.org/>. Acesso em: 13 jun. 2025.

⁴ Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por. Acesso em: 13 jun. 2025.

ram uma reconfiguração significativa no ecossistema que envolve atores, normas e práticas científicas.

Do mesmo modo, a adoção de práticas de ciência aberta na América Latina tem sido objeto de estudo, revelando avanços e desafios nos últimos anos. Sánchez-Tarragó *et al.* (2016) realizaram um diagnóstico das políticas e estratégias editoriais de periódicos científicos na América Latina, destacando as barreiras e os fatores importantes para que a região adote o acesso aberto. Ribeiro, Oliveira e Santos (2021) em estudo de múltiplos casos sobre ações e estratégias de ciência aberta em universidades estaduais de São Paulo, lançando luz sobre os esforços e desafios enfrentados na implementação o acesso aberto, embora as práticas relacionadas a dados científicos abertos ainda sejam incipientes – realidade também observada no contexto brasileiro, onde apenas 0,48% da produção analisada por Caballero-Rivero, Sánchez-Tarragó e Santos (2019) mencionavam explicitamente o uso ou depósito de dados abertos. Além disso, Clinio (2019) discutiu as perspectivas contrastantes sobre o desenvolvimento da ciência aberta na América Latina, enfatizando os debates e as iniciativas em andamento na região. Ademais, Oliveira *et al.* (2022) realizaram uma pesquisa bibliográfica qualitativa para identificar o estado da arte da teoria de citação de dados na produção científica na América Latina, fornecendo um panorama sobre o cenário atual das práticas de citação de dados na região.

Esse cenário plural demonstra coletivamente o crescente interesse e os esforços para a adoção de práticas em ciência aberta, destacando a necessidade de estratégias e políticas personalizadas para promover o acesso aberto, a citação de dados e a transparência científica geral na região. Apesar dos desafios e das diferentes perspectivas, essas iniciativas significam uma mudança significativa em direção ao alinhamento aos movimentos globais de ciência aberta, promovendo um ecossistema científico mais inclusivo e transparente.

Essas tendências globais levantam diversas questões, desde a resposta às críticas sobre a mercantilização da ciência e o modelo lucrativo de publicação científica até a necessidade de abordar crescentes escândalos éticos e fraudes. O movimento em prol da ciência aberta, representa a comunicação e o compartilhamento transparente de pesquisa ao longo do ciclo de vida e ecossistema científico. A evolução recente em direção à ciência aberta mostra variações entre campos e disciplinas, com a publicação aberta ganhando espaço e se institucionalizando, resultando em infraestruturas inovadoras e novos atores identificáveis e legítimos, padrões reconhecidos e uma variedade de modelos e iniciativas.

No contexto específico dos dados científicos abertos, os princípios FAIR (*Findable, Accessible, Interoperable, and Reusable*) são a principal estrutura de orientação, complementados pelos princípios CARE (*Collective benefit, Authority to control, Responsibility, and Ethics*) da Global Indigenous Data Alliance (GIDA) (GIDA, 2019b). Contudo, a aplicação e o reconhecimento dos princípios CARE ainda são limitados. Além disso, persiste uma incerteza em relação ao significado prático dos dados abertos, FAIR e CARE nas comunidades acadêmicas e de pesquisa. Nestes ambientes, a abertura de dados é frequentemente recebida com preocupação ou resistência, destacando a necessidade de uma maior conscientização e diálogo para promover a aceitação e implementação eficaz desses princípios.

Argumenta-se que a abertura de dados deve visar o aprimoramento da ciência como um bem público global, promovendo a verificação de dados, a construção de confiança, o engajamento amplo e a aceleração de esforços para enfrentar desafios globais. No entanto, há barreiras significativas à abertura de dados, incluindo percepções equivocadas sobre as motivações para a abertura de dados, um sistema competitivo de pesquisa, desigualdades no acesso a recursos e capacidades, e os desafios específicos apresentados pelos diversos tipos de dados científicos a serem disponibilizados e preservados. Este estudo oferece uma visão geral sobre o ecossistema de atores, infraestruturas, padrões e princípios orientadores envolvidos na abertura de dados, bem como as motivações e barreiras para este processo.

4.2 ENSAIO SOBRE A ESSÊNCIA DAS MOTIVAÇÕES E BARREIRAS PARA A ABERTURA DE DADOS CIENTÍFICOS

Dados científicos abertos referem-se à prática de disponibilizar dados de pesquisa científica abertamente para o público, permitindo que qualquer pessoa os acesse, use e redistribua sem restrições. Esta prática faz parte do movimento mais amplo de dados abertos, que visa promover a transparência, a colaboração e a inovação ao tornar os dados disponíveis gratuitamente para todos (Rautenberg; Burda; Souza, 2018). O ecossistema de dados abertos envolve diversas partes interessadas, como pesquisadores, instituições, editores e o público. Ele inclui o desenvolvimento de padrões, repositórios e ferramentas para compartilhamento e reutilização de dados, bem como políticas e diretrizes para o gerenciamento e publicação de dados (Pinto *et al.*, 2017).

As motivações para a abertura de dados científicos são multifacetadas. Em primeiro lugar, os dados abertos promovem a transparência e a reprodutibilidade na pesquisa científica, permitindo que outros verifiquem e se baseiem nas descobertas existentes, aumentando assim a confiabilidade e a integridade do conhecimento científico (Rautenberg; Burda; Souza, 2018). Em segundo lugar, promovem a colaboração e a inovação ao permitir que os pesquisadores acessem e usem uma gama mais ampla de dados, levando a novas descobertas e percepções (Pinto *et al.*, 2017). Além disso, os dados abertos podem aumentar a visibilidade e o impacto da pesquisa, uma vez que os conjuntos de dados disponíveis abertamente podem ser citados e reutilizados por outros pesquisadores, potencialmente resultando em novas colaborações e citações para os autores originais (Rautenberg; Burda; Souza, 2018).

Para explorar de forma linear e crítica o tema das definições e do ecossistema que engloba os dados científicos abertos, é essencial iniciar pelo entendimento dos diversos tipos de acesso a dados. Este aspecto é fundamental, pois define a maneira como os dados são disponibilizados e acessados no universo científico.

O acesso totalmente aberto, caracterizado pela ausência de barreiras, promove a democratização do conhecimento, permitindo que qualquer pessoa, independentemente de sua filiação institucional, possa acessar, utilizar e compartilhar os dados. Contudo, é importante salientar que, embora essa abertura seja louvável em termos de promoção do conhecimento livre, ela pode levantar questões relacionadas à qualidade e segurança dos dados. A falta de restrições pode, em alguns casos, levar ao uso indevido ou à interpretação errônea dos dados, caso o processo descritivo não esteja em conformidade com os padrões internacionais disponíveis.

Por outro lado, o acesso restrito, onde o uso dos dados é limitado a certos grupos ou indivíduos, é muitas vezes necessário para proteger informações sensíveis ou confidenciais considerando questões de propriedade intelectual. Esse tipo de restrição de acesso é crucial para dados que envolvem privacidade pessoal ou segredos comerciais. No entanto, essa restrição pode ser vista como um obstáculo à colaboração e ao avanço científico, já que limita o escopo de pesquisadores que podem beneficiar-se e contribuir para o campo de estudo. Para estes casos, as orientações são de que, ao menos, os metadados devem ser disponibilizados, dando visibilidade à pesquisa realizada contemplando todo o seu escopo.

Além disso, no contexto dos dados científicos abertos, as categorias de dados baseadas na frequência de uso também merecem atenção. Dados frequentemente utilizados e referenciados, tendem a ser mais acessíveis e, consequentemente,

mais expostos a análises e críticas, o que pode levar a uma melhoria contínua na sua qualidade. Por outro lado, dados menos utilizados podem não receber a mesma atenção, potencialmente podem também não receber avaliações ou apre-moramentos.

Outro aspecto crítico é o papel dos metadados, que são informações essenciais para a descoberta e acessibilidade dos dados. Ao serem utilizados seguindo uma orientação e padronização adequada visando enriquecer o processo descritivo dos dados, os metadados permitem que os usuários compreendam a origem, o contexto e a estrutura dos dados, facilitando sua interpretação e uso corretos. A qualidade dos metadados é, portanto, um ponto crucial, pois insuficiências ou imprecisões podem levar a mal-entendidos ou a conclusões errôneas, prejudicando a confiabilidade da pesquisa científica.

Finalmente, o Plano de Gestão de Dados (PGD) surge como um elemento vital neste ecossistema uma vez que orienta a coleta, o armazenamento e o acesso aos dados, de maneira ética e eficiente. No entanto, a elaboração e implementa-ção de um PGD eficaz requerem um equilíbrio entre a proteção de dados sensí-veis em conformidade com as normas legais e éticas e a não imposição de restri-ções excessivas que possam impedir o compartilhamento e a colaboração efetiva. Agências de financiamento frequentemente exigem um PGD que detalha como os dados da pesquisa serão coletados, armazenados e acessados. Como uma fer-amenta desenvolvida especificamente para auxiliar na elaboração desses planos temos o PGD BR⁵.

Diferentes repositórios, como DANS⁶ (*Data Archiving and Networked Services*), Da-taHub⁷, Harvard-Dataverse⁸, Dryad⁹, Zenodo¹⁰ e Figshare¹¹, atendem a propósitos distinos, cada um adequado a diferentes tipos de dados e disciplinas. Todos são repositórios de dados científicos que promovem o acesso aberto e apoiam os princípios FAIR, oferecendo armazenamento, preservação e compartilhamento de dados de pesquisa em diversas disciplinas. Eles facilitam a localização, acesso, in-teroperabilidade e reutilização de dados, com metadados detalhados e formatos padronizados. Esses repositórios também suportam a preservação de longo pra-zo dos dados e oferecem funcionalidades para colaboração científica. Além disso, contribuem para novas formas de avaliar o impacto da pesquisa científica, indo

⁵ Disponível em: <https://pgd.ibict.br/>. Acesso em: 13 jun. 2025.

⁶ Disponível em: <https://dans.knaw.nl/en>. Acesso em: 13 jun. 2025.

⁷ Disponível em: <https://datahub.io>. Acesso em: 13 jun. 2025.

⁸ Disponível em: <https://dataverse.harvard.edu/>. Acesso em: 13 jun. 2025.

⁹ Disponível em: <https://datadryad.org/>. Acesso em: 13 jun. 2025.

¹⁰ Disponível em: <https://zenodo.org>. Acesso em: 13 jun. 2025.

¹¹ Disponível em: <https://figshare.com>. Acesso em: 13 jun. 2025.

além dos métodos tradicionais baseados em publicações, reforçando assim seu papel crucial na promoção da ciência aberta e na eficiência da pesquisa.

Adicionalmente, o ecossistema de dados abertos pode ser compreendido por meio de seus atores envolvidos. Destacam-se os produtores de dados, como engenheiros de dados em empresas de tecnologia brasileiras, equipes técnicas em instituições de pesquisa como o Instituto Nacional de Pesquisas Espaciais (INPE), especializado em dados ambientais e climáticos, e pesquisadores vinculados a Instituições de Pesquisa (IP), atuando em diversos campos científicos. Outros participantes incluem editores e curadores de dados, como organizações que conduzem levantamentos censitários, a exemplo do Instituto Brasileiro de Geografia e Estatística (IBGE), além de repositórios de dados científicos.

Financiadores e patrocinadores também desempenham um papel crucial, impondo requisitos de acessibilidade, exemplificados pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Os usuários de dados, por sua vez, abrangem acadêmicos de IP, formuladores de políticas públicas em ministérios governamentais e cidadãos engajados, como ativistas de Organizações Não Governamentais (ONG) que utilizam dados abertos para monitorar questões ambientais, sociais, dentre outras temáticas.

Nesse contexto, o entendimento do ciclo de vida e dos fluxos de trabalho associados aos dados científicos torna-se essencial para compreender a reutilização e valorização desses dados a longo prazo.

Durante a fase de coleta de dados, sejam em trabalhos de campo ou experimentos, os pesquisadores coletam ou geram dados, realizam pesquisas ou entrevistas, e consultam arquivos, documentando também os métodos de coleta e manipulação empregados. Após a coleta/geração, os dados passam por processamento, limpeza, combinação e, se necessário, seleção e exclusão. Documentar esse processamento é essencial para garantir controle de qualidade e replicabilidade, possibilitando a subsequente análise e interpretação dos dados.

Para publicação, pode ser necessário disponibilizar tanto os dados brutos quanto os processados, visualizações de dados e documentação, acessíveis aos revisores por meio de plataformas de dados com direitos de acesso variados. A validação dos dados pode ocorrer por meio do processo tradicional de revisão por pares ou por um processo de validação de dados específico. No final do projeto, práticas adequadas de preservação e arquivamento são essenciais para manter os dados acessíveis e reutilizáveis a longo prazo, o que inclui a descrição de metadados,

a atribuição de identificadores persistentes, como *Digital Object Identifier* (DOI), licenciamento para reutilização, armazenamento em repositórios de dados e gestão do acesso aos dados, dependendo da sua sensibilidade.

A elaboração e divulgação de artigos de dados, conhecidos como "*data papers*", que descrevem e acompanham conjuntos de dados, implicam em uma complexidade crescente de processos e fluxos de trabalho. Esses artigos são fundamentais para a disseminação e compartilhamento de conjuntos de dados, contribuindo significativamente para a transparência e reproduzibilidade da pesquisa. A complexidade intrínseca a esses procedimentos é evidenciada pelas extensas discussões relacionadas à gestão, divulgação e compartilhamento de informações científicas. Além disso, destaca-se a abordagem de desafios atuais e novas práticas que estão moldando o cenário da pesquisa científica.

4.2.1 EVOLUÇÃO DOS PRINCÍPIOS DE DADOS CIENTÍFICOS: DA FORCE11 À PRÁTICAS ATUAIS

O ecossistema de dados abertos envolve diversos participantes, como produtores de dados, editores e curadores, financiadores e usuários, cada um desempenhando papéis específicos. O ciclo de vida dos dados científicos, desde a definição da pergunta de pesquisa até a preservação a longo prazo, é essencial para entender sua reutilização e valorização.

Em 2014, a FORCE11¹² publicou a *Joint Declaration of Data Citation Principles* (JDDC) (Data Citation Synthesis Group, 2014), um marco no reconhecimento de dados como produtos valiosos da pesquisa científica. Essa declaração estabeleceu diretrizes para a citação e reutilização de dados, definindo princípios fundamentais, conforme detalhado no Quadro 4.1, para garantir a atribuição adequada e a rastreabilidade das informações.

Quadro 4.1 – Princípios para Citação e Reutilização de Dados - FORCE11

PRINCÍPIO/ PRÁTICA	DESCRIÇÃO
Importância dos Dados	Reconhece os dados como resultados valiosos e significativos da pesquisa.
Citação de Dados	Incentiva a prática de citar dados de forma adequada em publicações científicas.
Acesso e Reutilização	Promove o acesso aberto aos dados e a sua reutilização em pesquisas futuras.

¹² The Future of Research Communications and e-Scholarship. Disponível em: <https://force11.org/>. Acesso em: 13 jun. 2025.

Padrões de Citação	Estabelece padrões para a citação de dados, garantindo consistência e rastreabilidade.
Reconhecimento e Atribuição	Enfatiza a importância do reconhecimento e atribuição adequados aos produtores de dados.
Integração com Publicações	Sugere a integração de dados com publicações científicas para maior transparência e verificabilidade.
Responsabilidade e Ética	Destaca a responsabilidade e considerações éticas na utilização e citação de dados.
Fomento à Colaboração	Encoraja a colaboração entre pesquisadores e instituições para promover o compartilhamento de dados.

Fonte: Adaptado de Data Citation Synthesis Group (2014).

Com base nisso, surgiu uma série de princípios que enquadram e orientam o funcionamento dos dados abertos científicos. Os princípios FAIR foram novamente desenvolvidos pela FORCE11¹³, em 2016. Esses princípios determinam que os dados de pesquisa abertos devem atender aos seguintes requisitos (Quadro 4.2).

Quadro 4.2 – Princípios FAIR

Princípio FAIR	Descrição
<i>Findable</i> (Encontrável)	Os dados devem ser fáceis de encontrar para humanos e computadores. Metadados e dados devem ter identificadores únicos e ser descritos com metadados ricos.
<i>Accessible</i> (Acessível)	Os dados devem ser acessíveis através de protocolos padrão, abertos e gratuitos. Metadados devem permanecer acessíveis mesmo quando os dados não estiverem mais disponíveis.
<i>Interoperable</i> (Interoperável)	Os dados devem ser compatíveis e poder ser integrados com outras plataformas e ferramentas. Eles devem usar uma linguagem formal e compartilhada para representação de conhecimento.
<i>Reusable</i> (Reutilizável)	Os dados devem ser reutilizáveis e acompanhados de informações claras sobre sua proveniência, uso permitido e atender aos padrões de domínio relevantes.

Fonte: FORCE11 (2016).

Alguns anos depois, em 2019, a GIDA delineou uma estrutura chamada de Princípios CARE¹⁴ para Governança de Dados Indígenas. Esses princípios foram estabelecidos para promover os direitos de dados no contexto da *United Nations Declaration on the Rights of Indigenous Peoples* (United Nations, [2007]) e para garantir que a produção, coleta, armazenamento, difusão e citação de dados beneficiem os povos indígenas. Os princípios CARE são descritos no Quadro 4.3. A operacionalização desses princípios é fundamental. A Infraestrutura de Dados Integrados (IDI) da Nova Zelândia aborda essa questão por meio de um protocolo desenvolvido para garantir que os dados estejam em conformidade com os princípios FAIR e CARE (Carroll *et al.*, 2021).

¹³ Disponível em: <https://force11.org/info/the-fair-data-principles/>. Acesso em: 13 jun. 2025.

¹⁴ Disponível em: <https://www.gida-global.org/care>. Acesso em: 13 jun. 2025.

Quadro 4.3 – Princípios CARE

Princípio CARE	Descrição
<i>Collective Benefit</i> (Benefício Coletivo)	Os dados devem ser gerenciados de maneira que beneficiem as comunidades indígenas. Deve haver um claro retorno para essas comunidades em termos de conhecimento, inovação e avanço.
<i>Authority to Control</i> (Autoridade para Controlar)	Os povos indígenas devem ter autoridade sobre a coleta, acesso e uso dos dados que os afetam. Eles devem ter controle sobre como esses dados são usados e compartilhados.
<i>Responsibility</i> (Responsabilidade)	Há uma responsabilidade ética em como os dados sobre povos indígenas são tratados. Isso inclui proteger sua privacidade e considerar o impacto cultural e social do uso desses dados.
<i>Ethics</i> (Ética)	Os dados devem ser gerenciados de forma ética, com respeito aos valores, direitos e interesses dos povos indígenas. Isso inclui respeitar suas tradições, culturas e relações com a terra e entre si.

Fonte: GIDA (2019a).

Esses princípios são conhecidos e aceitos, mas muitas vezes não são aplicados diretamente por toda a comunidade de pesquisadores. Existem casos em que os financiadores especificam que os dados da pesquisa devem ser "abertos", porém sem detalhar alguma orientação adicional. Tem-se por exemplo no contexto brasileiro, o caso da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), que exige que um PGD seja fornecido como requisito no ato da solicitação de financiamento de um projeto. Do ponto de vista de melhores práticas, no âmbito internacional, os princípios FAIR têm sido considerados nas exigências elaboradas em alguns contextos, como para garantir financiamentos da União Europeia^{15,16} e Estados Unidos da América^{17,18}.

4.2.2 TENDÊNCIAS E DESAFIOS NA POLÍTICA DE DADOS ABERTOS DAS REVISTAS CIENTÍFICAS

No contexto das revistas científicas, poucas já apresentam uma política de dados abertos, apesar do fato de que elas exigem cada vez mais que os conjuntos de dados sejam fornecidos para os processos de revisão e para permitir que os leitores consultem os conjuntos de dados online dos artigos aceitos. É de se esperar que essa tendência emergente se generalize para a maioria dos periódicos e pode até,

¹⁵ Exemplificado em: <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm>. Acesso em: 13 jun. 2025.

¹⁶ Exemplificado em: <https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-data>. Acesso em: 13 jun. 2025.

¹⁷ Exemplificado em: <https://new.nsf.gov/funding/data-management-plan>. Acesso em: 13 jun. 2025.

¹⁸ Exemplificado em: <chrome-extension://efaidnbmnnibpcajpcgkclefindmkaj/https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf>

em alguns casos, tornar-se uma exigência. Considerando esse contexto, quais são as motivações e desafios reais para a abertura de dados?

A abertura de dados científicos é uma questão que ganha cada vez mais relevância, defendida por uma variedade de entidades na América Latina e no Brasil, como a CAPES, o CNPq, a Rede Brasileira de Pesquisa e Gestão em Desenvolvimento Territorial (RETE) e a Comissão Econômica para a América Latina e o Caribe (CEPAL). Estas organizações promovem a abertura de dados baseando-se na necessidade de transparência, acessibilidade e responsabilidade, principalmente em relação a dados financiados por recursos públicos.

Os argumentos a favor desta abertura abrangem uma variedade de aspectos, desde a precisão na verificação dos dados até a promoção de pesquisas mais amplas e profundas. Estes argumentos também incluem a construção de confiança na ciência e a valorização do processo de coleta e produção de dados. Além disso, a patrimonialização, embora seja um conceito menos conhecido, emerge como um fator crucial, considerando os dados atuais como arquivos valiosos para futuras pesquisas.

Essa abordagem reflete um crescente reconhecimento na região da importância de uma ciência aberta e colaborativa, que pode acelerar o desenvolvimento científico e tecnológico, beneficiando a sociedade como um todo. A OpenAIRE, uma infraestrutura europeia de acesso aberto para pesquisa, destaca benefícios adicionais dos dados científicos abertos para diferentes grupos: pesquisadores, financiadores, o público em geral, organizações e ONGs, e governos nacionais. Estes benefícios vão desde o aumento da capacidade de descoberta científica até a redução de custos governamentais. No entanto, é crucial reconhecer que tais benefícios dependem da capacidade de todos os envolvidos de usar e reutilizar os dados abertos, uma premissa que nem sempre é atendida. Destaca-se a importância de diversos fatores, como a patrimonialização e o agrupamento de dados. A abertura de dados não é apenas uma questão técnica, mas também envolve um aspecto cultural significativo na ciência.

4.2.3 DESAFIOS E BARREIRAS PARA A ABERTURA

O processo de abertura de dados científicos enfrenta desafios. Barreiras incluem preocupações com privacidade, confidencialidade e direitos de propriedade intelectual, além da falta de práticas padronizadas para documentação, curadoria e preservação de dados (Rautenberg; Burda; Souza, 2018). Além disso, questões

práticas relacionadas ao compartilhamento de dados também incluem a necessidade de planejamento de todo o processo de gerenciamento de dados, desde definições sobre os dados a serem coletados/gerados, padrões de metadados, até infraestrutura para armazenamento e acesso aos dados em longo prazo (Pinto *et al.*, 2017).

Novamente, o PGD neste contexto inicial de planejamento é um componente essencial nesta discussão. Ele abrange desde a descrição e coleta de dados até os requisitos legais e éticos, passando pelo compartilhamento de dados e as responsabilidades de gerenciamento. Este plano é vital para garantir que a abertura de dados seja realizada de maneira eficiente e ética, contribuindo significativamente para o avanço da ciência como um bem público global.

Ao enfrentar os desafios e obstáculos associados à abertura de dados, é imperativo concentrar a atenção em aspectos menos explorados. Um dos principais desafios reside nas motivações para a pesquisa aberta, frequentemente mal interpretadas, especialmente no contexto das ciências sociais, onde pode haver redução do debate a uma dicotomia entre obrigação institucional versus altruísmo científico. Outra questão crítica é a tensão entre a colaboração e a competitividade no ambiente acadêmico. A pesquisa de dados abertos, que tende a ser mais colaborativa, pode entrar em conflito com a cultura competitiva alimentada por rankings e publicações em periódicos de alto impacto.

Os custos e a escassez de recursos para gerenciar dados abertos também representam barreiras consideráveis. Dados frequentemente são produzidos por equipes pequenas e com recursos limitados. O gerenciamento de dados abertos requer tempo e esforço adicionais, criando uma possível desconexão entre as expectativas dos pesquisadores e as realidades práticas do trabalho.

Além disso, é relevante destacar os desafios específicos enfrentados por pesquisadores e periódicos independentes. Aspectos como a preparação de dados para compartilhamento, a seleção de um repositório apropriado e a criação de uma declaração de disponibilidade de dados tornam-se fundamentais no processo de abertura de dados.

A questão da abertura de dados científicos, no contexto da comunicação científica formal, cada vez mais relevante, coloca desafios específicos para revistas independentes ou sem fins lucrativos, principalmente aquelas que lidam com a publicação de dados científicos abertos. Gerenciar e revisar esses dados é uma tarefa especializada e distinta da publicação acadêmica tradicional. Por isso, não se pode

esperar que essas revistas, muitas das quais operam com recursos limitados, forneçam a infraestrutura necessária para esse fim.

Este contexto levanta discussões importantes sobre a ciência aberta e a sua relação com a publicação acadêmica. Uma das questões centrais é se devemos permitir a obtenção de lucro com Dados Científicos ou com a publicação de pesquisa. Há uma crescente necessidade de assegurar que qualquer uso comercial de dados científicos beneficie principalmente organizações públicas, universidades e cidadãos. Buscando tentar mitigar eventuais dificuldades quanto à abertura de dados por parte das revistas, o Quadro 4.4 a seguir apresenta orientações quanto a uma possível política de abertura de dados em um periódico científico.

Quadro 4.4 - Roteiro para definir uma Política de Dados para periódicos

Tópico Geral	Questões e assuntos a serem abordados
Definição de Dados Científicos e exceções	Quais são as condições de acesso e embargo? Quais dados estão inclusos na política? Em que etapa do processo? Submissão? Aceitação?
Padrões e formatos de dados e metadados	Quais padrões devem ser usados e quais são os protocolos de metadados?
Acesso, hospedagem e publicação de dados	Quais são os protocolos e diretrizes para depositar dados e escolher um repositório?
Procedimentos de disponibilidade de dados	Quais são os procedimentos de disponibilidade de dados (prazos, etapas, etc.)? Os dados são revisados por pares e por quem? Existe uma declaração de disponibilidade de dados?
Acessibilidade dos dados	Como os dados serão conectados à publicação? Identificadores permanentes são usados?
Não conformidade	O que acontece quando os autores não cumprem a política de dados do periódico?
Supporte para autores, revisores e editores	O periódico ou sua associação acadêmica oferece algum tipo de suporte?

Fonte: Os autores (2023).

Para a gestão de dados, é recomendável que as revistas incentivem o uso de formatos abertos e padronizados, bem como metadados estruturados usando padrões reconhecidos, como o Dublin Core. Além disso, aconselha-se que os dados que contribuíram para a elaboração do artigo sejam depositados em repositórios confiáveis que garantam o armazenamento seguro e o acesso aos dados.

O procedimento para a disponibilidade de dados pode ser dividido em três fases: a fase da submissão, a fase de revisão por pares e a fase de aceitação. Em todas as etapas, é imperativo garantir a anonimização dos dados e aderir rigorosamente às normas de privacidade e confidencialidade.

Ao concluir o processo, é esperado que os autores incorporem uma declaração específica de acessibilidade de dados em suas publicações, descrevendo os métodos para acessar os dados e fornecendo *links* permanentes para os mesmos. Esta prática é vital para promover a transparência e a integridade na pesquisa científica, alinhando-se com as tendências globais de ciência aberta e colaborativa.

No que tange à política de dados abertos, é importante que os dados associados a publicações científicas estejam acessíveis de maneira aberta, sempre que possível, e sob as condições de restrição necessárias. Os princípios FAIR e CARE são essenciais nesse processo. Enquanto algumas revistas incentivam a disponibilidade de dados, outras adotam uma postura mais ativa para aumentar a conscientização sobre a importância de integrar esses princípios.

Na América Latina, as políticas de dados abertos podem se inspirar em diretrizes internacionais, como as da Research Data Alliance¹⁹ ou do Comitê para a Ciência Aberta da França²⁰, adaptando-as à realidade e aos desafios locais. Isso inclui considerar as evoluções institucionais e as especificidades do campo científico da região.

Finalizando, a definição de dados científicos abertos deve ser clara e abrangente, incluindo todo o escopo do conjunto de dados necessários para verificar e reproduzir os resultados apresentados nas publicações. Estes dados podem variar em formato - imagem, vídeo, texto, código, tabela estatística, dentre outros - e devem ser produzidos e compartilhados de acordo com os princípios FAIR e CARE apresentados anteriormente.

4.3 CONSIDERAÇÕES FINAIS

As tendências recentes na pesquisa científica internacional destacam os esforços para promover a abertura ao longo de todo o ecossistema e ciclo de vida dos dados científicos, desde a coleta até a publicação. No entanto, essa abertura apresenta desafios e preocupações específicas. Este capítulo proporciona uma visão geral do ecossistema de atores, infraestruturas, padrões e princípios orientadores envolvidos na abertura de dados, juntamente com as motivações e barreiras associadas a esse processo.

¹⁹ Disponível em: <https://www.rd-alliance.org/>. Acesso em: 13 jun. 2025.

²⁰ Disponível em: <https://www.ouvrirlascience.fr/accueil/>. Acesso em: 13 jun. 2025.

Diversos desafios significativos foram identificados. Entre eles, destaca-se uma compreensão inadequada das motivações para a abertura de dados. Outro obstáculo reside em um sistema de pesquisa permeado pela competição, que desestimula a colaboração e o compartilhamento. Adicionalmente, a escassez de recursos e capacidades técnicas tende a acentuar desigualdades no acesso e na participação. Por fim, persistem riscos inerentes à abertura de dados, como os custos associados ao uso de diversos formatos de dados ou a perda de acesso a determinadas áreas de pesquisa.

Uma reflexão foi apresentada sobre as condições operacionais e as questões relacionadas à quando e como abrir dados científicos, tanto do ponto de vista do pesquisador quanto do periódico. O estudo enfatiza a importância de abrir dados sempre que possível, especialmente diante de emergências socioecológicas e na luta contra desigualdades. No entanto, reconhece-se o esforço e o tempo necessários para desenvolver projetos de pesquisa, implementar Planos de Gestão de Dados (PGDs) e publicar em revistas de acesso aberto, enquanto se promove a abertura dos dados.

Defende-se neste estudo, portanto, a ideia de uma "academia sustentável" e "ciéncia lenta", acrescentando a necessidade de reflexão e cuidado com uma preocupação ética na pesquisa, onde o sujeito ou objeto de estudo pode ser indevidamente influenciado ou dominado pelos interesses ou perspectivas do pesquisador. Isso pode envolver questões como viés na pesquisa, falta de consentimento informado dos participantes ou imposição de narrativa, ou conclusão, que não reflete adequadamente os dados ou as experiências dos participantes, enfatizando a importância de explorar e abrir espaço para perspectivas diversas e não dominantes.

A reflexão sobre abordagens éticas para a governança de dados é um exemplo que nos auxilia a repensar métodos atuais e reconhecer como produzimos e perpetuamos relações de poder e hierarquias de conhecimento na geração, utilização e gerenciamento de dados científicos. Concluímos, portanto, que a abertura de dados transcende a mera transparência ou reproduzibilidade, representando também um compromisso com o acesso inclusivo e equitativo ao conhecimento para todos.

REFERÊNCIAS

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília: Presidência da República, [2020]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Lei/L13709.htm. Acesso em: 25 maio 2025.

CABALLERO-RIVERO, A.; SÁNCHEZ-TARRAGÓ, N.; SANTOS, R. N. M. Práticas de Ciência Aberta da comunidade acadêmica brasileira: estudo a partir da produção científica. **TransInformação**, Campinas, v. 31, e190029, p. 1-14, 2019. DOI: <https://doi.org/10.1590/2318-0889201931e190029>. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/5930>. Acesso em: 27 maio. 2025.

CARROLL, S. R.; HERCZOG, E.; HUDSON, M.; RUSSELL, K.; STALL, S. Operationalizing the CARE and FAIR principles for Indigenous data futures. **Scientific data**, [s. l.], v. 8, n. 108, p. 1-6, 2021. DOI: <https://doi.org/10.1038/s41597-021-00892-0>. Disponível em: <https://www.nature.com/articles/s41597-021-00892-0>. Acesso em: 27 maio 2025.

CLINIO, A. Ciência Aberta na América Latina: duas perspectivas em disputa. **TransInformação**, Campinas, v. 31, e190028, 2019. DOI: <https://doi.org/10.1590/238180889201931e190028>. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/5908>. Acesso em: 6 dez. 2023.

DANS. **Data Archiving and Networked Services**. [s. l.], 2021. Disponível em: <https://dans.knaw.nl/en>. Acesso em: 7 dez. 2023.

DATA CITATION SYNTHESIS GROUP. **Joint Declaration of Data Citation Principles**. Martone M. (ed.) San Diego, CA: FORCE11; 2014. DOI: <https://doi.org/10.25490/a97f-egyk>. Disponível em: <https://force11.org/info/joint-declaration-of-data-citation-principles-final/>. Acesso em: 10 dez. 2023.

DATAHUB. Datahub. [s. l.], 2021. Disponível em: <https://datahub.io>. Acesso em: 7 dez. 2023.

DATAVERSE. **The Dataverse Project**. [s. l.], 2021. Disponível em: <https://dataverse.org>. Acesso em: 7 dez. 2023.

DORA. San Francisco Declaration on Research Assessment. Atena, 2012. Disponível em: <https://sfdora.org/read/>. Acesso em: 6 dez. 2023.

DRYAD. **Dryad Digital Repository**. [s. l.], 2021. Disponível em: <https://datadryad.org/stash>. Acesso em: 7 dez. 2023.

EUROPEAN UNION. **General Data Protection Regulation.** [s. l.]: Intersoft Consulting, 2016. Disponível em: <https://gdpr-info.eu/>. Acesso em: 27 nov. 2023.

FIGSHARE. **Figshare.** [s. l.], 2023. Disponível em: <https://figshare.com>. Acesso em: 7 dez. 2023.

FORCE11. **The FAIR Data Principles.** Davis: Force11, [2016]. Disponível em: <https://force11.org/info/the-fair-data-principles/>. Acesso em: 10 dez. 2023.

GLOBAL INDIGENOUS DATA ALLIANCE. **CARE Principles for Indigenous Data Governance.** [s. l.], 2019. Disponível em: <https://www.gida-global.org/care>. Acesso em: 10 dez. 2023.

GLOBAL INDIGENOUS DATA ALLIANCE. **Promoting Indigenous Control of Indigenous Data.** [s. l.], 2019. Disponível em: <https://www.gida-global.org/>. Acesso em: 7 dez. 2023.

OLIVEIRA, C. C.; SILVA, M. C.; PAVÃO, C. M. G.; SILVA, F. C. C.; MOURA, A. M. M.; BARROS, T. H. B. A teoria da citação de dados: uma revisão da produção científica na América Latina. **TransInformação**, Campinas, v. 34, e210062, 2022. DOI: <https://doi.org/10.1590/2318-0889202234e210062>. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/6487>. Acesso em: 22 abr. 2025.

PINTO, D. M.; BRANDÃO, V. V. S.; DRUCKER, D. P.; DOMPIERI, M. H. G.; RASCHE, F.; HOLLER, W. A.; BETTIOL, G. M.; CUSTÓDIO, D. O.; VICTORIA, D. C.; GONÇALVES, L. M. P. B. Descrição dos dados da pesquisa geoespacial: a experiência da Embrapa. **RECIIS: Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, [s. l.], v. 11, supl., nov. 2017. DOI: <https://doi.org/10.29397/reciis.v11i0.1418>. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/1418>. Acesso em: 11 dez. 2023.

PLAN S. **About Plan S.** France: Plan S, [2018]. Disponível em: <https://www.coalition-s.org>. Acesso em: 6 dez. 2023.

RAUTENBERG, S.; BURDA, A. C.; SOUZA, L. Um workflow para compartilhamento de dados científicos primários baseado em Dados Abertos Conectados.

Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação, [s. l.], v. 23, n. 53, p. 110-123, set./dez. 2018. DOI: <https://doi.org/10.5007/1518-2924.2018v23n53p110>. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p110>. Acesso em: 7 dez. 2023.

RIBEIRO, N. C.; OLIVEIRA, D. A.; SANTOS, S. R. O. Ações e estratégias voltadas para a ciência aberta em universidades estaduais paulistas: um estudo multicaso.

Em Questão, Porto Alegre, v. 27, n. 2, p. 164-192, abr./jun. 2021. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/102078>. Acesso em: 06 dez. 2023.

SÁNCHEZ-TARRAGÓ, N.; CABALLERO-RIVERO, A.; TRZESNIAK, P.; DEROY DOMÍNGUEZ, D.; SANTOS, R. N. M.; FERNÁNDEZ-MOLINA, J-C. Las revistas científicas en América Latina hacia el camino del acceso abierto: un diagnóstico de políticas y estrategias editoriales. **TransInformação**, Campinas, v. 28, n. 2, p. 159-172, maio/ago. 2016. DOI: <https://doi.org/10.1590/2318-08892016000200003>. Disponível em: <https://www.scielo.br/j/tinf/a/bg4phnKWMYkhvxxbVFMmFHf/?lang=es>. Acesso em: 6 dez. 2023.

UNESCO. **Ciência aberta no Brasil**. Brasília: DF: UNESCO, 2023. Disponível em: <https://www.unesco.org/pt/fieldoffice/brasilia/expertise/open-science-brazil>. Acesso em: 6 dez. 2023.

UNITED NATIONS. **Declaration on the Rights of Indigenous Peoples**. [s. l.], [2007]. Disponível em: https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDROP_E_web.pdf. Acesso em: 7 dez. 2023.

ZENODO. **Zenodo**. [s. l.], 2021. Disponível em: <https://zenodo.org>. Acesso em: 7 dez. 2023.

Como citar este capítulo:

SILVA, Fabiano Couto Corrêa da; REZENDE, Laura Vilela Rodrigues; LUCAS, Elaine Rosangela de Oliveira. Além do acesso aberto: desafios e perspectivas na era dos dados científicos abertos. In: ARAÚJO, Paula Carina de; LIMA, Karolayne Costa Rodrigues de (org.). **Práticas de ciência aberta**. Brasília, DF: Editora Ibit, 2025. Cap. 4, p. 78-95. DOI: 10.22477/9788570131966.cap4.