



GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

ISSN 2177-3688

ORIENTAÇÕES PARA A FORMATAÇÃO DE ARTIGOS DE DADOS: PRIMEIROS PASSOS PARA A CONSTRUÇÃO DE UM PADRÃO ADAPTÁVEL

GUIDELINES FOR FORMATING DATA PAPERS: FIRST STEPS TOWARDS BUILDING AN ADAPTABLE STANDARD

Lyvia Rocha de Jesus Araujo - Instituto Brasileiro de Informação em Ciência e Tecnologia - Universidade Federal do Rio de Janeiro (IBICT/UNIRIO)

Maria Simone de Menezes Alencar - Universidade Federal do Rio de Janeiro (UNIRIO)

Modalidade: Trabalho Completo

Resumo: Esta pesquisa compara as características de templates de artigos de dados fornecidos por revistas de dados indexadas nas bases de dados Scopus e Web of Science, com o objetivo de elucidar os avanços encontrados na padronização das informações essenciais fornecidas por este tipo de publicação sobre os conjuntos de dados de pesquisa que descrevem. Foi realizada uma análise do conteúdo de 18 templates de artigos de dados extraídos de revistas de dados indexados na Scopus e na Web of Science. Foram utilizadas as seguintes categorias para descrição dos elementos de artigos de dados: propriedades gerais do conjunto de dados, descrição da produção/coleta dos dados e descrição sobre compartilhamento e reuso dos dados. Verificou-se que, em média, metade dos elementos tidos como ideais na composição de artigos de dados vêm sendo implementados nos templates de revistas de dados internacionais analisadas. Conclui-se que a investigação que a melhor compreensão da visão de editores científicos quanto à elaboração dos modelos de artigos de dados adotados nas revistas em que atuam pode vir a esclarecer as informações que consideram relevantes para a descrição, interpretação e reuso de conjuntos de dados compartilhados. Apesar disso, entende-se que os primeiros passos para o surgimento de um padrão de campos e recomendações para a publicação de artigos de dados já vêm sendo fundamentados.

Palavras-chave: artigos de dados; periódicos de dados; dados abertos de pesquisa; comunicação científica.

Abstract: This study compares the characteristics of data article templates provided by data journals indexed in the Scopus and Web of Science databases, with the aim of elucidating the progress made to standardize the essential information provided by this type of publication about the research data sets they describe. A content analysis of 18 data article templates extracted from data journals indexed in Scopus and Web of Science was performed. The following categories were used to describe the data article elements: General properties of the dataset, Description of the production/collection of the data, Description of the sharing and reuse of the data. It was found that, on average, half of the elements considered ideal in the composition of data articles are implemented in the templates of international data journals. It is concluded that research to better understand the views of scientific editors on the development of data article models adopted in the journals in which they work may shed light on the information they consider relevant for the description, interpretation and reuse of shared datasets. Nevertheless, it is clear that the first steps towards the emergence of a standard of fields and recommendations for the publication of data articles are already being established.

Keywords: data papers; data journals; ppen research data; scientific communication.

1 INTRODUÇÃO

A Ciência Aberta fez com que os atores envolvidos no processo de comunicação científica, como pesquisadores, editores, agências de fomento, começassem a repensar o modo de se conduzir a ciência. De modo que a construção do conhecimento não tivesse foco apenas nas publicações dos resultados das investigações conduzidas, mas também fosse orientada às suas ferramentas, recursos e dados de pesquisa. Assim, tendo em consideração que estes elementos podem ser disponibilizados sem restrições de acesso, além de produzidos, avaliados, publicados e preservados de maneira aberta, tornando possível seu uso e reaproveitamento dentro dos direitos e deveres de autoria e divulgação.

Compreende-se como dado de pesquisa “[...] todo e qualquer tipo de registro coletado, observado, gerado ou usado pela pesquisa científica, tratado e aceito como necessário para validar os resultados da pesquisa pela comunidade científica.” (SALES; SAYÃO, 2019, p. 36). Dentre os objetivos que atualmente norteiam a publicação de dados, os principais são o de torná-los encontráveis, acessíveis, reproduzíveis e reutilizáveis, seguindo os princípios FAIR (*Findable, Accessible, Interoperable, and Reusable*) para a gestão de dados de pesquisa, publicados em 2016 na revista *Scientific Data*, por um conjunto diversificado de representantes do meio acadêmico-científico, editoras científicas, e afins.

Os princípios FAIR têm como base fundamental a capacidade dos sistemas digitais de operar dados que sejam encontráveis, acessíveis, interoperáveis e reusáveis (GO FAIR, [2016]). Para Kim (2020) estes objetivos são cumpridos pelo compartilhamento de conjuntos de dados em repositórios e a elaboração de *data papers* (artigos de dados) que fornecem documentação extensa abordando informações contextuais suficientes para a interpretação dos conjuntos de dados compartilhados por pesquisadores.

Callaghan *et al.* (2012), explicam que um *data paper* fornece informações sobre o quê, onde, por quê, como e de quem são os dados de pesquisa compartilhados, além de garantir o acesso a estes materiais através de um identificador permanente (*Digital Object Identifier - DOI*) que direcione o pesquisador ao repositório em que os dados de pesquisa que procuram estão armazenados. Os artigos de dados ou *data papers*, são orientados para apresentação dos metadados que descrevem a estrutura, o formato, os métodos utilizados na coleta, organização e processamento dos dados.

Os artigos de dados são publicados em revistas de dados (*data journals*) que aceitem sua submissão, publicações ampliadas ou revistas de dados puras (que só aceitam a submissão de artigos de dados). As revistas de dados surgem a partir da discussão sobre os benefícios do compartilhamento e o reuso de dados de pesquisa e as barreiras metodológicas, legais e técnicas que fazem com que pesquisadores não se sintam muito inclinados a publicar seus dados. (CANDELA *et al.*, 2015).

Chavan e Penev (2011) discorrem sobre a função das revistas de dados como publicações de documentos de metadados pesquisáveis, que descrevem conjuntos de dados disponibilizados em formato digital dentro dos padrões de ações de divulgação científica já existentes. Essa função se soma às características comuns aos periódicos científicos convencionais, como a revisão por pares e a possibilidade de abranger conteúdos de uma ou mais áreas do conhecimento, mas os difere destes mesmos periódicos, na medida em que não compartilham da intenção de propagar os resultados de pesquisas acadêmicas e estudos científicos, limitando-se à descrição detalhada dos conjuntos de dados e de sua localização.

Uma revista de dados permite que os conjuntos de dados de pesquisa produzidos sejam compartilhados como uma “produção de pesquisa de primeira classe” publicada, revisada por pares, adequadamente indexada nas principais bases bibliográficas, que os tornem citáveis, facilmente detectáveis e reutilizáveis. Para fins desta pesquisa, considera-se como revistas de dados os periódicos mistos, que publicam qualquer tipo de artigo, incluindo artigos de dados, bem como periódicos de dados puros, que publicam somente artigos de dados (CANDELA *et al.*, 2015; CHEN, 2017).

Tanto as revistas quanto os próprios artigos de dados podem ser indexados e recuperados nas bases de dados tradicionais utilizadas para a disponibilização de periódicos convencionais de assuntos gerais ou especializados. O que os diferencia é a intenção que as revistas e artigos de dados têm de tornar um conjunto de dados reutilizável e reproduzível, mediante descrição aprofundada que permita sua exploração livre, garantindo os direitos de quem o tenha produzido (CANDELA *et al.*, 2015).

De acordo com Kim (2020), a maioria dos *data journals* “puros” foi lançada nos últimos quinze anos, a partir de 2008. No Brasil o número de revistas que informam aceitar a publicação de artigos de dados ainda é muito reduzido. Até o momento poucos estudos foram realizados em âmbito nacional discutindo este assunto e apenas nos últimos dois anos

observou-se a inclusão de artigos de dados no escopo de submissão de revistas científicas brasileiras (ARAUJO; PRINCIPE; ALENCAR, 2021).

Apresentado o contexto acima, e guiada pela seguinte questão: “Quais elementos essenciais dos modelos de artigos de dados atuais podem ser considerados padrão para este formato de publicação?”, a presente pesquisa tem o objetivo de comparar as características dos *templates* de artigos de dados fornecidos pelas revistas de dados indexadas nas bases de dados *Scopus* e *Web of Science*, com vista à elucidar os avanços encontrados na padronização das informações essenciais fornecidas por este tipo de publicação sobre os conjuntos de dados de pesquisa que descrevem.

2 PROCEDIMENTOS METODOLÓGICOS

Foram coletados os títulos das revistas indexadas nas bases *Scopus* e *Web of Science* que publicam artigos de dados (revistas de dados). Em seguida, sucedeu-se a exploração das instruções aos autores, ou sessão correspondente de cada revista, tendo em vista a seleção daquelas que, até o momento desta pesquisa, forneciam um *template* para a elaboração de artigos de dados em suas diretrizes.

A busca foi conduzida em fevereiro de 2023, nas bases *Scopus* e *Web of Science*. A escolha destas bases se deu devido à ampla abrangência de cobertura em diferentes áreas do conhecimento e a possibilidade de recuperação direta de artigos de dados pelo filtro *data papers* existente em ambas as bases.

A recuperação dos dados na *Scopus* se deu através do uso da query “*DOCTYPE (dp) AND PUBYEAR <2023*” no campo de pesquisa avançada da base, restringindo a busca a artigos de dados publicados antes de 2023. Os registros dos artigos recuperados na *Scopus* foram exportados em formato CSV para organização posterior.

Na *Web of Science* a busca foi realizada com o uso da query “*DT=(Data Paper) AND PY=(1900-2022)*”, restringindo-se à recuperação de artigos de dados publicados desde o primeiro ano de indexação das fontes da base até o fim do ano de 2022. A partir dos resultados obtidos nesta busca foram exportados os títulos das revistas nas quais os artigos foram publicados, estes títulos foram organizados em uma planilha única no *Google Sheets* junto aos títulos das revistas encontrados na *Scopus*, para deduplicação.

A exploração das orientações aos autores das revistas listadas foi realizada manualmente. Foram considerados para fins de análise desta pesquisa somente os periódicos

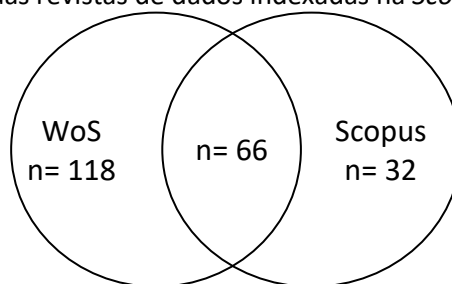
que disponibilizavam, até a data da consulta, um *template* específico para a elaboração de artigos de dados. Foram desconsideradas nesta pesquisa, revistas que não fornecem um *template* próprio ou somente direcionam para exemplos de artigos já publicados.

Foi realizada uma análise do conteúdo dos *templates* para elaboração de artigos de dados fornecidos pelas revistas selecionadas, considerando as categorias desenvolvidas por Kim (2020) para a descrição dos elementos recomendados para os artigos de dados, a saber: propriedades gerais do conjunto de dados, contextualização da produção/coleta dos dados, detalhes sobre compartilhamento e reuso dos dados.

3 RESULTADOS E DISCUSSÃO

Foram recuperados 282 títulos de revistas científicas, 98 delas indexadas na *Scopus* e 184 na *Web of Science*, 66 títulos foram excluídos por serem duplicatas, estando indexados em ambas as bases de dados consultadas. Foram listados 216 títulos de revistas de dados ao fim da deduplicação.

Figura 2 – Distribuição das revistas de dados indexadas na *Scopus* e na *Web of Science*



Fonte: Dados da pesquisa (2023).

Uma vez listadas todas as revistas encontradas, foram analisadas manualmente as sessões de orientações aos autores ou sessões correspondentes de cada uma delas. As revistas investigadas nesta pesquisa são mantidas por 11 (onze) editoras ou instituições científicas (Quadro 1).

Quadro 1 – Títulos, editoras e codificação das revistas incluídas

Título da revista	Editora	Código
Annals Of Forest Science	Springer	R1
Big Earth Data	Taylor & Francis	R2
Bmc Genomic Data	BioMed Central	R3
Bmc Research notes	BioMed Central	R4
Chemical Data Collections	Elsevier	R5
Computer Networks	Elsevier	R6

XXIII Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB
Aracaju-SE – 06 a 10 de novembro de 2023

Título da revista	Editora	Código
Data	Multidisciplinary Digital Publishing Institute (MDPI)	R7
Data In Brief	Elsevier	R8
Ecology	Wiley	R9
Epidemiology And Health	Korean Society of Epidemiology	R10
F1000research	Taylor & Francis	R11
Geoscience Data Journal	Wiley	R12
Geus Bulletin	Geological Survey of Denmark and Greenland (GEUS)	R13
Grey Journal	GreyNet	R14
Limnology And Oceanography Letters	Wiley	R15
Open Quaternary	Ubiquity Press	R16
Scientific data	Nature	R17
Tribology Letters	Springer	R18

Fonte: Dados da pesquisa (2023).

Esta etapa resultou em um total de 18 *templates* para a elaboração de *data papers*, *data articles*, *data notes*, *dataset articles*, *data descriptors*, *data profiles*, de acordo com a nomenclatura adotada por cada revista para se referir aos artigos de dados (Tabela 1).

Tabela 1 – Terminologia adotada pelas revistas sobre artigos de dados

Termo adotado	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18
<i>Data article</i>		X			X			X				X	X		X			
<i>Data descriptor</i>							X										X	
<i>Data note</i>			X	X							X							X
<i>Data paper</i>	X								X					X		X		
<i>Data profiles</i>										X								
<i>Dataset article</i>						X												

Fonte: Dados da pesquisa (2023).

Os dados apresentados nas tabelas 2, 3, 4 e 5 foram analisados de acordo com os 20 elementos explorados por Kim (2020) para a elaboração de artigos de dados.

Para fins desta pesquisa, ao tratar sobre o formato dos dados, levou-se em consideração o formato pelo qual os dados foram compartilhados e podem ser exportados, como: CSV (*Comma-separated values*), XLS (*Excel*), PDF (*Portable Document Format*), RTF (*Rich Text Format*), etc. Por natureza dos dados entende-se a tipologia dos dados, se estes são textuais, figuras, vídeos, entre outros. Por fim, a versão dos dados trata sobre as etapas do processamento dos dados, se o conjunto de dados contém os dados brutos, tratados, analisados, etc.

Nessa conjectura, 10 *templates* recomendam o preenchimento do formato dos dados compartilhados, sete (7) *templates* recomendam a especificação da natureza dos dados e nove (9) *templates* solicitam a descrição da versão dos dados. Outro dos elementos mais recorrentes identificados nesta categoria foi a solicitação da adoção de um nome específico para cada arquivo de dados do conjunto, observado em oito (8) *templates* (Tabela 2).

Dentre os elementos menos recorrentes nos *templates* consultados, se encontram a apresentação da linguagem ou idioma dos dados (5), do tamanho do arquivo de dados disponibilizado (1) e a data de criação dos conjuntos de dados (5). Este último podendo impactar na reprodutibilidade dos métodos implementados para coleta/produção dos dados, uma vez que as datas de captura de dados auxiliam na identificação de mudanças ocorridas no contexto de produção dos dados (CANDELA *et al.*, 2015).

Tabela 2 – Descrição das propriedades gerais dos conjuntos de dados segundo os *templates*

Propriedades gerais do conjunto de dados	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18
Data de criação/coleta dos dados	X									X			X	X		X		
Formato dos dados		X	X	X			X		X			X	X	X	X	X		
Natureza dos dados	X		X	X	X	X		X										X
Versão dos dados	X				X	X		X	X			X	X			X		X
Nome do arquivo	X		X	X			X		X			X		X		X		
Tamanho do arquivo									X									
Linguagem	X					X		X						X		X		

Fonte: Dados da pesquisa (2023).

No que tange ao panorama da produção e coleta dos dados expostos nos artigos de dados publicados, conforme a Tabela 3, a maioria dos *templates* sugere pelo menos a explicação dos métodos de coleta/criação dos conjuntos de dados compartilhados (14), bem como dos métodos adotados na análise/leitura dos dados para a pesquisa originária dos mesmos (15) e sua conexão com um ou mais projetos de desenvolvimento de pesquisa, financiados ou não, quando esta existir (13), os objetivos que nortearam a aquisição dos dados (13), e as limitações encontradas em sua organização e disposição (12).

Pouco menos presente nos modelos observados foi a orientação direcionada à explanação da cobertura dos dados, considerando limites temporais, de espécie e geográficos

(10), e a especificação do(s) responsável(eis) pela produção do(s) conjunto de dados publicados (9)¹.

Tabela 3 – Contextualização da produção/coleta dos dados segundo os *templates*

Produção e coleta dos dados	R1	R2	R3	R4	R5	R6	R7	R8	R9	R 10	R 11	R 12	R 13	R 14	R 15	R 16	R 17	R 18
Coleta/criação dos dados	X					X	X	X	X	X	X	X	X	X	X	X	X	X
Espécimes e artefatos (Cobertura)	X		X	X			X		X	X			X	X	X	X		
Produtor dos dados	X						X	X	X			X	X	X	X	X		
Análise dos dados	X	X	X	X	X	X	X	X	X	X	X		X		X	X	X	
Limitações	X		X	X	X				X	X	X	X		X		X	X	X
Objetivos da pesquisa	X	X	X	X				X	X	X	X	X	X		X		X	X
Projeto (financiamento)	X	X	X	X		X	X	X	X	X			X			X	X	X

Fonte: Dados da pesquisa (2023).

No conjunto das categorias relacionadas ao compartilhamento e reuso dos dados, os elementos de destaque foram a discriminação da proveniência ou localização dos dados para consulta (15), do reuso primário dos dados (12) - que compreende a descrição de trabalhos e projetos nos quais o(s) conjunto(s) de dados já tenham sido utilizados até o momento da publicação do artigo de dados, e do seu reuso potencial (13), e da definição dos termos ou licença de uso dos dados (16).

Somente um dos *templates* investigados apontou como sugestão o registro de atualização/curadoria dos dados após serem compartilhados/publicados em repositório. Por fim, nenhum dos modelos consultados solicita qualquer descrição específica sobre a reputação ou histórico do(s) repositório(s) utilizado(s).

Tabela 4 – Detalhes sobre compartilhamento e reuso dos dados segundo os *templates*

Compartilhamento e reuso dos dados	R1	R2	R3	R4	R5	R6	R7	R8	R9	R 10	R 11	R 12	R 13	R 14	R 15	R 16	R 17	R 18
Sobre repositórios																		
Proveniência	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X	
Reputação e histórico																		
Curadoria		X																
Sobre reuso																		
Reuso primário	X		X	X	X	X	X	X		X		X	X		X	X		X
Reuso potencial	X	X				X		X	X	X		X	X	X	X	X	X	X
Termos de uso (licença/conflitos de interesse)	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X

Fonte: Dados da pesquisa (2023).

¹ Para a contabilização do requerimento desta última informação nos *templates* consultados, foi levada em consideração a solicitação da especificidade dos papéis dos autores dos artigos de dados e o uso da taxonomia CrEdit.

Na maior parte dos casos, verifica-se que são preconizados pelas revistas a descrição dos termos ou licença de uso dos conjuntos de dados relacionados aos artigos de dados publicados, a proveniência ou localização dos *datasets*, quando podem ser compartilhados, e a descrição dos métodos adotados na coleta/criação dos dados.

O único elemento não recomendado em nenhum dos *templates* selecionados foi a descrição da reputação e o histórico do repositório de dados onde os conjuntos de dados se encontram depositados. Fenômeno observado por Faniel *et al.* em 2019, que argumentaram que isso pode se dar devido ao fator subjetivo do mesmo, que implica na avaliação das aplicações de acordo com o julgamento dos seus usuários.

Candela *et al.* (2015) apontam como problemática o fato de que a explanação do reuso potencial dos dados de pesquisa é frequentemente negligenciada pelas diretrizes de publicação para artigos de dados, mas assim como observado por Kim (2020), esta pesquisa apontou que cerca de 30% dos modelos consultados não solicita explicitamente o preenchimento desta informação.

Outros pontos pouquíssimos abordados foram a curadoria dos dados e o tamanho aceitável dos arquivos de dados compartilhados - caso estes precisassem ser anexados ao artigo de dados publicado. O primeiro sendo o elemento para o qual foi considerado que atualizações periódicas dos conjuntos de dados poderiam ser realizadas e deveriam ser documentadas em diferentes artigos de dados e/ou nos metadados dos conjuntos de dados. Enquanto o segundo versava sobre as limitações de espaço para submissão de arquivos suplementares junto dos artigos de dados.

Salienta-se que foram incluídos nesta análise os *templates* de metadados suplementares das revistas *Annals Of Forest Science* e *Ecology*, disponibilizados junto aos seus modelos de artigos de dados.

Em estudo anterior foram identificados dois *templates* de revistas de dados brasileiras: 1) o modelo da revista “Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação”, publicação do Programa de Pós-graduação em Ciência da Informação da Universidade Federal de Santa Catarina (UFSC), ampliou seu escopo para o aceite da submissão de artigos de dados a partir de 2020; e 2) o modelo do *Latin American Data in Science Journal* (LADS), periódico de dados lançado em julho de 2021, de caráter multidisciplinar (ARAUJO; PRINCIPE; ALENCAR, 2021).

XXIII Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB
Aracaju-SE – 06 a 10 de novembro de 2023

Em comparação com os *templates* verificados nesta pesquisa, a Tabela 5 demonstra que os modelos disponibilizados pelas revistas brasileiras recomendam especificamente menos da metade dos 20 elementos divididos nas categorias organizadas por Kim (2020).

Tabela 5 – Templates de revistas de dados brasileiras e os elementos organizados por Kim (2020)

Elementos	Encontros Bibli	LADS
Propriedades gerais do conjunto de dados		
Data de criação/ coleta		
Formato		X
Natureza	X	
Versão	X	X
Nome do arquivo		
Tamanho		
Linguagem	X	
Produção e coleta dos dados		
Coleta/ criação	X	X
Cobertura	X	X
Produto		
Análise	X	
Limitações		
Objetivos	X	
Projeto		
Compartilhamento e reuso dos dados		
Proveniência	X	X
Reputação e histórico		
Curadoria		
Reuso primário	X	X
Reuso potencial	X	X
Termos de uso		

Fonte: Dados da pesquisa (2023).

Em média, os *templates* internacionais analisados nesta pesquisa recomendam 10,2% dos elementos essenciais para a construção de artigos de dados efetivos segundo Kim (2020). Destes, o que demonstrou maior conformidade com os elementos foi o modelo da revista *Annals Of Forest Science*, da editora *Springer*, contemplando 16 dos elementos recomendados. Enquanto isso, o modelo da revista *F1000research*, da *Taylor & Francis*, teve o menor registro de elementos recomendados, apenas cinco.

Em paralelo, as revistas brasileiras recomendam em média apenas 8,5% destes elementos. Acredita-se que a novidade da presença dos artigos de dados em periódicos nacionais, pode fazer com que orientações menos aprofundadas e campos menos extensos para preenchimento podem parecer mais convidativos a pesquisadores que tenham pouca ou nenhuma familiaridade com este modelo de publicação. No entanto, entende-se que talvez seja preciso encontrar maneiras de melhor averiguar a visão dos editores científicos ao

elaborar os modelos e guias de preenchimento de artigos de dados para melhor compreender as diferenças entre as prioridades e formatação das recomendações de cada um.

4 CONSIDERAÇÕES FINAIS

A presente pesquisa comparou as características dos *templates* de artigos de dados fornecidos pelas revistas de dados indexadas na *Scopus* e na *Web of Science*, com o propósito de elucidar os avanços encontrados na padronização dos elementos recomendados para a publicação de artigos de dados. O processo teve como principal referencial o trabalho de Kim (2020), com objetivo de identificar os componentes de um artigo de dados, cujos resultados facilitam a compreensão das melhorias necessárias às orientações oferecidas por revistas de dados para a produção destes artigos.

Dos 18 *templates* recuperados e explorados nesta pesquisa, somente cinco (5) estiveram presentes na pesquisa de Kim (2020): *Data*, *Data In Brief*, *Ecology*, *Geoscience Data Journal* e *Scientific data*. A análise atual destas revistas demonstrou diferenças sutis nos modelos disponibilizados para a elaboração de artigos de dados. Foi percebida a solicitação de novos elementos nos *templates* das revistas *Data*, *Data In Brief*, *Ecology* e *Scientific data*, indicando melhorias. Enquanto isso, na revista *Geoscience Data Journal* a ausência de alguns dos elementos identificados anteriormente por Kim (2020) se destacam.

Essas pequenas alterações parecem comprovar o avanço e adaptabilidade dos moldes fornecidos pelas revistas para a elaboração de artigos de dados nos últimos anos. O que talvez possa abrir espaço para indagações futuras a respeito das motivações de editores científicos ao planejarem tais mudanças, ou ainda à exploração da opinião dos autores quanto ao detalhamento das informações dos seus conjuntos de dados nestes modelos de publicação, e se de alguma maneira isso pode interferir na decisão dos editores científicos a respeito do tema.

Nesse sentido, Kim (2020) expressa que a principal motivação para a publicação de artigos de dados é fazer com que os dados de pesquisa se tornem facilmente reutilizáveis e reproduzíveis. Para tanto, estes artigos devem fornecer uma descrição extensa dos dados que ofereçam informações contextuais e aplicáveis o suficiente para o reaproveitamento dos dados fornecidos.

Um exame breve dos *templates* fornecidos pelas duas primeiras revistas de dados brasileiras demonstrou uma proximidade entre os elementos recomendados por ambas.

Apesar disso, ao serem comparadas com as revistas investigadas nesta pesquisa, notou-se que estas se encontram abaixo da média de elementos recomendados para a composição artigos de dados.

Conclui-se que a investigação que a melhor compreensão da visão de editores científicos quanto à elaboração dos modelos de artigos de dados adotados nas revistas em que atuam pode vir a esclarecer as informações que consideram relevantes para a descrição, interpretação e reuso de conjuntos de dados compartilhados. Apesar disso, entende-se que os primeiros passos para o surgimento de um padrão de campos e recomendações para a publicação de artigos de dados já vêm sendo fundamentados.

Por fim, a presente pesquisa sugere investigações mais apuradas e aprofundadas em contexto acadêmico-científico para melhor compreender as motivações que norteiam a construção das diretrizes para publicação de artigos de dados. Nesse cenário, ainda é possível indagar a grupos de pesquisadores que já tenham publicado artigos de dados sobre suas experiências com o processo pelo qual passaram, desde o compartilhamento dos dados em repositório até a submissão, revisão e publicação dos artigos.

Ao buscar-se observar a visão de editores científicos sobre os elementos que compõem os modelos desenvolvidos em suas revistas, e de pesquisadores-autores a respeito de sua vivência no processo de publicação de artigos de dados, espera-se facilitar a identificação de melhorias possíveis e padrões que possam ser efetivamente aplicáveis. De modo geral, entende-se que o processo de comunicação científica se torna eficaz na medida em que os objetivos dos meios de comunicação e o conteúdo disseminado por eles conseguem ser claros e úteis àqueles que os consomem.

REFERÊNCIAS

ARAUJO, Lyvia Rocha de Jesus; PRÍNCIPE, Eloisa; ALENCAR, Maria Simone de Menezes. Data papers brasileiros: análise da publicação de artigos de dados indexados na web of science e scopus. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 21., 2021, Rio de Janeiro. **Anais [...]**. Rio de Janeiro: IBICT/UFRJ, 2021. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxienancib/paper/viewFile/177/479>. Acesso em: 11 jun. 2023.

CALLAGHAN, Sarah. et al. Making data a first class scientific output: data citation and publication by NERC's environmental data centres. **Int J Digit Curation**, [s./l.], v. 7, n. 1, p. 107-113, 2012. Disponível em: <http://www.ijdc.net/article/view/208>. Acesso em: 13 jun. 2023.

CANDELA, Leonardo *et al.* Data journals: a survey. **Journal of the Association of Information Science and Technology**, [s.l.], v. 66, n. 9, p.1747-1762, 2015.

CHAVAN, Vishwas; PENEV, Lyubomir. The data paper: a mechanism to incentivize data publishing in biodiversity science. **BMC Bioinformatics**, [s.l.], v. 12, Supl. 15, 2011.

CHEN, Ya-Ning. An analysis of characteristics and structures embedded in data papers: a preliminary study. **Libellarium**. 2017.

FANIEL, Ixchel M. *et al.* Context from the Data Reuser's Point of View. **Journal of Documentation**, [s.l.], v. 75, n. 6, p. 1274–97, 2019.

GO FAIR. FAIR principles. **Go FAIR**, [2016]. Disponível em: <https://www.go-fair.org/fair-principles/>. Acesso em: 24 jun. 2023.

KIM, Jihyun. An analysis of data paper templates and guidelines: types of contextual information described by *data journals*. **Science Editing**, [s.l.], v. 7, n. 1, p. 16-23, 2020.

SALES, Luana Farias; SAYÃO, Luís Fernando. Uma proposta de taxonomia para dados de pesquisa. **Revista Conhecimento em Ação**, Rio de Janeiro, v. 4, n. 1, 2019.