

QUALIDADE DOS DADOS NA WEB: REVISÃO INTEGRATIVA SOBRE DIRETRIZES PARA PUBLICAÇÃO

Morgana Carneiro de Andrade
 <http://lattes.cnpq.br/4008145322343111> –  <https://orcid.org/0000-0002-4055-5882>
morganaandrade@hotmail.com
Universidade Federal do Espírito Santo (UFES)
Vitória, ES, Brasil

Maria José Baños Moreno
 <http://lattes.cnpq.br/4816805184682339> –  <https://orcid.org/0000-0001-9137-1330>
mbm41963@um.es
ODILO Plataforma Digital - Aprendizaje a Medida

Juan-Antonio Pastor-Sánchez
 <https://orcid.org/0000-0002-1677-1059>
pastor@um.es
Universidad de Murcia
Espanha

RESUMO

O aumento exponencial de dados publicados na Web e a diversidade de sistemas requerem adoção de boas práticas para alcançarem índices de qualidade que possibilitem a descoberta, o acesso e o reuso. Para identificar as boas práticas, utilizou-se a revisão integrativa, assim como procedimentos da metodologia ProKnow-C. Após a aplicação dos procedimentos do ProKnow-C aos documentos recuperados nas bases de dados Web of Science, Scopus e Library, Information Science & Technology Abstracts, foi realizada a análise de 31 itens. Essa análise permitiu observar que, nos últimos 20 anos, as diretrizes para publicação de dados governamentais abertos tiveram grande impacto na implementação do modelo Linked Data em diversos domínios e, atualmente, os princípios FAIR e as Data on the Web Best Practices são as mais destacadas na literatura. Essas diretrizes apresentam orientações em relação a vários aspectos para publicação de dados, de forma a contribuir para a otimização da qualidade, independente do contexto em que são aplicadas. Já os princípios CARE e FACT, que embora não tenham sido formulados com o mesmo objetivo do FAIR e das Best Practices, representam grandes desafios para os cientistas da informação e da tecnologia no que se refere a ética, responsabilidade, confidencialidade, imparcialidade, segurança e transparência dos dados.

Palavras-chave: Boas práticas. Publicação de dados na Web. Linked Open Data. Qualidade de dados.

DATA QUALITY ON THE WEB: INTEGRATIVE REVIEW OF PUBLICATION GUIDELINES

ABSTRACT

The exponential increase of published data and the diversity of systems require the adoption of good practices to achieve quality indexes that enable discovery, access, and reuse. To identify good practices, an integrative review was used, as well as procedures from the ProKnow-C methodology. After applying the ProKnow-C procedures to the documents retrieved from the Web of Science, Scopus and Library, Information Science & Technology Abstracts databases, an analysis of 31 items was performed. This analysis allowed observing that in the last 20 years the guidelines for publishing open government data had a great impact on the Linked Data model implementation in several domains and currently the FAIR principles and the Data on the Web Best Practices are the most highlighted in the literature. These guidelines presents orientations in relation to various aspects for the publication of data in order to contribute to the optimization of quality, independent of the context in which they are applied. The CARE and FACT principles, on the other hand, although they were not formulated with the same objective as FAIR and the Best Practices, represent great challenges for information and technology scientists regarding ethics, responsibility, confidentiality, impartiality, security, and transparency of data.

Keywords: Best practices. Data publishing on the Web. Linked Open Data. Data quality.

DOI <http://dx.doi.org/10.1590/1981-5344/40663>

Recebido em: 30 julho 2022.

Aceito em: 27 março 2023.



INTRODUÇÃO

O avanço das tecnologias de informação tem proporcionado importantes alterações em vários setores, principalmente no que se refere a geração, disponibilização, uso e reuso dos dados, a exemplo das plataformas de sistemas abertos (*Open Web Platforms*), das redes sociais e da Internet das Coisas (*Web of Things*).

No contexto da pesquisa científica, houve mudanças estruturantes na comunicação científica e, em relação ao governo eletrônico, registrou-se a ampliação da transparência e da governança participativa. Pode-se dizer que essas transformações tiveram como marco três episódios: a publicação do relatório da American Scientific Agency, em 1995, que menciona pela primeira vez o termo *open data* (dados abertos) (CHIGNARD, 2013); o posicionamento de ativistas, em 2007, que promoveram o acesso aos dados públicos do governo americano (CHIGNARD, 2013); e o aprimoramento da Web com o uso de novas ferramentas que proporcionaram avanços em relação aos dados ligados (*Linked Data*) (BIZER *et al.*, 2008; BIZER; HEATH; BERNERS-LEE, 2008, 2011).

A produção e a publicação de dados ligados são “processos de engenharia intensivos”, que demandam por uma alta qualidade dos dados. Há alguns anos existia certo receio de que as diretrizes existentes não fossem suficientes para gerar dados ligados ou até mesmo viabilizar a replicação dos processos (RADULOVIC *et al.*, 2015). Por outro lado, ainda não é possível afirmar se as comunidades em geral têm conhecimento sobre as boas práticas para publicação de dados mais adequadas às suas necessidades, apesar dessa preocupação estar presente desde o início da Internet (BORGMAN, 2012; KLUMP *et al.*, 2006).

A forma como os dados passaram a ser processados após o advento da Web mostra que esses são criados e compartilhados de forma desenfreada, o que requer uma gestão mais sofisticada (BATINI; SCANNAPIECA, 2006). A baixa aderência, a dificuldade em implementar boas práticas, assim como o treinamento deficiente e a forte pressão pela produção de dados são fatores que interferem diretamente na qualidade dos dados publicados, o que pode inviabilizar a sua localização e até mesmo o seu reuso.

O termo qualidade traz consigo alta carga de ambiguidade e, muitas vezes, pode ser interpretado de forma subjetiva, daí a necessidade de indicadores formalmente definidos. No que se refere a este estudo, é considerada a seguinte definição para qualidade: “[...] é o grau em que um conjunto de características inerentes atende aos requisitos” (ISO, 2020, tradução nossa). Já a definição adotada para o termo diretriz, neste estudo, corresponde a “[...] instrução ou orientação que deve ser seguida para levar a bom termo determinada tarefa” (DIRETRIZ, 2023). Neste estudo, esse termo é adotado para se referir a boas práticas, recomendações ou princípios.

É imprescindível ainda tornar claro outros conceitos, visto que esses podem ter significados distintos em determinadas comunidades/domínios. O termo dado, para cuja definição não há um consenso, neste trabalho é adotado como “[...] representações de observações, objectos, ou outras entidades utilizadas como evidências de fenômenos para fins de investigação ou de aprendizagem [...]” (BORGMAN, 2017, tradução nossa). Lovett *et al.* (2019) acrescentam ainda que dados como informação são registrados de várias formas, nomeadamente quando no que toca ao ambiente digital.

Jayawardene, Sadiq e Indulska (2015) trazem, em seu artigo, uma definição para dados, apresentada por Liebenau e Backhouse (1990), como sendo a “[...] representação de objetos ou fenômenos no mundo real”, e relacionam a qualidade desses dados com a satisfação das necessidades dos usuários de sistemas de informação. Para os autores, a má qualidade dos dados, à luz dessa definição, é “o resultado de uma má representação do mundo real” e, quando essa representação é moderada pelas necessidades dos usuários dos sistemas de informação, ela está sujeita às particularidades de cada usuário. Nesse sentido, há de se refletir sobre a qualidade de dados em domínios específicos (JAYAWARDENE; SADIQ; INDULSKA, 2015, tradução nossa), e até que ponto isso poderia ser um fator para definir recomendações para publicação de dados.

Ao contrário da definição apresentada por Liebenau e Backhouse em 1990, Price e Shanks (2005, p. 662) sugerem o termo representação 'externa' para se referir às “manifestações representadas” por se tratar de um termo mais geral, uma vez que inclui fenômenos do mundo real. Dessa forma, na visão de Price e Shanks (2005, p. 662), é possível distinguir dado de metadado. O metadado é definido pelos autores como “[...] regra representada que condiciona ou um documento que descreve fenômenos externos relevantes para um domínio de

aplicação específico ou organização de dados relevantes para um modelo de dado específico¹.

A publicação de dados refere-se ao fato de tornar os dados disponíveis e acessíveis na Web de forma que outros possam (re)utilizá-los (SANTOS *et al.*, 2018). A publicação de dados corresponde à terceira etapa do Ciclo de Vida dos Dados na Web, que inclui: o planejamento, a criação, a publicação, o acesso, o consumo, o *feedback* e o refinamento. O presente estudo visa identificar diretrizes para a publicação de dados na Web com o objetivo de apresentar um panorama do que tem sido mais representativo ao longo dos últimos 20 anos, incluindo a identificação de lacunas de investigação e os benefícios oriundos da adoção de boas práticas de publicação de dados.

Contextualização

Seja no contexto da *Big Science* ou seja na Ciência da cauda-longa (*Long-tail Science*) (HEIDORN, 2008), os dados se impõem como um motor para o avanço da Ciência e, no ambiente governamental, têm proporcionado melhoria dos serviços, insumo para tomadas de decisão e empoderamento da sociedade (GIL-GARCIA; GASCO-HERNANDEZ; PARDO, 2020).

No entanto, existem várias barreiras que dificultam a realização desses objetivos de forma eficaz (TENOPIR *et al.*, 2011; BORGMAN, 2012; GAMA; LÓSCIO, 2014; WIECZORKOWSKI, 2019). São barreiras associadas às questões técnicas, metodológicas, legais e financeiras que atingem tanto os dados científicos quanto os governamentais (DOUGLASS *et al.*, 2014). O que alguns organismos e/ou pesquisadores têm feito é introduzir boas práticas e mecanismos de partilhas de dados para facilitar sua publicação, a compreensão e a potencialização do seu reuso.

Um dos primeiros projetos a focar na publicação e citação de dados foi apresentado pelo Grupo CODATA, fundado pela German Science Foundation DFG nos anos de 2003-2005. Uma das práticas adotadas foi o uso de identificadores persistentes para identificar conjuntos de dados no formato digital, bem como a utilização de um conjunto de metadados para descrição (KLUMP *et al.*, 2006). No contexto da pesquisa científica, em 2009, era possível observar que a maioria dos periódicos que aceitava a publicação de dados o fazia como material suplementar de artigos de pesquisa no formato online,

ainda com muitas restrições devido ao volume, número de itens suplementares e *copyright*. Associada a essa forma de publicar os dados, observava-se a dificuldade em acessar, interpretar, além de reutilizá-los, um agravante para aqueles que entendem o dado como motor propulsor da ciência (BORGMAN, 2012). O fato de os dados começarem a fazer parte da comunicação científica, sendo integrados às publicações científicas, promoveu o surgimento de modelos de publicação como:

- a) *Data paper* (artigo de dados) – é um documento pesquisável por metadados que descreve um conjunto de dados específico ou um conjunto de dados publicados em forma de artigos revistos por pares em um periódico científico. Tem como foco principal descrever o grupo de dados, deixando hipóteses e conclusão em segundo plano (CHAVAN; PENEV, 2011, p. 3; GLOBAL BIODIVERSITY INFORMATION FACILITY, [ca. 2023]);
- b) *Data journal* (revista de dados) – revista especializada em publicação de *data papers*;
- c) *Data repository* (repositório de dados) – “são infraestruturas de investigação digital para arquivamento e distribuição de conjuntos de dados de investigação” (ELGER *et al.*, 2016, p. 122, tradução nossa). Adotam práticas sistemáticas de gestão de dados, de modo a promover a recolha, preservação, disseminação e acesso aos conjuntos de dados (MARCIAL; HEMMINGER, 2010).

Em síntese, a publicação de dados de pesquisa é um componente essencial na investigação e na comunicação científica, que impacta na reprodução de resultados e na geração de novos conhecimentos. A publicação é a fase em que os consumidores também podem se tornar produtores, uma vez que, por se tratar de dados abertos, estão sujeitos a melhorias e refinamentos que possibilitam a sua republicação (LÓSCIO; OLIVEIRA; BITTENCOURT, 2015; SANTOS *et al.*, 2018). Entretanto, preocupações que sempre estiveram presentes na comunidade científica são intensificadas, já que uma vez publicados sempre estarão disponíveis na Web (SANTOS *et al.*, 2018, p. 3). O valor dos dados e seu reuso no âmbito científico estão intimamente vinculados à adoção de boas práticas que garantam a sua



qualidade. Entretanto, segundo alguns autores, os padrões universais de qualidade dos dados não são úteis (RADULOVIC *et al.*, 2015; PILAT; FUKASAKO, 2007). Por essa razão, Pilat e Fukasako (2007) sugerem que as normas sejam definidas de acordo com as necessidades dos pesquisadores das diferentes disciplinas, o que se traduz em imensos desafios.

Já os dados governamentais abertos (OGD) podem ser definidos como dados gerados ou publicados por órgãos públicos que, em geral, incluem vários conjuntos de dados correspondentes a finanças, saúde, educação, população, clima, transporte e geográficos, entre outros. A disponibilização dos dados por parte dos governos está associada ao movimento de E-Government, que tem como objetivo principal a disponibilização de informação, que pode promover a criação de redes de colaboração e cocriação, transparência e participação dos cidadãos no processo de democratização (ATTARD *et al.*, 2015; GIL-GARCIA; GASCO-HERNANDEZ; PARDO, 2020). Na maioria dos casos são adotados os portais de dados para esse fim. Os portais de dados são considerados interfaces na Web planejados para facilitar a localização de informações e dados que podem ser reutilizáveis. Contêm registros de metadados dos conjuntos de dados, bem como os próprios conjuntos de dados. Os conjuntos de dados, na maioria das vezes, são disponibilizados em forma de tabelas, dados numéricos e, em alguns casos, documentos textuais (EUROPEAN COMMISSION, 2022).

A qualidade dos dados admite distintas definições em diferentes domínios (científico, governamental) e períodos, mas é consenso que, independentemente do domínio, a qualidade é imprescindível para a publicação e a reutilização dos dados, e a sua identificação é apropriada para utilização ou para satisfazer as necessidades dos utilizadores (SIDI *et al.*, 2012; JAYAWARDENE; SADIQ; INDULSKA, 2015). Ao realizar uma busca no Scholar Google pelo termo "*data quality*", em dezembro de 2021, foram 1.930.000 documentos relacionados com "*management*", "*assessment*", "*dimensions*", "*control*", "*metrics*", "*framework*", "*problems*" e "*healthcare*", entre outros, o que reforça a importância de ações que propiciem a qualidade dos dados, e, conseqüentemente, de estudos sobre essa temática.

No sentido de avançar em relação a um cenário em que os dados precisam ser efetivamente disponibilizados, descobertos e reutilizados"

(ARZBERGER *et al.*, 2004; ATTARD *et al.*, 2015; CANDELA *et al.*, 2015; RADULOVIC *et al.*, 2015; WILKINSON *et al.*, 2016; TRENHAM; STEER, 2019), esforços vêm sendo empreendidos para criar formas para disponibilização e gestão. Entretanto, é necessário que esses dados cumpram um critério básico: qualidade. Em vista do que foi exposto, quais boas práticas são identificadas em relação à qualidade dos dados?

Na sequência, são apresentados os procedimentos metodológicos adotados (Seção 2, Metodologia); os principais resultados (Seção 3, Resultados), seguidos das considerações finais (Seção 4, Conclusão) e, por último, as publicações científicas utilizadas (Referências).

METODOLOGIA

Este estudo não corresponde a uma revisão sistemática da literatura ou a um mapeamento sistemático - como os estudos publicados pelos autores Attard *et al.* (2015), Churi e Pawar (2019), Devriendt, Shabani e Borry (2021), Santos *et al.* (2018), Feitosa *et al.* (2018), Rashid e Torchiano (2016), Van den Eynden e Corti (2017), Zhang, Hua e Yuan (2017) –, mas sim a uma revisão integrativa sobre o que tem sido divulgado em relação às diretrizes para publicação de dados na Web.

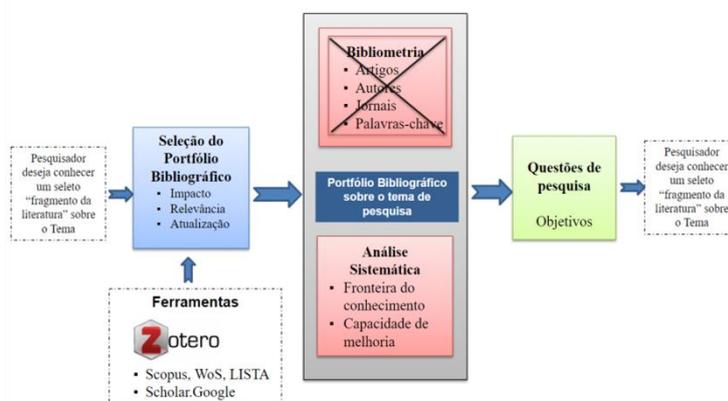
A revisão integrativa tem por objetivo analisar pesquisas relevantes em determinado campo de conhecimento (MENDES; SILVEIRA; GALVÃO, 2008; SOUZA; SILVA; CARVALHO, 2010; BOTELHO; CUNHA; MACEDO, 2011). Adotada com frequência para estudos na área médica tem se expandido para outras áreas devido à possibilidade de conclusões a respeito de um determinado tema, a partir de sínteses de múltiplos estudos (WHITTEMORE; KNAFL, 2005; MENDES; SILVEIRA; GALVÃO, 2008; BOTELHO; CUNHA; MACEDO, 2011; ESPER; CUNHA, 2015).

Em adição à proposta da revisão integrativa apresentada por Botelho *et al.* (2011), utilizou-se a metodologia Knowledge Development Process-Constructivist (ProKnow-C) (LACERDA; ENSSLIN; ENSSLIN, 2012; ENSSLIN *et al.*, 2014; VALMORBIDA *et al.*, 2016; VALMORBIDA; ENSSLIN; ENSSLIN, 2018), que busca produzir conhecimento com base em um fragmento da literatura científica. Para tanto, essa metodologia é estruturada em quatro fases:

- a) definição de um Portfólio Bibliográfico (PB) – conjunto de publicações científicas relevantes obtido a partir de uma seleção do banco de artigos bruto; filtragem do banco de artigos; e realização do teste de representatividade do Portfólio Bibliográfico;
- b) análise bibliométrica – busca identificar determinadas características nas publicações dos artigos do PB. O desenho deste estudo não contempla a análise bibliométrica;
- c) análise sistêmica – refere-se à análise crítica do PB selecionado, evidenciando lacunas e oportunidades de pesquisas; e
- d) identificação de onde e como se deseja intervir com base em justificativas científicas.

A Figura 1 apresenta, de forma concisa, etapas, ferramentas e procedimentos adotados para a realização deste estudo.

Figura 1 - Knowledge Development Process-Constructivist (ProKnow-C)



Fonte: adaptado de Valmorbidia *et al.* (2016, 2018).

Para a estratégia de busca, foi definida a seguinte equação: “*data publication*” OR “*data publishing*” AND “*principles OR practices*”. Os termos adotados na busca e para seleção do PB seguem a terminologia apresentada por Torino *et al.* (2020). Foram utilizadas as bases de dados Library, Information Science & Technology Abstract (LISTA), Scopus e Web of Science (WOS). As duas últimas são bases multidisciplinares de grande cobertura em várias áreas de conhecimento (SINGH *et al.*, 2021), enquanto que a base LISTA cobre a área de Ciência da Informação, alcançando assim um perfil de investigadores

apontado como relevante nesta temática. Também foi adotado o Scholar Google, uso também justificado pela literatura e que, neste estudo, acrescenta-se ainda a necessidade de obtenção de trabalhos que podem não ter sido indexados pelas bases de dados comerciais (relatórios, recomendações emitidas por organismos internacionais).

Para seleção do PB, foi definido que a busca pelos artigos incluiria documentos publicados a partir de 2001 por ter sido o período em que houve um incremento nas discussões sobre publicações de dados na Web. Foram incluídos artigos originais, revisões de literatura, relatórios ou outros documentos que permitissem extrair informações pertinentes ao objetivo do estudo, visto que se trata de um tema que pode ser resultante de um projeto, recomendações de organismos internacionais, ou de grupos de especialistas. Foram excluídos artigos de estudos de caso, propostas, ferramentas, modelos, e material não disponível em texto integral.

Cumprida essa fase, foi realizada a análise sistêmica dos documentos. Buscou-se extrair, além das recomendações para publicação de dados, conceitos norteadores de forma a clarificar algumas definições que possam ter interpretações diferenciadas por conta de domínios específicos. Assim como a revisão integrativa, o *Proknow-C* também recomenda que a análise da literatura seja realizada na perspectiva de lentes pré-definidas. Nesse sentido, foram definidas as seguintes lentes:

- a) Lente 1 – As diretrizes disponíveis atendem aos atuais paradigmas para publicação de dados na Web?
- b) Lente 2 – Qual é a tendência das diretrizes para publicação de dados na Web, universais ou específicas? Abrange os dados e metadados?
- c) Lente 3 – Qual tem sido o foco (ênfase) das diretrizes para publicação de dados na Web?

RESULTADOS

Definição do portfólio bibliográfico

A definição do Portifólio Bibliográfico é estruturada em três etapas. A primeira etapa corresponde à definição do Banco de Artigos Bruto, que inclui a formulação da estratégia de busca. Optou-se por iniciar a busca pelo Scholar Google. Devido à quantidade de artigos recuperados, os itens passaram por

uma análise inicial. Em seguida, os artigos não pertinentes foram eliminados (segunda etapa). Os itens selecionados foram exportados em formato BibTex para o Zotero®.

Os resultados das buscas no Scholar Google também indicaram que os termos adotados estavam adequados ao objetivo do estudo. A partir dessa constatação, prosseguiu-se com a pesquisa na Scopus, LISTA e WoS. A Tabela 1 mostra a estratégia de busca adotada e a quantidade de documentos obtidos no período de 12 a 21 de dezembro de 2021.

Tabela 1 – Quantitativo de documentos obtidos com a pesquisa bibliográfica

Termos	Scopus	WoS	LISTA	Scholar Google
"practices" AND "data publication" OR "data publishing"	205	52	10	990*
"principles" OR "data publication" OR "data publishing"	128	43	6	990*

Fonte: elaborado pelos autores (2022).

Nota: *A busca no Google Scholar gerou um resultado de 11.800 e 13.400, respectivamente. No entanto, só foi possível avaliar 990 itens por estratégia de busca, devido à limitação da ferramenta.

Todos os documentos recuperados foram exportados para o Zotero®, onde foram eliminadas as duplicidades e verificado o alinhamento dos títulos à proposta do estudo. Após essa triagem, 254 documentos foram submetidos à análise de reconhecimento científico. Nessa etapa, foram adotados os seguintes procedimentos:

- a) documentos com mais de dois anos de publicação – dos 146, foram selecionados 53 títulos. Esses títulos foram consultados quanto ao número de citações no Google Scholar. Para se obter o índice de relevância, foi aplicada a Regra de Pareto (LACERDA; ENSSLIN; ENSSLIN, 2012; ENSSLIN *et al.*, 2014). Dos 53 trabalhos submetidos a essa análise, 11 foram selecionados para leitura do texto na íntegra (no mínimo 23 citações); e
- b) análise dos resumos dos artigos com menos de dois anos. Foram obtidos 108 documentos publicados entre 2019-2021. Desses foram eliminados 54, pelo fato de o resumo não mostrar alinhamento com a proposta deste estudo; 1, em idioma chinês;

e 12, por não disponibilizar o texto completo. No total, foram selecionados 41 documentos para análise do texto na íntegra.

Assim, após leitura dos 52 documentos na íntegra para verificar o alinhamento com o objetivo e com os critérios de inclusão, concluiu-se a etapa de Filtragens do banco de artigo com 12 selecionados.

A última etapa do processo de obtenção do PB é o teste de representatividade dos denominados “artigos primários”. Nessa etapa, são avaliadas as referências dos artigos que passaram pelo processo de filtragem do banco de artigos (12). Para análise e verificação de pertinência das referências, foram considerados os seguintes procedimentos: leitura de título, resumo e leitura do texto completo. A partir dessa nova análise, foram recuperados mais 20 trabalhos, totalizando 32 documentos. Posteriormente, foi realizada a análise sistêmica dos 32 documentos (Tabela 2).

Tabela 2 – Portfólio bibliográfico

Autoria	Título
Attard <i>et al.</i> (2015)	A systematic review of open government data initiatives
Austin <i>et al.</i> (2017)	Key components of data publishing: using current best practices to develop a reference model for data publishing
Bennett e Harvey (2009)	Publishing Open Government Data. W3C Working Draft 8
Bizer <i>et al.</i> (2011)	The story so far
Daly, Mann e Devitt (2019)	Good data: theory on demand.
Dekker (2020)	Social data: CESSDA best practices.
Deppe (2020)	FAIR, CARE und mehr.
Derilinx, Lóscio e Archer (2015)	Data on the Web Best Practices Use Cases and Requirements.
Feitosa <i>et al.</i> (2018)	A systematic review on the use of best practices for publishing linked data.
GBIF Secretariat e IAIA (2020)	Best Practices for Publishing Biodiversity Data from Environmental Impact Assessments. Version 1.0.
Hasnain e Rebholz-Schuhmann (2018)	Assessing FAIR data principles against the 5-star open data principles
Hyland e Wood (2011)	The Joy of Data - A Cookbook for Publishing Linked Government Data on the Web.
Hyland, Atemezing e Villazon-Terrazas (2014)	Best practices for publishing linked data.
Klump <i>et al.</i> (2006)	Data publication in the open access initiative
Kučera <i>et al.</i> (2015)	Methodologies and best practices for Open Data publication
Lee, Cyganiak e Decker (2014)	Open Data Ireland: best practice handbook.
Lóscio <i>et al.</i> (2016)	Data on the Web Best Practices.
Lóscio <i>et al.</i> (2017)	Data on the Web Best Practices.
Lóscio <i>et al.</i> (2019)	Implementation of the Best Practices for Data on the Web in Brazil and Costa Rica
Lovett <i>et al.</i> (2019)	Good data practices for indigenous data sovereignty and governance.
Mons <i>et al.</i> (2020)	The FAIR principles: first generation implementation choices and



	challenges
Oliveira, Oliveira e Lóscio (2017)	Um Survey sobre Soluções para Publicação de Dados na Web sob a Perspectiva das Boas Práticas do W3C.
Penev <i>et al.</i> (2017)	Strategies and guidelines for scholarly publishing of biodiversity data
Penteado, Maldonado e Isotani (2021)	Methodologies for publishing linked open government data on the web: a systematic mapping and a unified process model
Radulovic <i>et al.</i> (2015)	Guidelines for Linked Data generation and publication: An example in building energy consumption
Redalyc <i>et al.</i> (2003)	Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities.
Santos <i>et al.</i> (2018)	Investigations into data published and consumed on the Web: a systematic mapping study
Tandy, Van den Brink e Barnaghi. (2017)	Spatial Data on the Web Best Practices
THE FUTURE OF RESEARCH COMMUNICATIONS AND E-SCHOLARSHIP FORCE11 (2014)	Guiding Principles for Findable, Accessible, Interoperable and Re-Usable Data Publishing Version B1.0
Trenham e Steer, 2019)	The Good Data Manifesto
van den Brink <i>et al.</i> (2019)	Best practices for publishing, retrieving, and using spatial data on the web
Wilkinson <i>et al.</i> (2016)	The FAIR Guiding Principles for scientific data management and stewardship

Fonte: elaborado pelos autores (2022).

Análise sistêmica

Observou-se, nas diretrizes selecionadas, que elas pontuam questões anteriores à própria Web, mas que ainda permanecem atuais e, mais do que nunca, relevantes para que a ciência e a própria sociedade possam usufruir da imensa quantidade de dados que é disponibilizada diariamente.

A publicação de dados na Web não se refere apenas ao fato de fornecer dados em formatos legíveis por máquina (KUČERA *et al.*, 2015), é preciso que sejam interoperáveis e reutilizáveis. Os "Objectos de Dados" não só devem ser devidamente licenciados como também os métodos de acesso e/ou descarregamento devem ser bem descritos e, de preferência, totalmente automatizados e utilizando protocolos bem estabelecidos (FORCE11, 2014).

É evidente que o panorama de publicação de dados é variado e dinâmico e que existem importantes lacunas e desafios. Os diferentes componentes de um sistema de publicação de dados precisam funcionar, na medida do possível, de forma integrada e sem descontinuidades para apoiar a evolução das normas normalmente compreendidas e utilizadas e -eventualmente - para aumentar a reprodutibilidade (AUSTIN *et al.*, 2017, tradução nossa).

De acordo com Wilkinson *et al.* (2016), facilitar a descoberta de conhecimento com o uso intensivo de dados se tornou um dos grandes desafios da ciência. Nesse contexto, tem-se trabalhado para que humanos e

máquinas descubram, acessem, integrem e analisem os dados científicos adequadamente “para as suas tarefas e os seus algoritmos e fluxo de trabalho associados”.

Para o reuso de dados abertos, é importante que seja garantida a qualidade dos dados, que, por sua vez, está associada às melhores práticas, como a adoção de licenças não proprietárias, metadados e formatos de dados legíveis por humanos e por máquinas, e identificadores únicos (LEE, CYGANIAK; DECKER, 2014; WILKINSON *et al.*, 2016; AUSTIN *et al.*, 2017; LÓSCIO; BURLE; CALEGARI, 2019). Vale destacar que o aumento da publicação de dados e a diversificação dos sistemas que os publicam terminam por impulsionar a evolução das boas práticas, até mesmo porque a Web é “um ecossistema tecnológico complexo em constante evolução” (PASTOR-SÁNCHEZ, 2016, p. 268).

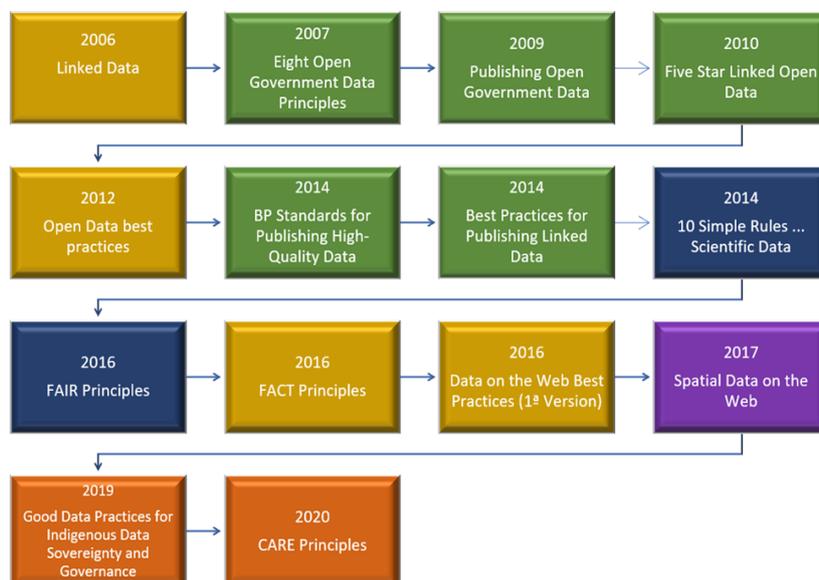
No entanto, um dos aspectos que merece reflexão se refere à formulação de diretrizes para domínios específicos ou gerais. Em 2011, Villazón-Terrazas *et al.* (2011) relataram que as diretrizes para publicação de dados ligados existentes à época não cobriam todas as etapas necessárias para publicação de dados. Nesse sentido, Radulovic *et al.* (2015) ratificaram essa posição ao afirmarem que diretrizes genéricas frequentemente não são suficientes para serem aplicadas para todos os domínios, daí o surgimento de iniciativas específicas em que os especialistas desenvolvem ferramentas, técnicas e diretrizes ajustadas às suas necessidades.

Não obstante, Austin *et al.* (2017) asseveram que “uma barreira significativa para avançar [na publicação dos dados] é a grande variação nas melhores práticas e normas entre e dentro das disciplinas”.

No sentido de apresentar o que tem sido elaborado para melhoria da qualidade de dados, foi desenvolvida a análise sistêmica dos artigos que fazem parte do PB. Inicialmente, é apresentada a Figura 2 que mostra boas práticas, recomendações e princípios (neste estudo, denominados de forma genérica como diretrizes) para publicação de dados, cujas cores correspondem ao domínio para o qual, inicialmente, foram projetadas. É possível observar que a maioria, quando formuladas inicialmente para um domínio específico, posteriormente tiveram seu uso ampliado para outros domínios. Na sequência, é apresentada a Tabela 3 com as 14 diretrizes

selecionadas, em que constam autores ou editores, título, uma síntese sobre elas e as fontes onde foram identificadas.

Figura 2 – Cronologia das diretrizes para publicação de dados na Web



Legenda: ● sem domínio específico; ● dados governamentais abertos; ● publicações científicas; ● dados indígenas; ● dados geoespaciais.

Fonte: elaborado pelos autores (2022).

Vale destacar que dois princípios FACT e CARE presentes na Figura 2 e na Tabela 3 não são voltados para descrição ou publicação dos dados, mas para aspectos, principalmente, éticos e responsáveis, que devem ser considerados ao publicar e/ou reutilizar os dados. A inclusão deles neste estudo será justificada na sequência.

Tabela 3 – Diretrizes para publicação de dados na Web

Título/Acrônimo	Autor/Editor	Resumo	Fonte
10 Simple Rules for the Care and Feeding of Scientific Data	Goodman <i>et al.</i> (2014)	Trata-se de um conjunto de 10 recomendações para gestão e publicação de dados. Além de incentivar a publicação de dados, sugere o uso de métodos e ferramentas seguras. Os dados devem ser publicados na Web usando identificadores persistentes; devem ser fornecidos metadados e documentação, adotar boas práticas de curadoria; fornecer informações sobre proveniência e etapas de processamento dos dados para que possam ser reutilizados; e conectar os dados às publicações.	
Best Practice Standards for Publishing High-Quality Data	Lee, Cyganiak e Decker (2014)	Conjunto de 10 práticas que inclui: realizar uma auditoria de dados, selecionar os dados a publicar; assegurar o cumprimento das leis de proteção de dados; associar dados com uma licença aberta; publicar dados como dados abertos de 3 a 5 estrelas; associar dados com metadados normalizados; utilizar padrões de dados; utilizar identificadores únicos; proporcionar acesso aos dados; publicar dados no Portal Nacional de Dados Abertos.	Kučera <i>et al.</i> (2015)
Best Practices for Publishing Linked Data	Hyland, Atemezing e Villazon-Terrazas (2014)	O documento apresenta uma compilação das práticas mais relevantes para gestão e publicação de dados governamentais com alta qualidade. Esse processo resultou em 10 práticas: qualificação de equipes envolvidas na gestão de dados; publicação de dados originais; modelagem de dados; especificação de licença; bons URIs para dados ligados; uso de vocabulários padronizados; conversão de dados; as máquinas devem ter acesso aos dados; divulgação dos dados para o público; e reconhecimento do contrato social.	Kučera <i>et al.</i> (2015)
CARE Principles	Carroll <i>et al.</i> (2020)	Os princípios CARE estão associados ao tratamento responsável de dados de modo que o pesquisador e o pesquisado tenham controle sobre eles, seja em relação aos benefícios, seja em relação ao uso. Prevê o Benefício coletivo; Autoridade para controlar; Responsabilidade e Ética.	Deppe (2020) Lovett <i>et al.</i> (2019)
Data on the Web Best Practices	Lóscio, Burle e Calegari (2017)	É apresentado um conjunto de 35 boas práticas relativas a 14 dimensões: Metadados; Licença de dados; Proveniência dos dados; Qualidade de dados; Versão dos dados; Identificadores de dados; Formatos de dados; Vocabulários de dados; Acesso aos dados; APIs de Acesso a dados;	Penteado, Maldonado e Isotani (2021); Derilinx; Lóscio; Archer (2015)

		Preservação de dados; <i>Feedback</i> ; Enriquecimento de dados; Republicação de dados. Associados às boas práticas são elencados benefícios como: reuso, interoperabilidade, compreensão, descoberta, processabilidade, confiabilidade, conectividade, interoperabilidade e acessibilidade. O documento indica elementos de metadados e ferramentas que podem ser adotados na representação e para publicação de dados	
Eight Open Government Data Principles	Malamud <i>et al.</i> (2014)	Esses princípios orientam que os dados sejam: Completo - todos os dados públicos devem ser disponibilizados respeitando as limitações de privacidade, segurança ou privilégios; Primário - os dados devem ser originais (como estão disponíveis nas fontes, sem agregações ou modificações); Temporal - os dados criados devem ser disponibilizados em tempo real ou o mais rápido possível; Acessível – os dados são disponibilizados para todos sem limitações de uso; Não-discriminatório – os dados devem estar disponíveis para uso de todos sem registro prévio; Sem licença proprietária – os dados não devem possuir limitação de uso devido a direitos autorais, patentes, marcas registradas ou regulamentos de segredo comercial	Attard <i>et al.</i> (2015) Tauberer (2014)
FACT principles	van der Aalst (2016)	Acrônimo para quatro princípios (<i>fairness, confidentiality, accuracy e transparency</i>) apresentados como meio de se obter a “ <i>Green Data Science</i> ” (GDS). São vinculados às iniciativas de <i>Responsible Data Science</i> que tem o objetivo de buscar soluções tecnológicas para proteção ao indivíduo, às organizações e a sociedade contra “os efeitos negativos dos dados.	Deppe (2020)
FAIR Principles	Wilkinson <i>et al.</i> (2016)	O termo FAIR foi cunhado durante o Workshop de Lorentz em 2014. Esse conjunto de 4 princípios e 14 subprincípios, publicados em 2016, tem por objetivo orientar a disponibilização de dados para que sejam Encontráveis, Acessíveis, Interoperáveis e Reutilizáveis.	Hasnain e Rebolz-Schuhmann (2018) Mons <i>et al.</i> (2020) THE FUTURE OF RESEARCH..., (2014)

Five Star Linked Open Data	Berners-Lee (2006)	São cinco sugestões para dados abertos apresentadas pelo criador da Web Tim Berners-Lee em 2010. Os dados devem ser disponibilizados na Web em qualquer formato com licença aberta (<i>open data</i>); como dados estruturados legíveis por máquina; em formato não proprietário; utilizando padrões abertos (RDF/SPARQL) e adotando o modelo Linked Open Data.	Attard <i>et al.</i> (2015) Penteado, Maldonado e Isotani (2021) Hasnain e Rebholz-Schuhmann (2018)
Good Data Practices for Indigenous Data Sovereignty and Governance	Lovett <i>et al.</i> (2019)	Essas práticas têm por objetivo proporcionar clareza sobre as definições de Indigenous Data Sovereignty (IDS) e Indigenous Data Governance (IDG); fornece uma visão geral do contexto histórico onde o IDS surgiu; e fornece exemplos de IDS e IDG em todo o espectro da comunidade, bem como política e práticas.	Daly, Mann e Devitt (2019)
Linked Data	Berners-Lee (2006)	Em 2006, Berners-Lee apresentou quatro orientações para publicação de dados, de modo que os dados pudessem ser conectados e legíveis por máquinas e humanos.	Kučera <i>et al.</i> (2015) Penteado, Maldonado e Isotani (2021) Deppe (2020)
Open Data best practices	Romain <i>et al.</i> (2012)	São defendidas 72 práticas organizadas em 13 categorias e em 3 níveis para publicação de dados em sites Web.	
Publishing Open Government Data	Bennett e Harvey (2009)	O documento apresenta três etapas onde são descritas normas e metodologias para a publicação de dados governamentais, de modo que sejam identificáveis, documentáveis, conectados, preservados e legíveis por máquinas e humanos.	Attard <i>et al.</i> (2015)
Spatial data on the Web	Tandy, Van Den Brink e Barnaghi (2017)	Advoga um conjunto de 14 boas práticas para os dados espaciais. São direcionadas sobretudo aos desenvolvedores da Web e peritos em geoespacialidade. São compiladas com base em provas de aplicação no mundo real e baseadas em normas gerais da Web.	van den Brink <i>et al.</i> (2019)

Fonte: elaborado pelos autores (2022).

Observa-se que os princípios Linked Data, defendidos por Berners-Lee (2006), têm como proposta transformar a World Wide Web numa grande base de dados global em que os dados podem ser conectados e/ou combinados com outros dados e com outros tipos de recursos. Como mencionado anteriormente, o volume de dados e as novas tecnologias impactam na definição de boas práticas que atendam às necessidades de determinados setores ou domínios. Isso pode ser visualizado na Figura 2 e na Tabela 3,

quando se constata a preocupação em formular recomendações que propiciem qualidade aos dados.

No período de 2009 a 2014, observa-se o desenvolvimento de várias diretrizes, na maioria voltadas para os dados publicados por órgãos públicos, são elas a *Eight Open Government Data Principles*, *Publishing Open Government Data*, *Five Star Linked Open Data*, *Best Practices for Publishing Linked Data*, *Best Practices Standards for Publishing High-Quality Data* (BENNETT; HARVEY, 2009; HYLAND *et al.*, 2014, LEE; CYGANIAK; DECKER, 2014; MALAMUD *et al.*, 2014; TAUBERER, 2014; 5-STAR OPEN DATA, 2015; KUČERA *et al.*, 2015; PENTEADO; MALDONADO; ISOTANI, 2021). Entre elas, pode-se destacar as *Best Practices for Publishing Linked Data*, uma compilação das práticas mais relevantes em todo o mundo para o gerenciamento e a publicação de dados governamentais, como dados ligados (HYLAND *et al.*, 2014). Vale destacar que a publicação de dados por governos e/ou agências governamentais têm avançado significativamente desde as primeiras iniciativas (BENNETT; HARVEY, 2009), o que propicia um aumento na qualidade da prestação de serviços públicos, maior transparência e o incremento de novos serviços oferecidos pelo setor público e/ou privado.

Em 2012, foram apresentadas 75 boas práticas para publicação de dados em sites Web (ROMAIN *et al.*, 2012; ROMAIN, 2013). À época, foi o maior número de itens abordados para publicação de dados. Foram incluídas orientações tanto para os dados como para os sites que os disponibilizam, assim como um *checklist* para verificação de cumprimento das práticas (ROMAIN *et al.*, 2012; ROMAIN, 2013). Ainda nesse período, é apresentada uma das primeiras recomendações ("conselhos") para comunidade científica para publicação de dados (*10 Simple Rules for the Care and Feeding of Scientific Data*), em que são descritas orientações sobre procedimentos e ferramentas (GOODMAN *et al.*, 2014). No ano de 2016, inicia-se a discussão de três diretrizes: os princípios FAIR, as recomendações da W3C com a publicação de uma versão preliminar das *Data on the Web Best Practices* (DWBP) e os princípios FACT.

Em 2017, são divulgados os Princípios FAIR como resultado de debates impulsionados pelo *Netherlands eScience Center and the Dutch Techcentre for the Life Sciences* (DTL) do Lorentz Center em Leiden, Holanda (WILKINSON *et*

al., 2016). Ao contrário do que Villazón-Terrazas *et al.* (2011) e Radulovic *et al.* (2015) apresentaram em seus artigos, os princípios FAIR inicialmente foram propostos para as necessidades da comunidade científica, mas rapidamente difundidos para outros domínios. O FAIR tem sido amplamente aceito e adotado por organismos internacionais e diferentes comunidades (PEREZ-LUQUE; ROS-CANDEIRA, 2019; POMMIER *et al.*, 2019; SHAW *et al.*, 2020). Esses princípios são apresentados como uma compilação das melhores práticas transdisciplinares e de alto nível para tornar os dados acessíveis, interoperáveis e reutilizáveis (KOESTEN *et al.*, 2020).

O FAIR vem sendo detalhado por meio de artigos, documentos, inclusive com disponibilização de ferramentas que auxiliam os gestores, publicadores e usuários a se inteirarem e adotarem esses princípios (WILKINSON *et al.*, 2016, 2018, 2019; MONS *et al.*, 2017, 2020; BAHIM *et al.*, 2020; DEVARAJU *et al.*, 2020)

Ainda no ano de 2017, é publicada a atual versão das DWBP (LÓSCIO; BURLE; CALEGARI, 2016, 2017). Essas recomendações têm por base o trabalho de pesquisa que antecedeu a sua publicação intitulado “Data on the Web Best Practices: use cases & requirements” (DERILINX; LÓSCIO; ARCHER, 2015). De acordo com Pastor-Sánchez (2016), as DWBP representam uma evolução no que se refere à publicação de dados na Web e em dados ligados. O documento inclui indicações de padrões para metadados e formatos de dados, identificadores, licença, proveniência, vocabulários, tecnologias, exemplos de casos de uso e um conjunto de benefícios que podem ser alcançados a partir da adoção de cada boa prática (LÓSCIO; BURLE; CALEGARI, 2016, 2017).

A W3C, com o objetivo de integrar os dados espaciais na Web de forma mais consistente, publicou também em 2017 as orientações Spatial Data on the Web. Essa diretriz tem como base as DWBP e se refere a padrões e soluções para publicação de informações sobre localização com o propósito de solucionar problemas de descoberta, acessibilidade e interoperabilidade (TANDY; VAN DEN BRINK; BARNAGHI, 2017; VAN DEN BRINK *et al.*, 2019).

Como mencionado no início desta seção, os princípios CARE e FACT não foram formulados para gestão de dados, mas sim sobre valores que

devem ser contemplados pelos dados como ética, imparcialidade, transparência, confiabilidade, responsabilidade e cuidado.

Em relação aos princípios FACT, o consórcio holandês Responsible Data Science (RDS), fundado em 2016, em resposta às crescentes preocupações da sociedade sobre o uso de dados e com o objetivo de desenvolver técnicas de ciência de dados, infraestrutura e abordagens para a RDS apresentou quatro princípios norteadores: *Fairness, Accuracy, Confidentiality and Transparency* (FACT) (VAN DER AALST *et al.*, 2017). Esses princípios estão associados à ideia de uma Green Data Science para o contexto de Big Data (VAN DER AALST, 2016).

Segundo Deppe (2020), embora se trate mais de questões metodológicas do que de gestão de dados de investigação no sentido mais estrito, os princípios FACT relacionam potenciais riscos quando se trata do *big data*. A preocupação com esses princípios é demonstrada por Helliwell (2019) ao mencionar que os dados do seu estudo estão alinhados aos princípios FAIR e FACT.

Os princípios FACT estão em consonância com a preocupação da violação da segurança, privacidade e restrições legítimas, questões que culminaram com a aprovação da Lei Geral de Proteção de Dados na Europa. A crescente preocupação com a ética, a responsabilidade, a transparência e a necessidade desses valores estarem representados na descrição dos dados ao serem publicados (WIECZORKOWSKI, 2019) são fatores que levaram a inclusão de tais princípios neste estudo, assim como os princípios CARE.

Os princípios CARE dizem respeito à governança dos dados que “[...] é o poder e a autoridade sobre a concepção, propriedade, acesso e utilização dos dados” pelos indígenas e por governos. A governança começa na fase de desenvolvimento, e continua pela fase de aplicação da ética, recolha, análise e relato de dados, e através da tradução de políticas (SNIPP, 2016).

Ainda que sejam poucas as publicações que abordam os princípios CARE e FACT em conjunto com outras diretrizes, vislumbra-se a necessidade de pensar como podem ser complementados. Esse pensamento é compartilhado por Carroll *et al.* (2021) ao propor essa ação para os princípios CARE e FACT.

A proposta externada por Carroll *et al.* (2021) em relação aos princípios CARE e FAIR deve-se ao fato de os princípios CARE contemplar um

ecossistema de dados indígenas extenso em que incluem os dados gerados ou armazenados por vários setores como organizações indígenas, governos, organizações não governamentais e instituições de pesquisa (RAINIE *et al.*, 2019). Esse pensamento não deveria estar restrito apenas ao FAIR, mas às boas práticas em geral, a exemplo das DWBP, como forma de se complementarem e assim ampliar o nível de qualidade dos dados conferido pelas atividades de publicadores e gestores. Ainda sobre o CARE, Carroll *et al.* (2021) destacam que esses princípios não se restringem aos indígenas, podem contemplar outras comunidades.

Como mencionado por Wilkinson *et al.* (2016), facilitar a descoberta de conhecimento com o uso intensivo de dados se tornou um dos grandes desafios da ciência, e, para o reuso de dados abertos, é importante que seja garantida a qualidade dos dados que, por sua vez, está associada às melhores práticas (LEE; CYGANIAK; DECKER, 2014).

CONSIDERAÇÕES FINAIS

Foram analisados 32 documentos e identificadas 14 diretrizes que trazem orientações no que se refere ao tratamento e à publicação dos dados, de forma a proporcionar qualidade, ética e responsabilidade para viabilizar o seu reuso com mais segurança.

Observa-se que alguns aspectos abordados pelas diretrizes analisadas vêm sendo discutidos desde o início da Web e ainda continuam sendo um desafio, principalmente quando estão relacionados à ética e à responsabilidade. Tal fato, provavelmente, decorre em razão de aumento do número de dados publicados, diversidade de sistemas e capacitação de pessoal. Fatores que, de certa forma, impulsionam a evolução das boas práticas e ferramentas. No entanto, é possível observar que, desde as primeiras diretrizes para publicação de dados na Web, os preceitos básicos continuam sendo ratificados pelas mais atuais, como as recomendações DWBP e os princípios FAIR, ao fornecerem orientações que auxiliam a publicação de dados de maneira a propiciar ou potencializar a qualidade desses dados.

Ao analisar as 14 diretrizes, foi possível identificar que são apresentadas boas práticas em relação a: Acesso, APIs, Dados e Dados ligados, Documentação, *Feedback*, Formalismo, Formatos, Identificadores, Licenças,

Metadados, Modelagem de dados, Políticas, Preservação, Processamento, Protocolos de comunicação, informações sobre Proveniência, Qualidade, Reúso, Serialização, Sintaxe, Versionamento e Vocabulários. No entanto, vale salientar que, embora sejam mencionadas a relevância da Confidencialidade, Controle sobre os dados, Ética, Imparcialidade, Justiça, Precisão e Transparência, ainda é necessário ampliar e formalizar as descrições e representações dessas práticas por meio de metadados e vocabulários. A esse respeito, observa-se que a Lei de Proteção Geral de Dados da União Europeia tem impulsionado algumas iniciativas em relação a esse objetivo.

Observa-se ainda o uso por diferentes comunidades de diretrizes universais, como as DWBP e o FAIR, em que as ferramentas, tecnologias, metadados e formatos podem ser ajustados às necessidades de domínios específicos.

Por outro lado, há questões que merecem ser analisadas, como: a) se a existência de várias diretrizes é uma questão facilitadora ou pode levar a desgastes desnecessários por parte dos atores envolvidos; b) se é justificável a formulação de diretrizes específicas para domínios específicos; c) se é viável reunir essas diretrizes em uma única diretriz, de modo que cada domínio pudesse adequá-las às suas necessidades ou às suas possibilidades técnicas e estruturais?

Por fim, como melhorar a qualidade dos seus dados? Adotando uma das boas práticas abordadas, a depender das necessidades do seu domínio e do nível de complexidade que pode implementar. Todas as diretrizes apresentadas trazem os pilares para publicação de dados na Web, sendo que as mais recentes – DWBP e FAIR – têm avançado em relação a questões sobre interoperabilidade, proveniência e uso de APIs. Observa-se ainda a preocupação em contemplar a ética e a responsabilidade dos dados a partir dos princípios FACT e CARE.

Vale destacar que as DWBP trazem informações mais detalhadas, o que facilita a compreensão de cada boa prática e seus benefícios, e os princípios FAIR vêm adicionando ferramentas que propiciam a publicação e a análise da qualidade dos dados de forma mais ágil.

Como trabalho futuro, pretende-se buscar respostas para as questões formuladas no texto e nomeadamente sobre a viabilidade de reunir essas diretrizes em uma única, de modo que cada domínio possa se adequar às suas necessidades ou às suas possibilidades técnicas e estruturais.

REFERÊNCIAS

5-STAR OPEN DATA. [s.l.]: 5-star open data, 2015. Disponível em: <http://5stardata.info/en/>. Acesso em: 19 dez. 2021.

ARZBERGER, P. *et al.* Promoting access to public research data for scientific, economic, and social development. **Data Science Journal**, [s.l.], v. 3, [s.n.], p. 135–152, 2004. DOI <https://doi.org/10.2481/dsj.3.135>.

ATTARD, J. *et al.* A systematic review of open government data initiatives. **Government Information Quarterly**, [s.l.], v. 32, n. 4, p. 399–418, 2015. DOI <https://doi.org/10.1016/j.giq.2015.07.006>.

AUSTIN, C. *et al.* Key components of data publishing: using current best practices to develop a reference model for data publishing. **International Journal on Digital Libraries**, [s.l.], v. 18, n. 2, p. 77–92, jun. 2017. DOI 10.1007/s00799-016-0178-2.

BAHIM, C. *et al.* The FAIR data maturity model: An approach to harmonise FAIR assessments. **Data Science Journal**, [s.l.], v. 19, n. 1, p. 1-7, 2020. DOI 10.5334/DSJ-2020-041.

BATINI, C.; SCANNAPIECA, M. Introduction to Data Quality. *In.*: BATINI, C.; SCANNAPIECA, M. (ed.). **Data quality: concepts, methodologies and techniques. data-centric systems and applications**. Berlin, Heidelberg: Springer, 2006. p. 1–18. DOI https://doi.org/10.1007/3-540-33173-5_1.

BENNETT, D.; HARVEY, A. **Publishing Open Government Data**. [s.l.]: W3C Working Draft, 2009. Disponível em: <https://www.w3.org/TR/gov-data/>. Acesso em: 17 dez. 2021.

BERNERS-LEE, T. **Linked Data: design issues**. [s.l.]: [s.n.], 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 17 dez. 2021.

BERNERS-LEE, T. **5 Star Linked Data: Government Linked Data**. [s.l.]: [s.n.], 2013. Disponível em: https://www.w3.org/2011/gld/wiki/5_Star_Linked_Data. Acesso em: 17 dez. 2021.

BIZER, C. *et al.* Linked data on the web (LDOW2008). *In.*: WWW '08: 17th INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, 8., 2008, Beijing. **Proceedings...** New York (NY): Association for Computing Machinery, 2008. DOI <https://doi.org/10.1145/1367497.1367760>

BIZER, C.; HEATH, T.; BERNERS-LEE, T. **Linked data**: principles and state of the art. In: WWW '08: 17th INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, 8., 2008, Beijing. **Proceedings...** New York (NY): Association for Computing Machinery, 2008. Disponível em: <https://www.w3.org/2008/Talks/WWW2008-W3CTrack-LOD.pdf>. Acesso em: 19 dez. 2021.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data: the story so far. **International Journal of Semantic Web Information Systems**, [s.l.], [s.n.], 2011. Disponível em: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>. Acesso em: 15 dez. 2021.

BORGMAN, C. L. The conundrum of sharing research data. **Journal of the American Society for Information Science and Technology**, v. 63, n. 6, p. 1059–1078, 2012. DOI <https://doi.org/10.1002/asi.22634>.

BORGMAN, C. L. **Big data, little data, no data**: scholarship in the networked world. [s.l.]: MIT press, 2017.

BOTELHO, L. L. R.; CUNHA, C. J. C. de A.; MACEDO, M. O método da revisão integrativa nos estudos organizacionais. **Gestão e sociedade**, v. 5, n. 11, p. 121–136, 2011. DOI <https://doi.org/10.21171/ges.v5i11.1220>.

CANDELA, L. *et al.* Data journals: a survey. **Journal of the Association For Information Science and Technology**, [s.l.], v. 66, n. 9, p. 1747–1762, set. 2015. DOI 10.1002/asi.23358.

CARROLL, S. R. *et al.* The CARE Principles for indigenous data governance. **Data Science Journal**, [s.l.], v. 19, n. 43, p. 1–12, 2020. DOI <https://doi.org/10.5334/dsj-2020-0432020>.

CARROLL, S. *et al.* Operationalizing the CARE and FAIR Principles for Indigenous data futures. **Scientific Data**, [s.l.], v. 8, n. 1, 16 abr. 2021. DOI 10.1038/s41597-021-00892-0.

CHAVAN, V.; PENEV, L. The data paper: a mechanism to incentivize data publishing in biodiversity science. **BMC Bioinformatics**, [s.l.], v. 12, n. 15, p. S2, 15 dez. 2011. Disponível em: <http://www.biomedcentral.com/1471-2105/12/S15/S2>. Acesso em: 19 dez. 2021.

CHIGNARD, S. **A brief history of Open Data**. [s.l.]: Paris Tech Review, 2013. Disponível em: <http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data>. Acesso em: 18 dez. 2021.

CHURI, P. P.; PAWAR, A. V. A systematic review on privacy preserving data publishing techniques. **Journal of Engineering Science & Technology Review**, [s.l.], v. 12, n. 6, 2019. DOI 10.25103/jestr.126.03.

DALY, A.; MANN, M.; DEVITT, S. K. **Good data**: theory on demand. Amsterdam: Institute of Network Cultures, 2019. Disponível em: https://eprints.qut.edu.au/125605/1/Good_Data_book.pdf. Acesso em: 17 dez. 2021.

DEKKER, R. Social data: Cessda best practices. **Data Intelligence**, [s.l.], v. 2, n. 1–2, p. 220–229, 2020. DOI 10.1162/dint_a_00044.

DEPPE, A. FAIR, CARE und mehr: prinzipien für einen verantwortungsvollen Umgang mit Forschungsdaten. In.: Matthias Schulze (Hrsg.). **Historisches Erbe und zeitgemäße Informationsinfrastrukturen: bibliotheken am Anfang des 21. Jahrhunderts**. Kassel: Kassel University Press, 2020. DOI 10.17170/kobra-202010131934-299.

DERILINX, D. L.; LÓSCIO, B. F.; ARCHER, P. **Data on the web best practices use cases and requirements**. [s.l.]: W3C, 2015. Disponível em: <http://www.w3.org/TR/2015/NOTE-dwbp-ucr-20150224/>. Acesso em: 17 dez. 2021.

DEVARAJU, A. *et al.* FAIRSFAR data object assessment metrics. **Zenodo**, 2020. DOI 10.5281/zenodo.408121.

DEVRIENDT, T.; SHABANI, M.; BORRY, P. Data sharing in biomedical sciences: a systematic review of incentives. **Biopreservation and Biobanking**, v. 19, n. 3, p. 219–227, 2021. DOI 10.1089/bio.2020.0037.

DIRETRIZ. In.: INFOPÉDIA. [s.l.]: Dicionário Porto Editora, 2023. Disponível em: <https://www.infopedia.pt/dicionarios/lingua-portuguesa/diretrizes>. Acesso em: 2 abr. 2022.

DOUGLASS, K. *et al.* Managing scientific data as public assets: data sharing practices and policies among full-time government employees. **Journal of the Association for Information Science and Technology**, [s.l.], v. 65, n. 2, p. 251–262, 2014. DOI <https://doi.org/10.1002/asi.22988>.

ELGER, K. *et al.* Open research data, data portals and data publication—an introduction to the data curation landscape. **Polarforschung**, [s.l.], v. 85, n. 2, p. 119–133, 2016. DOI 10.2312/polfor.2016.009.

ENSSLIN, S. R. *et al.* Processo de mapeamento das publicações científicas de um tema: portfólio bibliográfico e análise bibliométrica sobre avaliação de desempenho de cooperativas de produção agropecuária. **Revista de Economia e Sociologia Rural**, [s.l.], v. 52, [s.n.], p. 587–608, 2014. DOI 10.22004/ag.econ.211729.

ESPER, A. J. F.; CUNHA, C. J. C de A. Liderança autêntica: uma revisão integrativa. **Navus-Revista de Gestão e Tecnologia**, [s.l.], v. 5, n. 2, p. 60–72, 2015. DOI 10.18815/navus.v5i2.254.

EUROPEAN COMISSION. **Open data portals: shaping Europe's digital future**. [s.l.]: European Comission, 2022. Disponível em: <https://digital-strategy.ec.europa.eu/en/policies/open-data-portals>. Acesso em: 15 jan. 2022.
FEITOSA, D. *et al.* A systematic review on the use of best practices for publishing linked data. **Online Information Review**, 2018. DOI 10.1108/OIR-11-2016-0322.

GAMA, K.; LÓSCIO, B. F. Towards ecosystems based on open data as a service. **ICEIS (2)**, [s.l.], v. 10, [s.n.], p. 0004974506590664, 2014.



GBIF SECRETARIAT; IAIA. Best practices for publishing biodiversity data from environmental impact assessments. Copenhagen: GBIF Secretariat. 2020. DOI <https://doi.org/10.35035/doc-5xdm-8762>

GIL-GARCIA, J. R.; GASCO-HERNANDEZ, M.; PARDO, T. A. Beyond transparency, participation, and collaboration? A reflection on the dimensions of open government. **Public Performance & Management Review**, [s.l.], v. 43, n. 3, p. 483–502, 2020. DOI <https://doi.org/10.1080/15309576.2020.1734726>.

GLOBAL BIODIVERSITY INFORMATION FACILITY. **Data papers**. Denmark: Global Biodiversity Information Facility, [ca. 2023]. Disponível em: <https://www.gbif.org/data-papers>. Acesso em: 15 dez. 2021.

GOODMAN, A. *et al.* Ten simple rules for the care and feeding of scientific data. **PLOS Computational Biology**, [s.l.], v. 10, n. 4, p. e1003542, 24 abr. 2014. DOI <https://doi.org/10.1371/journal.pcbi.1003542>.

HASNAIN, A.; REBHOLZ-SCHUHMANN, D. Assessing FAIR data principles against the 5-star open data principles. *In.*: THE SEMANTIC WEB: ESWC 2018 SATELLITE EVENTS, 2018, Greece. **Proceedings...** Greece: Springer, 2018. DOI 10.1007/978-3-319-98192-5_60.

HEIDORN, P. B. Shedding light on the dark data in the long tail of science. **Library Trends**, [s.l.], v. 57, n. 2, p. 280–299, 2008. DOI 10.1353/lib.0.0036.

HELLIWELL, J. R. FACT and FAIR with Big Data allows objectivity in science: the view of crystallography. **Structural Dynamics**, [s.l.], v. 6, n. 5, 2019. DOI 10.1063/1.5124439.

HYLAND, B.; ATEMEZING, G.; VILLAZON-TERRAZAS, B. **Best practices for publishing linked data**. [s.l.]: W3C, 2014. Disponível em: <http://www.w3.org/TR/2014/NOTE-ld-bp-20140109/>. Acesso em: 15 dez. 2021.

HYLAND, B.; WOOD, D. The joy of data: a cookbook for publishing linked government data on the web. *In.*: WOOD, D. **Linking government data**. [s.l.]: Springer, 2011. p. 3–26. DOI 10.1007/978-1-4614-1767-5_1.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION [ISO]. ISO 8000-2:2020 (en). **Data quality – Part 2: Vocabulary**. Disponível em: <https://www.iso.org/obp/ui/#iso:std:iso:8000:-2:ed-4:v1:en>. Acesso em: nov. 2021.

JAYAWARDENE, V.; SADIQ, S.; INDULSKA, M. An analysis of data quality dimensions. **ITEE Technical Report**. Queensland: University of Queensland, 2015. Disponível em: <https://espace.library.uq.edu.au/view/UQ:312314>. Acesso em: 20 jan. 2022.

KLUMP, J. *et al.* Data publication in the open access initiative. **Data Science Journal**, [s.l.], v. 5, p. 79–83, 2006. DOI 10.2481/dsj.5.79

KOESTEN, L. *et al.* Dataset reuse: translating principles to practice. **Patterns**, [s.l.], v. 1, n. 8, 2020. Disponível em: <https://doi.org/10.1016/j.patter.2020.100136>. Acesso em: 16 jan. 2022.

KUČERA, J. *et al.* Methodologies and best practices for open data publication. *In.*: NECASKY M.; MORAVEC P.; POKORNY J. (ed.). CEUR Workshop Proc. **Proceedings...** CEUR-WS, 2015. Disponível em: <http://ceur-ws.org/Vol-1343/paper5.pdf>. Acesso em: 15 dez. 2021.

LACERDA, R. T. DE O.; ENSSLIN, L.; ENSSLIN, S. R. Uma análise bibliométrica da literatura sobre estratégia e avaliação de desempenho. **Gestão & Produção**, [s.l.], v. 19, p. 59–78, 2012. DOI <https://doi.org/10.1590/S0104-530X2012000100005>.

LEE, D.; CYGANIAK, R.; DECKER, S. **Open data Ireland**: best practice handbook. [s.l.]: [s.n.], 2014. Disponível em: https://data.gov.ie/uploads/page_images/2019-04-24-104746.825980Open-Data-Publication-Handbook.pdf. Acesso em: 15 dez. 2021.

LÓSCIO, B. F.; BURLE, C.; CALEGARI, N. **Data on the web best practices**. [s.l.]: W3C, 2017. Disponível em: <https://www.w3.org/TR/2016/PR-dwbp-20161215/>. Acesso em: 15 dez. 2021.

LÓSCIO, B. F.; BURLE, C.; CALEGARI, N. **Data on the web best practices**. [s.l.]: W3C, 2016. Disponível em: <https://www.w3.org/TR/dwbp/>. Acesso em: 15 dez. 2021.

LÓSCIO, B. F.; BURLE, C.; CALEGARI, N. Implementation of the Best Practices for Data on the Web in Brazil and Costa Rica. *In.*: WWW '19: THE 2019 WORLD WIDE WEB CONFERENCE, 19., 2019, San Francisco. **Proceedings...** New York (NY): Association for Computing Machinery, 2019. DOI 10.1145/3308560.3316463

LÓSCIO, B. F.; OLIVEIRA, M. I. S.; BITTENCOURT, I. I. Publicação e consumo de dados na Web. *In.*: HARA, C. S.; PORTO, F.; OGASAWARA, E. (ed.). **Tópicos em gerenciamento de dados e informações 2015**. Petrópolis: Laboratório Nacional de Computação Científica, 2015. DOI 10.1145/3308560.3316463

LOVETT, R. *et al.* Good data practices for Indigenous data sovereignty and governance. *In.*: DALY, A.; DEVITT, S. K.; MANN, M. (ed.). **Good data**: theory on demand. Amsterdam: Institute of Network Cultures, 2019. p. 26–36. Disponível em: <https://eprints.qut.edu.au/125605/>. Acesso em: 19 dez. 2021.

MALAMUD, C. *et al.* The 8 Principles of Open Government Data. **Open Government Working Group**. 2013. Disponível em: <http://www.opengovdata.org/home/8principles>. Acesso em: 15 dez. 2021.

MARCIAL, L. H.; HEMMINGER, B. M. Scientific data repositories on the Web: An initial survey. **Journal of the American Society for Information Science and Technology**, [s.l.], v. 61, n. 10, p. 2029–2048, out. 2010. <https://doi.org/10.1002/asi.21339>.



MENDES, K. D. S.; SILVEIRA, R. C. de C. P.; GALVÃO, C. M. Revisão integrativa: método de pesquisa para a incorporação de evidências na saúde e na enfermagem. **Texto & contexto-enfermagem**, [s.l.], v. 17, p. 758–764, 2008. DOI <https://doi.org/10.1590/S0104-07072008000400018>.

MONS, B. *et al.* Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. **Information Services & Use**, [s.l.], v. 37, n. 1, p. 49–56, 2017. DOI 10.3233/ISU-170824.

MONS, B. *et al.* The FAIR principles: first generation implementation choices and challenges Data Intelligence. **Data Intelligence**, [s.l.], v. 2, n. 1-2, p. 1-9, 2020. DOI https://doi.org/10.1162/dint_e_00023.

OLIVEIRA, L. E. R. A.; OLIVEIRA, M. I. S.; LÓSCIO, B. F. Um Survey sobre soluções para publicações de dados na Web sob a perspectiva das boas práticas do W3. 32nd SBBB, 32., 2017, Uberlândia. **Anais...** Uberlândia: [s.n.], 2017. Disponível em: <https://sbbd.org.br/2017/wp-content/uploads/sites/3/2018/02/p148-159.pdf>. Acesso em: 18 dez. 2021.

PASTOR-SÁNCHEZ, J.-A. Quince años de web semántica: de las tecnologías a las buenas prácticas. **Anuario ThinkEPI**, [s.l.], v. 10, [s.n.], p. 264–268, 2016. DOI <https://doi.org/10.3145/thinkepi.2016.58>.

PENEV, L. *et al.* Strategies and guidelines for scholarly publishing of biodiversity data. **Research Ideas and Outcomes**, [s.l.], v. 3, [s.n.], p. e12431, 2017. DOI <https://doi.org/10.3897/rio.3.e12431>.

PENTEADO, B. E.; MALDONADO, J. C.; ISOTANI, S. Methodologies for publishing linked open government data on the web: a systematic mapping and a unified process model. **Semantic Web Journal**, [s.l.], [s.n.], [s.n.], 2020. Disponível em: <http://www.semantic-web-journal.net/content/methodologies-publishing-linked-open-government-data-web-systematic-mapping-and-unified>. Acesso em: dez. 2021.

PEREZ-LUQUE, A.; ROS-CANDEIRA, A. Sharing data in Ecology: how to add more value to data. **Ecosistemas**, [s.l.], v. 28, n. 3, p. 150–159, set. 2019. DOI 10.7818/ECOS.1838

PILAT, D.; FUKASAKU, Y. OECD principles and guidelines for access to research data from public funding. **Data Science Journal**, [s.l.], v. 6, p. OD4–OD11, 2007.

POMMIER, C. *et al.* Applying FAIR Principles to Plant Phenotypic Data Management in GnpIS. **Plant Phenomics**, [s.l.], v. 2019, [s.n.], 2019. DOI <https://spj.science.org/doi/10.34133/2019/1671403>.

PRICE, R.; SHANKS, G. A Semiotic information quality framework: development and comparative analysis. **Journal of Information Technology**, [s.l.], v. 20, p. 658–672, jun. 2005. DOI 10.1057/palgrave.jit.2000038.

RADULOVIC, F. *et al.* Guidelines for linked data generation and publication: an example in building energy consumption. **Automation in Construction**, [s.l.], v. 57, p. 178–187, 2015. DOI 10.1016/j.autcon.2015.04.002.

RAINIE, S. C. *et al.* **Indigenous data sovereignty**. Cape Town: African Minds and International Development Research Centre, 2019. Disponível em: <https://press-files.anu.edu.au/downloads/press/n2140/pdf/book.pdf>. Acesso em: 19 dez. 2021.

RASHID, M.; TORCHIANO, M. A systematic literature review of open data quality in practice. *In.*: 2nd OPEN DATA RESEARCH SYMPOSIUM (ODRS), 2., 2016, Madrid. **Proceedings...** Madrid: [s.n.], 2016. Disponível em: <http://hdl.handle.net/11583/2648966>. Acesso em: 16 dez. 2021.

REDALYC *et al.* **Berlin Declaration on open access to knowledge in the sciences and humanities**. Berlin: MaxPlank, 2003. Disponível em: <https://openaccess.mpg.de/Berlin-Declaration>. Acesso em: 19 dez. 2021.

ROMAIN, P. *et al.* **Open Data best practices Version 1.1**. [s.l.]: [s.n.], 2012. Disponível em: https://res.cloudinary.com/opquast/image/upload/v1556788845/checklists/Opquast-Opedata_20120831_1.pdf. Acesso em: 19 DEZ. 2021.

ROMAIN, P. **A good practices checklist for opendata**. [s.l.]: W3C, 2013. Disponível em: <https://www.w3.org/2013/04/odw/eliie.pdf/>. Acesso em: 19 dez. 2021.

SANTOS, H. D. A. *et al.* Investigations into data published and consumed on the Web: a systematic mapping study. **Journal of the Brazilian Computer Society**, [s.l.], v. 24, n. 1, p. 1–22, 2018. DOI <https://doi.org/10.1186/s13173-018-0077-z>.

SHAW, F. *et al.* COPO: a metadata platform for brokering FAIR data in the life sciences. **bioRxiv 782771**, [s.l.], v. 9, n. 495, p. 495, 2020. DOI <https://doi.org/10.1101/782771>

SIDI, F. *et al.* Data quality: A survey of data quality dimensions. *In.*: 2012 INTERNATIONAL CONFERENCE ON INFORMATION RETRIEVAL KNOWLEDGE MANAGEMENT. [s.n.], 2012, Kuala Lumpur. **Proceedings...** Kuala Lumpur: IEEE xplore, 2012. p. 300-304, DOI 10.1109/InfRKM.2012.6204995.

SINGH, V. K. *et al.* The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. **Scientometrics**, [s.l.], v. 126, n. 6, p. 5113–5142, 1 jun. 2021. DOI <https://doi.org/10.1007/s11192-021-03948-5>.

SNIPP, C. M. What does data sovereignty imply: What does it look like. *In.*: WALTER, M. *et al.* (ed.). **Indigenous data sovereignty: toward an agenda**. Canberra: ANU Press, 2016. Disponível em: <https://press-files.anu.edu.au/downloads/press/n2140/pdf/book.pdf>. Acesso em: 19 dez. 2021.



SOUZA, M. T.; SILVA, M. D.; CARVALHO, R. Revisão integrativa: o que é e como fazer. **Einstein**, São Paulo, v. 8, [s.n.], p. 102–106, 2010. Disponível em: <https://www.scielo.br/j/eins/a/ZQTBkVJZqcWrTT34cXLjtBx/?format=pdf&lang=pt>. Acesso em: 05 dez. 2021.

TANDY, J.; VAN DEN BRINK, L.; BARNAGHI, P. **Spatial Data on the Web Best Practices**. [s.l.]: W3C, 2017. Disponível em: <https://www.w3.org/TR/sdw-bp/>. Acesso em: 15 dez. 2021.

TAUBERER, J. Open government data definition: the 8 principles of open government data. In.: Tauberer, J. **Open government data: the book**. 2. ed. New Hampshire: [s.n.]. 2014. Disponível em: <https://opengovdata.io/2014/8-principles/>. Acesso em: 15 dez. 2021.

TENOPIR, C. *et al.* Data sharing by scientists: practices and perceptions. **PLoS one**, [s.l.], v. 6, n. 6, p. e21101, 2011. DOI <https://doi.org/10.1371/journal.pone.0021101>.

THE FUTURE OF RESEARCH COMMUNICATIONS AND E-SCHOLARSHIP FORCE11. **Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data** [s.l.]: FORCE11, 2014. Disponível em: <https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-data-publishing-version-b1-0/>. Acesso em: 19 dez. 2021.

TORINO, E.; ROA-MARTÍNEZ, S. M.; VIDOTTI, S. A. B. G. Dados de pesquisa: disponibilização ou publicação? In.: Shintaku, M.; SALES, L. F.; Costa, M. (org.). **Tópicos sobre dados abertos para editores científicos**. Botucatu (SP): ABEC, 2020. p. 183-201. DOI 10.21452/978-85-93910-04-3.cap15.

TRENHAM, C.; STEER, A. The good data manifesto. In: DALY, A.; DEVITT, S. K.; MANN, M. (ed.). **Good Data**. Amsterdam: Institute of Network Cultures, 2019. p. 37. Disponível em: https://www.researchgate.net/publication/331985183_The_Good_Data_Manifesto. Acesso em: 19 dez. 2021.

VALMORBIDA, S. M. I. *et al.* Rankings universitários mundiais. Que dizem os estudos internacionais? **REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, [s.l.], v. 14, n. 2, p. 5–29, 2016. DOI <https://doi.org/10.15366/reice2016.14.2.001>.

VALMORBIDA, S. M. I.; ENSSLIN, S. R.; ENSSLIN, L. Avaliação de desempenho e contabilidade gerencial: revisão integrativa da literatura para superar as dificuldades de aplicação prática da avaliação de desempenho na gestão organizacional. **Revista Contabilidade, Gestão e Governança**, [s.l.], v. 21, n. 3, p. 339–360, 2018. DOI https://doi.org/10.51341/1984-3925_2018v21n3a3.

VAN DEN BRINK, L. *et al.* Best practices for publishing, retrieving, and using spatial data on the web. **Semantic Web**, [s.l.], v. 10, n. 1, p. 95–114, 2019. Disponível em: <http://www.semantic-web-journal.net/system/files/swj1711.pdf>. Acesso em: 15 dez. 2021.

VAN DEN EYNDEN, V.; CORTI, L. Advancing research data publishing practices for the social sciences: from archive activity to empowering researchers. **International Journal on Digital Libraries**, [s.l.], v. 18, n. 2, p. 113–121, 2017. DOI 10.1007/s00799-016-0177-3.

VAN DER AALST, W. Green Data Science: using big data in an “environmentally friendly” manner. *In.*: 18TH INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS. 18., 2016, Portugal. **Anais...** Portugal: SCITEPRESS, 2016. Disponível em: <https://www.scitepress.org/Link.aspx?doi=10.5220/0006806900010001>. Acesso em: 23 dez. 2021.

VAN DER AALST, W. M. P.; BICHLER, M.; HEINZL, A. Responsible Data Science. **Business and Information Systems Engineering**, [s.l.], v. 59, n. 5, p. 311–313, 2017. DOI 10.1007/s12599-017-0487-z

VILLAZÓN-TERRAZAS, B. *et al.* Methodological guidelines for publishing government linked data. *In.*: WOODS, D. (ed.). **Linking government data**. New York (NY): Springer, 2011. p. 27–49. DOI https://doi.org/10.1007/978-1-4614-1767-5_2.

WHITEMORE, R.; KNAFL, K. The integrative review: updated methodology. **Journal of advanced nursing**, [s.l.], v. 52, n. 5, p. 546–553, 2005. DOI 10.1111/j.1365-2648.2005.03621.x.

WIECZORKOWSKI, J. Barriers to using open government data. *In.*: 3RD INTERNATIONAL CONFERENCE ON E-COMMERCE, E-BUSINESS AND E-GOVERNMENT. 3., 2019, Lyon. **Proceedings...** Lyon, 2019. p. 15-20. DOI <https://doi.org/10.1145/3340017.3340022>.

WILKINSON, M. D. *et al.* A design framework and exemplar metrics for FAIRness. **Scientific data**, [s.l.], v. 5, n. 1, p. 1–4, 2018. DOI 10.1038/sdata.2018.118.

WILKINSON, M. D. *et al.* Evaluating FAIR maturity through a scalable, automated, community-governed framework. **Scientific data**, [s.l.], v. 6, n. 1, p. 1–12, 2019. DOI <https://doi.org/10.1038/s41597-019-0184-5>.

WILKINSON, M. D. *et al.* The FAIR guiding principles for scientific data management and stewardship. **Scientific data**, [s.l.], v. 3, n. 1, p. 1–9, 2016. DOI 10.1038/sdata.2016.18.

ZHANG, Y.; HUA, W.; YUAN, S. Mapping the scientific research on open data: A bibliometric review. **Learned Publishing**, [s.l.], v. 31, n. 2, p. 95–106, 2017. DOI <https://doi.org/10.1002/leap.1110>.

