Tendências da Pesquisa Brasileira em Ciência da Informação

DISPONIBILIDADE DE DADOS DE PESQUISA: UMA ANÁLISE DOS DATA PAPERS PUBLICADOS POR AUTORES FILIADOS ÀS INSTITUIÇÕES BRASILEIRAS[1]

AVAILABILITY OF RESEARCH DATA: AN ANALYSIS OF DATA PAPERS PUBLISHED BY AUTHORS AFFILIATED WITH BRAZILIAN INSTITUTIONS

DISPONIBILIDAD DE DATOS DE INVESTIGACIÓN: UN ANÁLISIS DE ARTÍCULOS DE DATOS PUBLICADOS POR AUTORES AFILIADOS A INSTITUCIONES BRASILEÑAS

Crislaine Silveira[2]
Thiago Magela Rodrigues Dias[3]
Patrícia Neubert[4]

Submetido em: 23/03/2026
Aprovado em: 10/07/2026
Publicado em: 31/07/2026

Artigo submetido ao sistema de similaridade turnitin

[1] O presente artigo amplia e aprofunda versão anterior que foi submetida, avaliada, aprovada e apresentada no XXV ENANCIB (Encontro Nacional de Pesquisa em Ciência da Informação) ocorrido no Rio de Janeiro (RJ), no período de 3 a 7 de novembro de 2025.

[2] Doutora em Ciência da Informação. E-mail: crislaine.silveira@ufsc.br. ORCID: https://orcid.org/0000-0003-3081-9968.

[3] Doutor em Modelagem Matemática e Computacional. Professor no Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG). E-mail: thiagomagela@gmail.com. ORCID: https://orcid.org/0000-0001-5057-9936.

[4] Docente do Departamento de Ciência da Informação, da Universidade Federal de Santa Catarina (CIN/UFSC). E-mail: patricia.neubert@ufsc.br. ORCID: https://orcid.org/0000-0002-8909-1898.

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional.

Resumo: os *data papers* têm como finalidade descrever detalhadamente os conjuntos de dados de pesquisa e disponibilizá-los em formato bruto para os pesquisadores interessados. Entender como funciona a disponibilidade dos dados associados a eles é importante pois pode ajudar a editores de periódicos a elaborarem políticas de compartilhamento de dados. Este artigo tem por objetivo verificar de que maneira os conjuntos de dados de pesquisa associados aos *data papers* estão sendo disponibilizados e se o compartilhamento destes dados segue as diretrizes específicas do periódico. Para a coleta dos dados foi realizada uma pesquisa bibliográfica cujo objetivo foi identificar *data papers* com pelo menos um autor filiado a uma instituição brasileira, nas bases de dados Scopus e Web of Science. Em seguida foram realizadas duas pesquisas documentais, sendo a primeira nos 634 *data papers* identificados na coleta anterior para identificar o tipo de disponibilidade dos dados de pesquisa e a segunda nas políticas dos periódicos onde os *data papers* foram publicados para verificar se as diretrizes foram seguidas. Foram identificadas diversas formas de disponibilização, por exemplo: repositórios, material suplementar, páginas pessoais ou institucionais. Em relação a adesão as diretrizes dos periódicos quanto a tipologia de repositórios, foi verificado que embora os periódicos indiquem os repositórios, os autores tendem a não as seguir, sendo influenciados pelo comportamento da área. Portanto, a publicação de *data papers* carece de maior aderência às políticas editoriais relacionadas ao compartilhamento de dados, particularmente no que tange à escolha adequada dos repositórios.

Palavras-chave: repositório de dados; publicação de dados de pesquisa; artigo de dados.

Abstract: *Data papers aim to describe research datasets in detail and make them available in raw format to interested researchers. Understanding how the availability of the data associated with them works is important because it can help journal editors develop data sharing policies. This article aims to verify how research datasets associated with data papers are being made available and whether the sharing of this data follows the specific guidelines of the journal. Data collection involved a bibliographic search to identify data papers with at least one author affiliated with a Brazilian institution in the Scopus and Web of Science databases. Subsequently, two documentary searches were conducted: the first on the 634 data papers identified in the previous collection to identify the type of availability of the research data, and the second on the policies of the journals where the data papers were published to verify if the guidelines were followed. Several forms of availability were identified, for example: repositories, supplementary material, personal or institutional pages. Regarding adherence to journal guidelines concerning repository types, it was found that although journals indicate repositories, authors tend not to follow them, being influenced by the behavior of the field. Therefore, the publication of data papers lacks greater adherence to editorial policies related to data sharing, particularly regarding the appropriate choice of repositories.*

Keywords: *data repository; research data publication; data paper.*

***Resumen**: Los artículos de datos buscan describir conjuntos de datos de investigación en detalle y ponerlos a disposición de los investigadores interesados en formato original. Comprender cómo funciona la disponibilidad de los datos asociados a ellos es importante, ya que puede ayudar a los editores de revistas a desarrollar políticas de intercambio de datos. Este artículo busca verificar cómo se ponen a disposición los conjuntos de datos de investigación asociados a los artículos de datos y si su intercambio cumple con las directrices específicas de la revista. La recopilación de datos implicó una búsqueda bibliográfica para identificar artículos de datos con al menos un autor afiliado a una institución brasileña en las bases de datos Scopus y Web of Science. Posteriormente, se realizaron dos búsquedas documentales: la primera sobre los 634 artículos de datos identificados en la recopilación anterior para identificar el tipo de disponibilidad de los datos de investigación, y la segunda sobre las políticas de las revistas donde se publicaron los artículos de datos para verificar si se siguieron las directrices. Se identificaron diversas formas de disponibilidad, por ejemplo: repositorios, material suplementario, páginas personales o institucionales. En cuanto al cumplimiento de las directrices de las revistas sobre los tipos de repositorio, se observó que, si bien las revistas indican los repositorios, los autores tienden a no seguirlas, influenciados por el comportamiento del campo. Por lo tanto, la publicación de artículos de datos carece de una mayor adhesión a las políticas editoriales relacionadas con el intercambio de datos, particularmente en lo que respecta a la elección adecuada de repositorios.*

***Palabras clave**: repositorio de datos; publicación de datos de investigación; artículo de datos.*

1 INTRODUÇÃO

Os *data papers* serão considerados documentos publicados por periódicos científicos, cuja finalidade é descrever detalhadamente os conjuntos de dados de pesquisa e disponibilizá-los em formato bruto para os pesquisadores interessados.

Considerando que o objetivo central dos *data papers* é descrever os conjuntos de dados conferindo visibilidade a eles, é mister a necessidade de entender como funciona o compartilhamento dos dados aos quais eles se referem e verificar se os autores seguem as recomendações dos periódicos quanto à sua disponibilização nos repositórios de dados.

Visando compreender de que maneira os dados de pesquisas são compartilhados por meio dos artigos científicos, Federer *et al.* (2018), realizaram uma pesquisa no periódico Plos One e identificaram diferentes maneiras, as quais são: a) a disponibilização dos dados como material suplementar; b) a disponibilização mediante a solicitação ao autor; c) a disponibilização no próprio artigo e d) de maneira combinada, na qual os dados são disponibilizados em repositórios de dados e como material suplementar.

Jiao, Li e Fang (2022) também examinaram essas diferentes maneiras de compartilhamento dos dados de pesquisa em artigos científicos, mas enfatizaram a análise nos repositórios de dados de pesquisa.

No entanto, as pesquisas desses autores, se limitaram a análise das declarações de disponibilidade de dados, disponíveis nos artigos científicos publicados pelo periódico Plos One. Neste trabalho, o objetivo é realizar essa análise a partir de um tipo documental diferente dos artigos científicos, os *data papers* para verificar se essa tendência permanece, uma vez que esse tipo documental prescinde da disponibilidade dos dados que esses o descrevem. Com isso, o diferencial deste trabalho é identificar a disponibilidade dos dados de pesquisa, tendo como recorte, os *data papers* publicados por autores filiados às instituições brasileiras.

Portanto, vislumbrando uma possível contribuição para os editores de periódicos, na elaboração das políticas relacionadas à disponibilidade dos dados de pesquisa vinculados às publicações, e para os pesquisadores, cuja disponibilidade pode contribuir para a encontrabilidade e reusabilidade dos dados de suas pesquisas e das publicações associadas aos conjuntos de dados, este trabalho tem por

objetivo verificar de que maneira os conjuntos de dados de pesquisa associados aos *data papers* estão sendo disponibilizados e se o compartilhamento destes dados segue as diretrizes específicas do periódico.

2 A DISPONIBILIDADE DE DADOS EM DATA PAPERS

Os *data papers* fornecem descrições abrangentes de conjuntos de dados, vinculam sua localização, criando um caminho para o uso dos dados de pesquisa, incentivando assim seu compartilhamento, a reprodutibilidade da pesquisa, aprimorando o seu valor desses dados dentro do ecossistema acadêmico, além de oferecer um mecanismo para crédito de citação aos autores (Li; Huang; Jeng, 2025).

Os *data papers* devem sempre ter um conjunto de dados depositado num repositório de dados online e estável, além de serem descritos por metadados, pois esses são parte essencial do ecossistema da publicação dos dados de pesquisa. Uma característica desse tipo de publicação, é que os dados devem estar disponíveis publicamente, pois representam uma garantia de acessibilidade e qualidade, além da revisão por pares (Candela *et al*., 2015; Puerta-Piñero; Pérez-Luque; Rodríguez-Echeverría, 2020).

Pode-se dizer que o caminho para o compartilhamento e uso dos dados é operacionalizado pelas políticas adotadas pelos periódicos científicos, que devem definir de que maneira os dados de pesquisa associados, seja a artigos científicos ou *data paper*, deve ser disponibilizada.

Federer *et al.* (2018) identificou que os mecanismos de disponibilidade de dados em periódicos variavam, perpassando pelo compartilhamento mediante solicitação aos autores, disponibilização no artigo e/ou como materiais suplementares, até o uso do repositório

de dados. Na primeira abordagem, mediante solicitação, o inconveniente é que muitos autores não respondem a tais solicitações enviadas por outros ou não estão dispostos a compartilhar (Federer *et al*., 2018; Federer, 2022). O compartilhamento de dados no artigo é insuficiente porque os dados compartilhados tendem a ser dados resumidos quando analisados, não permitindo o acesso aos dados brutos que foram coletados pelos pesquisadores (Federer *et al*., 2018).

A desvantagem de disponibilizar os dados de pesquisa como material suplementar, é que esses requerem curadoria e preservação, o compartilhamento dessa maneira não permite que os leitores encontrem o seu link independentemente da publicação principal (Candela *et al*., 2015).

Por conseguinte, os repositórios de dados acabam sendo a melhor opção, pois fornecem armazenamento a longo prazo e uso de identificadores persistentes, tirando o ônus do pesquisador de manter os links continuamente ativos, de modo que os dados estejam mais propensos a se manterem acessíveis ao longo do tempo, além de fazerem uso de metadados para a melhor descrição dos objetos (Federer *et al*., 2018; Federer, 2022).

Depositar conjuntos de dados em repositórios estáveis, públicos e que seguem padrões internacionais é fundamental para assegurar os princípios FAIR, que orientam que os dados sejam localizáveis (Findable), acessíveis (Accessible), interoperáveis (Interoperable) e reutilizáveis (Reusable) (Wilkinson *et al*., 2016). A ausência dessa prática compromete a transparência científica, limita o reuso dos dados e enfraquece a integridade da comunicação científica. Portanto,

os repositórios de dados se tornam agentes mediadores-chave entre as publicações e os dados de pesquisa (Jiao; Li; Fang, 2022).

Portanto, analisar a disponibilidade dos conjuntos de dados pode contribuir para a otimização, a preservação e a confiabilidade de todo o ecossistema de comunicação científica que abarca os dados de pesquisa.

3 PROCEDIMENTOS METODOLÓGICOS

Quanto aos objetivos a pesquisa se caracteriza como exploratória e descritiva, pois visa conhecer, a partir dos *data papers,* as maneiras utilizadas para disponibilizar os dados bem identificar se esses seguem as políticas adotadas pelos periódicos, a partir da descrição da tipologia de repositório a ser utilizada.

Para coleta dos dados foi utilizada a pesquisa bibliográfica e documental. Na primeira o objetivo foi identificar os *data papers* com pelo menos um autor filiado às instituições brasileiras. Essa pesquisa, ocorreu em duas etapas: a primeira nas bases de dados Scopus e Web of Science, para identificar os *data papers* publicados por autores filiados às instituições brasileiras e a segunda, no catálogo do OpenAlex, para extrair os relatórios com dados sobre *data papers* identificados na primeira coleta de modo a enriquecer os dados da pesquisa.

Na segunda, foi realizada uma pesquisa nos 634 *data papers* identificados na coleta anterior para: a) identificar a quantidade de conjuntos dados mencionada por cada um, essa contabilização se deu pela quantidade do Digital Object Identifier (DOI) ou Handles, mencionados como depósitos em cada *data paper*, procedimento possível, pois os pesquisadores frequentemente usam links da web e

DOIs para se referir a um repositório (Jiao; Li; Fang, 2022) e verificar as formas de disponibilização dos dados de pesquisa associados aos *data papers*.

Além disso, foi realizada uma pesquisa nas políticas de compartilhamento de dados ou outra política com nome similar dos periódicos que publicaram os *data papers* identificados na pesquisa bibliográfica, cujo objetivo foi identificar as orientações referentes a disponibilidade dos dados e os repositórios a serem utilizados.

Após identificar a forma de disponibilização dos conjuntos de dados nos data papers, elas foram classificadas de acordo com a proposta de Jiao, Li e Fang (2022), a qual contempla o tipo de disponibilização: a) em repositórios: essa caracterização se deu a partir da menção do nome do repositório ou do(s) link(s) de acesso aos dados de pesquisa; b) como material suplementar: quando os arquivos contendo os conjuntos de dados eram inseridos numa seção específica do *data paper*, por exemplo: apêndice, informação de suporte, registro e disponibilidade dos dados, etc.; c) em páginas pessoais ou institucionais: a identificação foi realizada a partir da visita ao link disponibilizado; d) não disponibilizado: quando não foi mencionado nenhum conjunto de dados ao longo do *data paper* ou quando o *data paper* informava que o conjunto estava adicionado ao longo do texto e por fim, e) como material suplementar e repositório: quando mencionava as duas formas para disponibilizar os dados.

A partir da identificação dos 26 periódicos, nos quais foram publicados os 349 *data papers* que publicaram os dados em repositórios, outro levantamento documental foi realizado, cujo objetivo foi verificar nas políticas de compartilhamento de dados dos periódicos as orientações referentes a disponibilidades dos dados e os

repositórios recomendados. A tipologia dos repositórios indicados para o depósito dos conjuntos de dados foi verificada no *Registry of Research Data Repositories* (R3data).

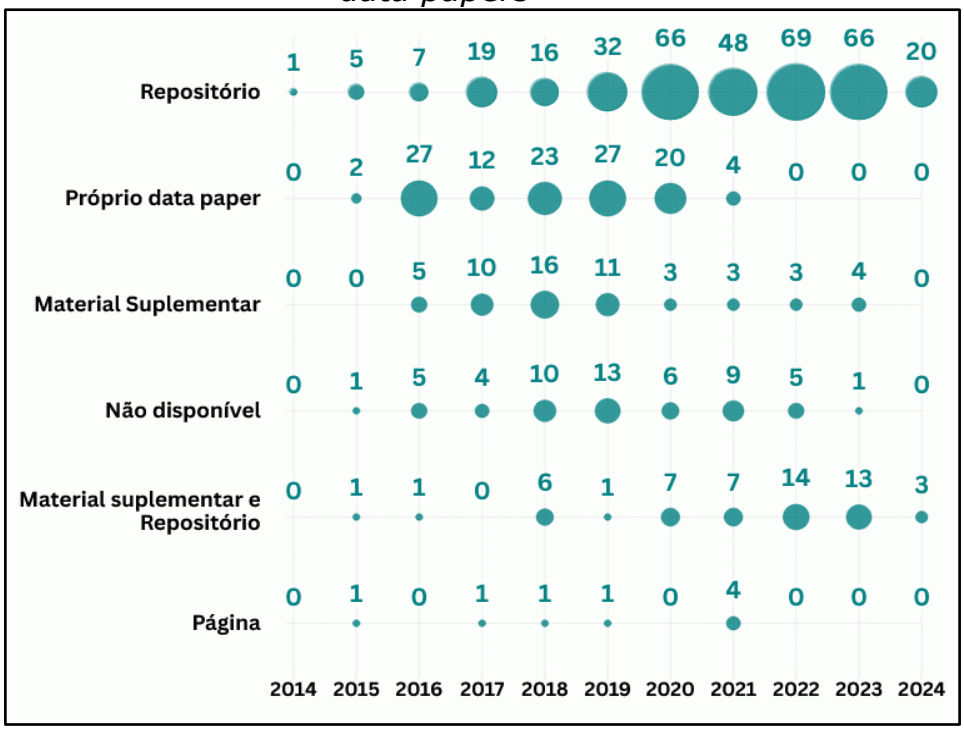
A análise dos dados se deu com a utilização dos métodos mistos, nos quais a abordagem quantitativa utiliza estatística descritiva, contabilizando a frequência, a porcentagem e o cálculo da média. E a análise qualitativa se baseou na literatura para fornecer sentido aos dados quantitativos.

Dentre as limitações desse estudo destaca-se: a análise apenas da tipologia dos repositórios e não dos repositórios em si, a consulta às diretrizes dos periódicos, considerou apenas a política vigente, não abarcando as políticas editoriais anteriores, o que pode influenciar na análise da disponibilidade, pois as orientações podem mudar ao longo do tempo. Nos casos em que os dados foram disponibilizados de forma combinada, ou seja, como material suplementar e em repositório, é importante mencionar que não foi realizada a análise qualitativa dos conjuntos de dados para verificar se eram os mesmos arquivos ou não.

4 RESULTADOS E DISCUSSÃO

Conforme mencionado anteriormente a disponibilidade dos dados de pesquisa pode variar nos artigos científicos, conforme identificado nos estudos conduzidos por Federer *et al*. (2018) e Jiao, Li e Fang (2022). No caso dos *data papers* essa tendência seguiu, sendo percebidas mudanças ao longo do tempo, de acordo com a Figura 1.

Figura 1 - Formas de publicação dos dados de pesquisa associados aos *data papers*



Fonte: elaborada pelos autores

A disponibilidade no **próprio *data paper***, teve um auge em 2016 com 27 *data papers* e manteve relevância até 2019, também com 27 *data papers*, mas começa a reduzir a partir de 2021, até ser uma prática não mais adotada a partir de 2022. No total foram identificados que em 18,14% (115) dos casos, os dados foram disponibilizados no próprio *data paper.*

A forma de disponibilizar como **material suplementar**, cresce até 2018, com 16 *data papers*, mas depois perde força nos anos seguintes. No total os dados foram disponibilizados como material suplementar em 8,68% (55) dos *data papers*.

A não disponibilização dos dados brutos, teve um pico de 13 *data papers*, em 2019, mas vem perdendo força nos anos posteriores. No total os dados não foram disponibilizados em 8,52% (54) dos *data papers*.

A disponibilização combinada como **material suplementar e repositórios** ganha força a partir de 2020, com sete *data papers* e picos em 2021 (14) e 2022 (13). Em 8,36% (53) dos *data papers*, os conjuntos de dados foram disponibilizados de maneira combinada, tanto em repositórios como em material suplementar.

O uso de **páginas pessoais ou institucionais** é irregular, com um pico de quatro *data papers*, em 2021. Em 1,26% (8) dos *data papers*, os dados foram disponibilizados em páginas pessoais ou institucionais.

Com isso é possível observar que a tipologia utilizada para a disponibilização dos dados, vem variando desde a publicação do primeiro *data paper*, em 2014, até o ano da coleta dos dados que originou o presente estudo, 2024. Essa heterogeneidade vem diminuindo ao longo do tempo, o que sugere que uma adesão irregular por parte dos *publishers* às práticas de abertura da Ciência.

O uso de **repositórios** para disponibilizar os dados de pesquisa é a categoria que vem crescendo e dominando, principalmente a partir de 2019, enquanto que as outras formas de disponibilização estão sendo reduzidas. E em 55,05% (349) dos casos os conjuntos de dados foram disponibilizados em repositórios.

Esses dados podem se relacionar com a adoção, pelos periódicos, de orientações claras a respeito da disponibilidade do conjunto de dados em repositórios. A tendência de crescimento do uso de repositórios ao longo dos anos também foi identificada nos estudos de Jiao, Li e Fang (2022).

O uso de repositórios para a publicação dos dados de pesquisa foi identificado em 26 dos 38 periódicos que publicaram os *data*

papers. O tipo de disponibilidade adotada por cada periódico, está descrito na Tabela 1.

Tabela 1 - Tipo de disponibilidade dos dados conforme os periódicos

...continua

Tipo de disponibilidade	F* periódicos	Periódicos (*Data Paper*)**
Repositórios	26	*Data in brief (164), Scientific data (82), Earth system science data (34), Data (16), Biodiversity Data Journal (15), Gigascience (5), BMC research notes (5), Chemical data collections (5), Geoscience data journal (30), Frontiers in ecology and evolution (2), Microbiology resource announcements (2), Frontiers in neuroscience (2), Proteomics (1), Frontiers in genetics (1), Annals of forest Science (1), Big earth data (1), Ecology (1), Ecosistemas (1), Frontiers in marine Science (1), Frontiers in psychology (1), Frontiers in public health (1), Global ecology and biogeography (1), Journal of open humanities data (1), Marine genomics (1), Scientific reports (1), ZooKeys (1).*
Dados não foram disponibilizados	14	*Data in brief* (27), *Scientific data* (7), *Chemical data collections* (7), *Frontiers in Marine Science* (2), *Frontiers in Earth Science* (2), *Earth system science data* (1), *Biodiversity Data Journal* (1), *Frontiers in ecology and evolution* (1), *Marine genomics* (1), *Cahiers Agricultures* (1), *Frontiers in microbiology* (1), *Frontiers in molecular biosciences* (1), *Human mutation* (1), *Taxon* (1).
Material suplementar	12	*Ecology* (19), *Scientific data* (13), *Data in Brief* (12), *Biodiversity Data Journal* (2), *Chemical data collections* (2), *Earth system science data* (1), *Data* (1), *Frontiers in genetics* (1), *ZooKeys* (1), *Frontiers in bioinformatics* (1), *Nature Conservation* (1), *Naturwissenschaften* (1).
Repositórios como em material suplementar	7	*Ecology* (31), *Global Ecology and Biogeography* (8), *Biodiversity Data Journal* (7), *Proteomics* (3), *Data in Brief*

Tipo de disponibilidade	F* periódicos	Periódicos (*Data Paper*)**
		(2), *Scientific data* (1), *Journal of biogeography* (1).
Páginas pessoais ou institucionais	5	*Data in brief* (3), *BMC bioinformatics* (2), *Biodiversity Data Journal* (1), *Frontiers in psychology* (1), *Comparative cytogenetics* (1).
Próprio *data paper*	2	*Data in brief* (107), *Chemical data collections* (8)

Fonte: elaborado pelas autorias. Notas: *F (frequência): refere-se a quantidade de periódicos

**Quantidade de *data papers* publicados pelos periódicos, conforme tipo de disponibilidade

Dito de outra forma, em pouco mais da metade dos *data papers* publicados, os repositórios foram a opção dos autores filiados às instituições brasileiras para disponibilizar os dados de pesquisa. No caso dos artigos científicos publicados na Plos One, Federer *et al.* (2018) havia identificado que o uso dos repositórios está longe do método dominante de partilha de dados, correspondendo a cerca de 20% dos artigos que foram analisados.

Esta variedade de disponibilização dos conjuntos de dados de pesquisa levou Candela *et al.*, (2015) a afirmar que, em geral, os *data papers* são os principais objetos de preocupação dos periódicos, tornando os conjuntos de dados objetos secundários. Jiao, Li e Fang (2022) corroboram com essa afirmação destacando que esses diferentes mecanismos de compartilhamento, impedem que os dados cumpram seus objetivos na comunicação acadêmica, além de prejudicar sua preservação.

A disponibilização dos dados como parte integrante dos *data papers* e a não disponibilização dos dados pesquisa foram aspectos observados em 14 dos 38 periódicos identificados, embora essa

prática esteja regredindo ao longo dos anos, conforme figura 1. Com isso, pode-se dizer que, em 26,66% (169) dos *data papers* os conjuntos de dados conjuntos não foram disponibilizados de fato, uma vez que esse tipo de comportamento não permite o acesso à totalidade dos dados de pesquisa.

Nesses casos, os interessados acabam não tendo acesso aos conjuntos de dados, exceto se entrarem em contato direto com os pesquisadores e se forem atendidos, o que acaba se tornando uma barreira para a verificabilidade e o reuso dos dados de pesquisa. Federer (2022) afirma que as diferenças disciplinares podem contribuir para variações nas taxas de disponibilidade de dados, o que pode explicar os resultados do presente estudo.

Na pesquisa de Jiao, Li e Fang (2022) os resultados revelaram que os métodos dominantes adotados na Plos One, os conjuntos de dados são compartilhados dentro do artigo e/ou nos materiais suplementares, seguidos pela categoria repositório. Ao comparar o trabalho desses autores com esta pesquisa, os resultados indicam uma tendência crescente na adesão do uso de repositórios de dados de pesquisa por parte dos pesquisadores, embora seja importante ressaltar que a tipologia documental analisada, os *data papers*, pode ter influência direta nos resultados.

4.1 Compartilhamento de dados em repositórios segundo as orientações dos periódicos

Com o objetivo de identificar se as orientações descritas pelos periódicos em relação a tipologia de repositório de dados foram seguidas, foram analisados os 638 conjuntos de dados, vinculados a 349 *data papers*, cujos dados foram disponibilizados apenas em repositórios de dados, conforme Tabela 2.

Tabela 2 - Disponibilidade dos conjuntos de dados segundo a tipologia dos repositórios

continua...

Periódico	Orientação do periódico	Data paper	Total de data sets			Repositório			
						Multidisciplinar		Disciplinar	
	Tipo de orientação	F*	F	%	MD**	F	%	F	%
Data in brief	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	164	204	31,97	1,24	139	21,79	65	10,19
Scientific data	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	82	249	39,03	3,04	117	18,34	132	20,69
Earth system science data	Repositórios temáticos	34	45	7,05	1,32	25	3,92	20	3,13
Data	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	16	20	3,13	1,25	19	2,98	1	0,16
Biodiversity Data Journal	Repositórios temáticos	15	19	2,98	1,27	3	0,47	16	2,51
BMC research notes	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	5	8	1,25	1,60	4	0,63	4	0,63
Chemical data collections	Recomenda o Mendeley Data	5	8	1,25	1,60	1	0,16	7	1,10
Gigascience	Recomenda o GigaDB	5	24	3,76	4,80	-	-	24	3,76
Geoscience data journal	Repositórios indexados no R3Data ou que estejam numa lista específica do periódico	3	21	3,29	7	2	0,31	19	2,98
Frontiers in ecology and evolution	Recomenda o Figshare	2	12	1,88	6	-	-	12	1,88
Frontiers in neuroscience	Recomenda o Figshare	2	2	0,31	1	2	0,31	-	-

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional. (cc) BY

Microbiology resource announcements	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	2	9	1,41	4,50	-	-	9	1,41
Annals of forest science	Repositórios indexados no R3Data	1	1	0,16	1	1	0,16	-	-
Big earth data	Repositórios indexados no R3Data	1	1	0,16	1	-	-	1	0,16
Ecology	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	1	1	0,16	1	1	0,16	-	-
Ecosistemas	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	1	1	0,16	1	1	0,16	-	-
Frontiers in genetics	Recomenda o Figshare	1	2	0,31	2	1	0,16	1	0,16
Frontiers in marine science	Recomenda o Figshare	1	1	0,16	1	1	0,16	-	-
Frontiers in psychology	Recomenda o Figshare	1	1	0,16	1	1	0,16	-	-
Frontiers in public health	Recomenda o Figshare	1	1	0,16	1	-	-	1	0,16
Global ecology and biogeography	Recomenda o Dryad	1	1	0,16	1	-	-	1	0,16
Journal of open humanities data	Possui uma lista própria de repositórios	1	1	0,16	1	1	0,16	-	-
Marine genomics	Preferencialmente em repositórios específicos	1	1	0,16	1	-	-	1	0,16
Proteomics	Repositórios que fazem parte do consórcio ProteomeXchange	1	3	0,47	3	-	-	3	0,47
Scientific reports	Sempre que possível em repositórios específicos, caso não haja usar multidisciplinar	1	1	0,16	1	1	0,16	-	-
ZooKeys	Recomenda o Dryad	1	1	0,16	1	-	-	1	0,16
Totais		**349**	**638**	**100**	**1,83**	**320**	**50,16**	**318**	**49,84**

Fonte: elaborado pelos autores

Notas: F data paper* (data papers depositado em repositório): quantidade de *data papers* cujos dados foram depositados em repositórios de dados

MD**(média): média de conjuntos de dados (*data sets*) por *data paper*.

Dos 26 periódicos identificados, 10 deles sugerem o depósito em **repositórios multidisciplinares** específicos, sendo que: seis sugerem o *Fighshare*, dois sugerem o *Dryad*, um recomenda o *GigaScience Data Base* e um sugere o *Mendeley Data*.

A recomendação pelo uso de **repositórios disciplinares**, temáticos ou específicos foi realizada por três periódicos, o *Earth System Science,* o *Biodiversity Data Journal,* e o *Marine genomics,* no entanto, esses periódicos não especificam qual ou quais os repositórios disciplinares seriam ideais para o depósito dos dados de pesquisa. O periódico *Proteomics*, restringiu a seleção aos repositórios que fazem parte do consórcio *ProteomeXchange*[5].

Os demais periódicos (12), deixam a seleção do repositório a critério dos autores, uma vez que apenas recomendam que o depósito "Sempre que possível seja realizado em repositórios específicos, caso não haja usar multidisciplinar" ou ainda que o repositório esteja indexado no R3Data ou lista específica definida pelo periódico.

Esses resultados indicam que não há uma padronização na recomendação clara dos periódicos em utilizar repositórios disciplinares, pois apenas dois periódicos realizaram essa recomendação.

[5] O ProteomeXchange uma colaboração internacional cujo objetivo é compartilhar conjuntos de dados de proteínas. Este foi um esforço internacional criado para evitar duplicação e facilitar o acesso e a reanálise de dados proteômicos (Stroe, 2024).

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional.

Ao realizar a comparação entre a sugestão do periódico e a efetivação do depósito dos dados de pesquisa, segundo a tipologia do repositório foi possível verificar que:

- nos **seis** periódicos que indicaram o **repositório multidisciplinar *Figshare***, em três títulos, os dados foram de fato depositados em repositórios multidisciplinares, em dois a opção foi pelo repositório disciplinar e um depositou os conjuntos nas duas tipologias de repositórios. Dessa maneira os autores seguiram parcialmente a orientação do periódico.
- nos periódicos que sugeriram **repositórios específicos**, **dois** periódicos que sugeriram o *Dryad* e um indicou ou *GigaScience Data Base*, os autores não seguiram a recomendação do periódico, compartilharam os dados em repositórios disciplinares e
- o único que indicou o *Mendeley Data*: publicou das duas formas, o que permite inferir também que os autores seguiram parcialmente a recomendação dos periódicos.

Com isso, dos 10 periódicos que **especificamente** indicaram o uso de **repositórios multidisciplinares**, em apenas três, os autores o fizeram especificamente, em quatro publicaram em repositórios disciplinares e três publicaram nas duas tipologias de repositórios.

Dos quatro periódicos que indicaram o uso de **repositórios disciplinares**, apenas os *data papers* publicados em dois dos títulos seguiram essa orientação e os demais publicaram em ambos os repositórios, multidisciplinares e disciplinares.

Nos demais periódicos (12), onde a indicação pela **tipologia do repositório não foi estabelecida**, em cinco periódicos os conjuntos de dados foram depositados em repositórios multidisciplinares, em dois periódicos o depósito ocorreu em repositórios disciplinares e, em cinco periódicos o depósito ocorreu nos dois tipos de repositórios.

Portanto, o uso apenas do repositório multidisciplinar ocorreu em oito periódicos, o mesmo se repetiu em repositórios disciplinares (8) e em 10 periódicos o depósito foi realizado em ambas às tipologias de repositórios. De maneira geral, esses resultados sugerem, que a falta de clareza nas orientações quanto aos tipos de repositórios a serem utilizados, têm levado os autores a selecionarem os repositórios com base em critérios alheios às recomendações dos periódicos.

Ao se realizar uma análise a partir da quantidade de conjuntos de dados que foram depositados nos tipos de repositórios indicados pelos periódicos, percebeu-se não houve uma diferença significativa entre o número de conjuntos de dados depositados em repositórios disciplinares e multidisciplinares, correspondendo a apenas dois conjuntos de dados. Com isso, pode-se inferir que novamente na prática que apesar das orientações dos periódicos, a escolha do tipo de repositório pode depender mais de fatores individuais dos autores ou da natureza dos dados do que das diretrizes dos periódicos.

4.2 A influência da área no compartilhamento dos dados

A média geral foi de 1,83 conjuntos de dados por *data paper*. As cinco maiores médias foram dos periódicos *Geoscience data journal* (7), *Frontiers in ecology and evolution* (6), *Gigascience* (4,80), *Microbiology resource announcements* (4,50), *Scientific data* (3,04), juntos esses periódicos correspondem a 61,64% (196) do

total de conjuntos de dados depositados em repositórios disciplinares (318). O fato dos repositórios disciplinares possuírem uma maior média de conjuntos por *data paper* pode ser influenciado pelo comportamento das áreas do conhecimento, embora neste estudo a análise tenha se concentrado apenas na tipologia do repositório e não no repositório especificamente.

Nesse sentido, Jiao, Li e Fang (2022) ao analisarem a relação entre os repositórios disciplinares e as áreas do conhecimento, verificaram que existem culturas de compartilhamento de dados diferentes entre os domínios de conhecimento, o que pode ser explicado por suas relações com dados e pelo paradigma de pesquisa orientado pela área.

Portanto, os resultados deste trabalho indicam que simplesmente exigir que os dados sejam compartilhados de alguma forma pode não ter o impacto desejado para tornar os dados de pesquisa FAIR, sendo assim, garantir que os dados estejam realmente disponíveis e permaneçam assim a longo prazo, pode exigir requisitos e especificações adicionais sobre como e onde os dados devam ser compartilhados (Federer, 2022). Além do estabelecimento de políticas e/ou critérios de revisão criteriosos para publicação dos dados de pesquisa associados aos *data papers*.

5 CONSIDERAÇÕES FINAIS

Esse trabalho se propôs verificar se a disponibilidade dos conjuntos de dados de pesquisa associados aos *data papers* de autores filiados às instituições brasileiras está condizente com as orientações dos periódicos, a partir da identificação de como são publicados os conjuntos de dados e da discussão se a publicação

segue as diretrizes do periódico no que condiz aos repositórios de dados.

Todos os periódicos aos quais os conjuntos de dados foram depositados em repositórios possuem orientações voltadas para o depósito dos dados de pesquisa. No entanto, o que se observou na prática é que, embora muitos periódicos exijam a disponibilidade de dados em repositórios, apenas a metade dos *data papers* o fez dessa maneira. A existência de múltiplas formas de disponibilização indica falta de padronização nas orientações. Tais resultados apontam para a necessidade de maior rigor nas políticas editoriais e no fortalecimento da cultura de compartilhamento de dados entre os pesquisadores.

O estudo revelou que a publicação de *data papers* por autores filiado às instituições brasileiras ainda carece de maior aderência às políticas editoriais relacionadas ao compartilhamento de dados, particularmente no que tange à escolha adequada dos repositórios. Incorporar a verificação da existência, acessibilidade e qualidade dos conjuntos de dados como parte do *peer review* se mostra essencial para fortalecer o ecossistema da publicação de dados.

Para estudos futuros recomenda-se analisar os repositórios especificamente e relacioná-los às temáticas abordadas nos *data papers*. Além de realizar uma busca retroativa nas políticas editoriais dos periódicos para verificar o quanto as orientações referentes às disponibilidades dos dados influenciaram o compartilhamento ao longo do tempo.

REFERÊNCIAS

CANDELA, Leonardo. *et al*. Data journals: a survey. **Journal of the Association for Information Science and Technology**, [*s.l.*], v. 66, n. 9, p. 1747-1762, 2015. Disponível em: https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/asi.23358. Acesso em: 8 abr. 2025.

FEDERER, Lisa. *et al*. Data sharing in Plos One: an analysis of data availability statements. **PLOS ONE**, [*s.l.*], v. 13, n. 5, p. 1-12, 2018. Disponível em: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194768. Acesso em: 8 abr. 2025.

FEDERER, Lisa. Long-term availability of data associated with articles in Plos One. **PLOS ONE**, [*s.l.*], v. 17, n. 8, p. e0272845, 2022. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC9401135/. Acesso em: 30 abr. 2025.

JIAO, Chenyue; LI, Kai; FANG, Zhichao. Data sharing practices across knowledge domains: a dynamic examination of data availability statements in Plos One publications. **Journal of Information Science**, [*s.l.*], v. 50, n. 3, p. 673-689, 2022. Disponível em: https://doi-org.ez46.periodicos.capes.gov.br/10.1177/0165551521101830. Acesso em: 14 abr. 2025.

LI, Kai; HUANG, Pao-Pei; JENG, Wei. Are data papers cited as research data? Preliminary analysis on interdisciplinary data paper citations. **Information Research**, [*s.l.*], v. 30, n. iConf, p. 1225-1233, 2025. Disponível em: https://doi.org/10.47989/ir30iConf46918. Acesso em 8 abr. 2025.

PUERTA-PIÑERO, Carolina; PÉREZ-LUQUE, Antonio Jesus; RODRÍGUEZ-ECHEVERRÍA, Susana. Ecosistemas is committed to the publication of data papers. **Ecosistemas**, [*s.l.*], v. 29, n. 3, p.1-2, 2020. Disponível em: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85099622212&doi=10.7818%2FECOS.2118&partnerID=40&md5=fe10a336677df22873eebbe41df5951f. Acesso em: 16 abr. 2025.

STROE, Oana. Proteomics: a different lens for precision medicine. **EMBL**, [*s.l.*], 2024. Disponível em:

https://www.embl.org/news/science-technology/proteomics-a-different-lens-for-precision-medicine/. Acesso em: 03 maio 2025.

WILKINSON, Mark. *et al.* The FAIR Guiding principles for scientific data management and stewardship. **Scientific Data**, [*s.l.*], v. 3, 160018, 2016. Disponível em: https://www.nature.com/articles/sdata201618. Acesso em: 16 abr. 2025.

LICENÇA DE USO

Direitos autorais das pessoas autoras, 2026. Licenciado sob Licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).

PUBLISHER

Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EQUIPE EDITORIAL

Martha Suzana Cabral Nunes, Maria Cleide Rodrigues Bernardino, Franciéle Carneiro Garcês da Silva.

COMO CITAR

SILVEIRA, Crislaine; DIAS, Thiago Magela Rodrigues; NEUBERT, Patrícia. Disponibilidade de dados de pesquisa: uma análise dos data papers publicados por autores filiados às instituições brasileiras. **Tendências da Pesquisa Brasileira e Ciência da Informação**, São Paulo, v. 19, p. 1-23, jan./jun. 2026.