

# Gestão de dados científicos: um estudo bibliométrico e altmétrico na Dimensions

Research data management: a bibliometric and altmetric study based on Dimensions

Marília Catarina Andrade Gontijo<sup>1</sup>, Raíssa Yuri Hamanaka<sup>2</sup>, Ronaldo Ferreira de Araujo<sup>3</sup>

**Como citar (APA):** Gontijo, M.C.A.; Hamanaka, R.Y.; & Araújo, R. F. (2021). Gestão de dados científicos: um estudo bibliométrico e altmétrico na Dimensions. *Iberoamerican Journal of Science Measurement and Communication*; 1(3), 1-19. <https://doi.org/10.47909/ijsmc.120>

**Recebido:** 12-10-2021

**Aceito:** 16-12-2021

**Copyright:** © 2021 Gontijo et al. This is an open access article distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

## RESUMO

**Objetivo.** A pesquisa objetivou analisar a produção científica sobre gestão de dados científicos indexada na base de dados Dimensions.

**Design/Metodologia/Abordagem.** A partir da busca pelo termo “research data management” foram recuperados 677 artigos, analisados por meio de indicadores bibliométricos e altmétricos. O sistema Altmetrics.com foi utilizado na coleta de dados em fontes virtuais alternativas para a aferição da atenção on-line recebida pelos artigos recuperados. Também foi utilizado o software VOSviewer para a elaboração de redes bibliométricas de acoplamento bibliográfico de periódicos e coocorrência de palavras-chave no título.

**Resultados/Discussão.** A partir da análise das 677 pesquisas recuperadas, observou-se um crescimento da produção científica ao longo dos anos (1970-2021). Os países/regiões com maiores índices de publicações foram os EUA (133 artigos), Alemanha (91) e Reino Unido (53). Entre os autores mais produtivos se destacam Andrew Martin Cox (13 artigos), Stephen Pinfield e Marta Teperek (sete cada) e Mary Anne Kennan e Amanda L. Whitmire (seis cada). Os periódicos científicos que mais apresentaram artigos publicados foram International Journal of Digital Curation, Journal of eScience Librarianship e Data Science

1 Universidade Federal de Minas Gerais, Escola de Ciencia da Informacao, Belo Horizonte, Minas Gerais, Brazil. Email: [mariliacgontijo@gmail.com](mailto:mariliacgontijo@gmail.com). ORCID: 0000-0002-9181-0302.

2 Universidade Estadual de Londrina, Departamento de Ciência da Informação, Rodovia Celso Garcia Cid, Pr 445 Km 380, Campus Universitário Cx. Postal 10.011, Londrina 86057-970, Brazil. Autor correspondente: [raissa.hamanaka@uel.br](mailto:raissa.hamanaka@uel.br). ORCID: 0000-0001-9516-5825.

3 Universidade Federal de Alagoas; Maceio, Alagoas; Brazil. Email: [ronaldfa@gmail.com](mailto:ronaldfa@gmail.com), ORCID: 0000-0003-0778-9561.

Journal, com 74, 38 e 25 publicações respectivamente, e entre as áreas de pesquisa: Information and Computing Sciences com 483, Information Systems com 362 e Library and Information Studies com 113 artigos. E as temáticas com maiores proximidades e ocorrências na produção sobre gestão de dados científicos, além do termo “research data management”, foram “university” e “data”.

**Conclusões.** A multidisciplinaridade em gestão de dados científicos foi demonstrada pelas publicações ocorrerem em diferentes campos de pesquisa, como Information and Computing Sciences, Information Systems, Library and Information Studies, Medical and Health Sciences e History and Archaeology. Os países com maiores índices de publicações foram Estados Unidos, Alemanha e Reino Unido. Cerca de 60% das publicações tiveram pelo menos uma citação, com um total de 3.598 citações encontradas, caracterizando-se um impacto acadêmico crescente.

**Originalidade/Valor.** Este estudo bibliométrico e altmétrico permitiu acompanhar o desempenho da literatura sobre gestão de dados científicos. A temática foi investigada na base Dimensions e analisada por meio de indicadores de produtividade, impacto e atenção on-line.

**Palavras-chave:** gestão de dados científicos, bibliometria, altmetria, produção científica

## ABSTRACT

**Objective.** This study aims to analyze the scientific production on research data management indexed in the Dimensions database.

**Design/Methodology/Approach.** Using the term “research data management” in the Dimensions database, 677 articles were retrieved and analyzed employing bibliometric and altmetric indicators. The Altmetrics.com system was used to collect data from alternative virtual sources to measure the online attention received by the retrieved articles. Bibliometric networks from journals bibliographic coupling and keywords co-occurrence were generated using the VOSviewer software.

**Results/Discussion.** Growth in scientific production over the period 1970-2021 was observed. The countries/regions with the highest rates of publications were the USA, Germany, and the United Kingdom. Among the most productive authors were Andrew Martin Cox, Stephen Pinfield, Marta Teperek, Mary Anne Kennan, and Amanda L. Whitmire. The most productive journals were the International Journal of Digital Curation, Journal of eScience Librarianship, and Data Science Journal, while the most representative research areas were Information and Computing Sciences, Information Systems, and Library and Information Studies.

**Conclusions.** The multidisciplinary in research data management was demonstrated by publications occurring in different fields of research, such as Information and Computing Sciences, Information Systems, Library and Information Studies, Medical and Health Sciences, and History and Archeology. About 60% of the publications had at least one citation, with a total of 3,598 citations found, featuring a growing academic impact.

**Originality/Value.** This bibliometric and altmetric study allowed the analysis of the literature on research data management. The theme was investigated in the Dimensions database and analyzed using productivity, impact, and online attention indicators.

**Keywords:** research data management, bibliometrics, altmetrics, scientific production

# 1 INTRODUÇÃO

A gestão de dados pode ser definida como uma disciplina que permite armazenar, utilizar e analisar dados, por meio da implementação de políticas, estratégias, procedimentos e práticas que mantenham a sua integridade, padronização e organização (Rodrigues, Nóbrega & Dias, 2017), envolvendo o desenvolvimento de diretrizes, planos e processos que gerenciem os dados mantendo sua segurança e usabilidade (Specht *et al.*, 2015).

Tal gestão pode também ser compreendida como uma cadeia de valor necessária à gestão de grandes volumes de dados: aquisição dos dados, análise dos dados, curadoria dos dados, armazenamento dos dados e uso dos dados (Di Martino *et al.*, 2014). No contexto acadêmico e em atividades de pesquisa, essa cadeia pode ser entendida como um conjunto de fases necessárias à gestão de dados científicos. Seu objetivo é permitir que os dados de pesquisa sejam autodescritivos, para que possam ser efetivamente reutilizados quando forem descobertos (Specht *et al.*, 2015).

Segundo Hey e Trefethen (2013), para lidar com o grande volume, ou mesmo dilúvio de dados, a infraestrutura científica precisará se adequar aos mecanismos de produção, compartilhamento, curadoria, preservação e reuso dos dados científicos. Os autores caracterizam o processo de automação da gestão de dados, automatizando-se a descoberta científica, com dados gerados por meio de simulações, experimentos e sensores, para posterior visualização e análise com apoio de intensa infraestrutura tecnológica.

No paradigma da ciência orientada a dados, vem se tornando comum a exigência do depósito de conjuntos de dados coletados antes mesmo da publicação dos artigos, assim como a exigência por algumas agências de fomento do Plano de Gestão de Dados para que um projeto de pesquisa seja aprovado. Nessa perspectiva, e devido ao crescente volume de estudos sobre gestão de dados científicos, esta pesquisa objetiva analisar a produção científica sobre gestão de dados científicos indexada na base de dados *Dimensions*, por meio de indicadores bibliométricos e altmétricos.

Pelo uso dos indicadores bibliométricos, é possível mapear a produção científica dos diversos campos do conhecimento, averiguando suas características e seu impacto acadêmico. Permite-se, também, que se atribua visibilidade aos autores, periódicos científicos, organizações e países, ao utilizar os indicadores como índices de acompanhamento do desempenho da literatura científica (Vanti & Sanz-Casado, 2016). O indicador de citação, por exemplo, é o mais recomendado para analisar o impacto acadêmico de pesquisas (Freitas, Rosas & Miguel, 2017), ao quantificar dados de citações de autores, títulos, origens geográficas, ano e idioma da publicação (Foresti, 1990).

Em complementaridade à bibliometria, a altmetria tem como objetivo aferir a atenção *on-line*, o impacto e a visibilidade das produções científicas em ambiente virtual, indo além do âmbito acadêmico (Freitas, Rosas & Miguel, 2017; Peters *et al.*, 2016). Os indicadores altmétricos permitem coletar informações sobre a produção científica como menções em postagens, comentários, *downloads*, *links*, entre outros, nas variadas fontes virtuais, como: mídias sociais, sites de redes sociais e de redes sociais acadêmicas, *blogs*, portais de notícias e gerenciadores de referências, por meio de programas, plataformas ou sistemas apropriados para a coleta desses dados (Barros, 2015; Costas, Zahedi & Wouters, 2015; Priem, Piwowar & Hemminger, 2012).

## 2 MATERIAIS E MÉTODOS

Trata-se de uma pesquisa de caráter descritivo e exploratório, com abordagem quantitativa, que tem como objetivo analisar a produção científica sobre gestão de dados científicos indexada na base de dados *Dimensions*, por meio de indicadores bibliométricos e altmétricos. Justificou-se a opção por essa base por sua ampla cobertura de dados internacionais, ao abranger diferentes tipos de fontes de publicações e uma alta cobertura de dados de diferentes campos de pesquisa e regiões, disponibilizando um grande volume de dados para análise (Bode *et al.*, 2019).

A pesquisa foi realizada em três etapas: busca na base de dados *Dimensions*; obtenção dos registros bibliométricos e altmétricos, e análise dos registros. Na primeira etapa, buscou-se pela palavra-chave “*research data management*”, nos campos de título e resumo, e foi aplicado o filtro por artigos científicos. Não foi realizado recorte temporal, abrangendo todos os anos de publicação disponíveis na *Dimensions*, desde 1970 até agosto de 2021 (período de realização da busca bibliográfica). Foram obtidos 677 artigos sobre gestão de dados científicos, que compuseram a população da pesquisa.

Na segunda etapa, os dados dos artigos científicos recuperados pela *Dimensions* foram classificados por quantidade de citação e exportados no formato .xlsx como registro completo, de modo a terem o *Digital Object Identifier* (DOI) selecionado e inserido no sistema *Altmetric* para análise dos dados de atenção *on-line*. Os registros bibliométricos também foram exportados no formato .csv para uso no *software VOSviewer*, na geração de mapas bibliométricos de redes de acoplamento bibliográfico de periódicos científicos e de coocorrência de termos nos títulos dos artigos científicos. Tais mapas permitem visualizar quais periódicos científicos tratam sobre temáticas conceitualmente aproximadas e os termos com maiores aparições nos títulos com relações entre si.

O sistema *Altmetric* foi utilizado para o levantamento dos dados altmétricos da produção científica, por meio da inserção do DOI de cada artigo no *Altmetric Explorer*. Esse sistema é considerado um dos mais importantes provedores desse tipo de dados no mundo (Robinson-García *et al.*, 2014), oportuno para o fornecimento e o fácil acesso aos dados altmétricos, ao oferecer um estudo sistemático de seus significados (Haustein *et al.*, 2013). É utilizado em estudos que buscam rastrear a atenção que os resultados da pesquisa recebem *on-line*, ao coletar dados de fontes virtuais, como mídias sociais, políticas e patentes, portais de notícias e *blogs*, fontes acadêmicas e outros tipos, como Wikipédia e vídeos (Adie & Roe, 2013; Digital Science & Research Solutions, 2021).

Na terceira etapa, a produção científica foi analisada segundo indicadores bibliométricos de produtividade e de citação, e altmétricos. Os indicadores bibliométricos foram: ano de publicação, tipo de acesso (aberto: via verde, dourada, bronze e híbrida, e fechado), instituições de pesquisa e países/regiões, autores, periódicos científicos e campos de pesquisa mais produtores, e coocorrência de termos nos títulos dos artigos. Para os indicadores de citação, tem-se: total de citações, distribuição de citações por ano e artigos mais citados. E os indicadores altmétricos foram: total de menções, distribuição por tipos de fontes virtuais (mídias sociais, notícias e *blogs*, políticas e patentes, fontes acadêmicas e outras fontes), atenção *on-line* de acordo com os países/regiões e as fontes utilizadas (*Twitter*, *Facebook*, notícias e políticas), periódicos científicos que apresentaram maior quantidade de menções nos artigos publicados e os artigos com mais menções.

Dentre os estudos bibliométricos correlatos estão: Guimarães e Bezerra (2019) que utilizaram o termo “*data management*”; Zhang e Eichmann-Kalwara (2019), Setyowati e

Heriyanto (2019) e Gupta, Arora e Chakravarty (2021) com “*research data management*” e, Peng e Ye (2021) com os termos “*data collection*” e “*data management*”. Também se encontraram estudos correlatos sobre o uso do sistema *Altmetric* para análises de dados altmétricos (Ortega, 2018; Robinson-García *et al.*, 2014; Tattersall & Carroll, 2018) e o uso do *VOSviewer* em análises bibliométricas (Codato, 2018; Lima & Leocádio, 2018; Palludeto & Felipini, 2019; Sousa *et al.*, 2019).

### 3 RESULTADOS E DISCUSSÕES

De acordo com os dados da pesquisa houve variação no aumento das publicações ao longo dos anos (1970 a 2021). Entre 1970 e 2011 manteve-se uma variação de uma a sete publicações, enquanto em 2012 houve um salto em comparação aos anos anteriores, com 22 publicações. A partir de 2012 houve grande crescimento da literatura sobre gestão de dados científicos, tendo como ápice das produções os anos de 2019 e 2020 (FIGURA 1).

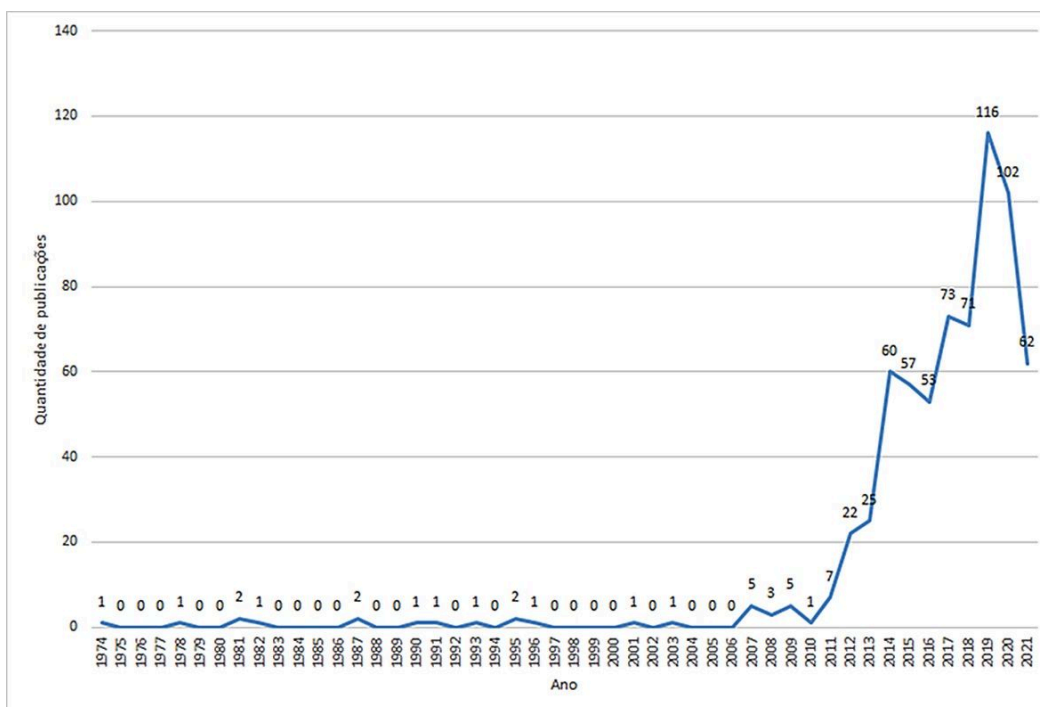


Figura 1. Evolução da produção sobre gestão de dados científicos (Fonte: dados da pesquisa, 2021).

O aumento nas publicações sobre o tema já havia sido evidenciado nos estudos bibliométricos correlatos. Guimarães e Bezerra (2019) demonstraram o aumento das publicações sobre gestão de dados em bases multidisciplinares no período de 1963 a 2018, havendo o auge das publicações em 2017. Zhang e Eichmann-Kalwara (2019), ao realizarem estudo bibliométrico na *Scopus* sem recorte temporal, recuperaram aproximadamente dois mil registros sobre o tema, tendo 90% da produção tendo sido publicada a partir de 2007, o que demonstra um crescimento recente nas publicações. Setyowati e Heriyanto (2019) apontaram o crescimento exponencial nas publicações sobre a temática nos últimos dez anos (2010-2019) na *Scopus*, enquanto Gupta, Arora e Chakravarty (2021) identificaram ápices na produção entre 1989 e 2021 na coleção principal da *Web of Science*, nos anos de 2018 e 2020 com 72 artigos cada.

Esses artigos estão disponibilizados conforme o tipo de acesso ao documento: aberto ou fechado. O acesso aberto, pode ocorrer pela via verde que tem como objetivo armazenar cópias do texto completo de publicações em repositórios digitais ou, pela via dourada que se refere à publicação de artigos científicos em periódicos de acesso aberto (Costa & Leite, 2015). Piowar *et al.* (2018) também se referem às vias bronze e híbrida, em que a primeira permite o livre acesso para leitura no site do próprio editor, mas sem licença e direitos de reuso, como *downloads*, e a segunda de acesso livre com licença de reuso em periódicos que não são considerados de acesso livre. O acesso fechado inclui todas as publicações não contempladas pelo acesso aberto. A distribuição quanto ao tipo de acesso dos 677 artigos divide-se em: 510 (75,33%) de acesso aberto, sendo 352 (69,01%) disponibilizados pela via dourada, 55 pela bronze (10,78%), 54 pela híbrida (10,58%) e 49 pela via verde (9,60%), e 167 (24,66%) publicações em acesso fechado.

As instituições e os países/regiões mais producentes na temática, com corte pelos 10 primeiros colocados podem ser visualizados na Tabela 1.

Organizações	Nº artigos	(%)	Países/regiões	Nº artigos	(%)
<i>University of Sheffield</i>	12	1,77	Estados Unidos	133	19,64
<i>University of Göttingen</i>	9	1,32	Alemanha	91	13,44
<i>Oregon State University</i>	7	1,03	Reino Unido	53	7,82
<i>Stanford University</i>	6	0,88	Austrália	22	3,24
<i>University Medical Center Göttingen</i>	6	0,88	Canadá	21	3,10
<i>University of Cologne</i>	6	0,88	Holanda	18	2,65
<i>RWTH Aachen University</i>	5	0,73	China	16	2,36
<i>University of Michigan</i>	5	0,73	África do Sul	11	1,62
<i>University of Washington</i>	5	0,73	Suíça	9	1,32
<i>Delft University of Technology</i>	5	0,73	Indonésia	9	1,32

**Tabela 1.** Organizações e países/regiões com maiores quantidades de artigos publicados (Fonte: dados da pesquisa, 2021).

Esses resultados se assemelham com os estudos correlatos, como o de Costa e Cunha (2015), que utilizaram o termo “e-Science” e identificaram que a maioria dos artigos científicos são produzidos por instituições dos Estados Unidos, seguidas por instituições do Reino Unido e da Alemanha. Gupta, Arora e Chakravarty (2021) identificaram que os países mais produtivos sobre *research data management* são Estados Unidos, seguidos por Inglaterra e Alemanha, havendo destaque para as instituições *University of Sheffield*, *University of Pittsburgh* e *University of Oxford*. Peng e Ye (2021), verificaram que as produções sobre “data collection” e “data management” no contexto de mídias sociais e do *Big Data* na *Web of Science* no período de 2010 a 2019, foram lideradas pelos Estados Unidos, seguidos pela China e Reino Unido, tendo como universidades mais produtivas: *University of Michigan*, *Chinese Academy of Sciences* e *University of Sydney*.

Para o indicador de autores, levantou-se um total de 1.288, destes apenas 10 (0,77%) publicaram cinco ou mais artigos (TABELA 2), tendo mais de 90% dos autores publicado entre um e quatro artigos, sendo que 24 (1,86%) publicaram quatro, 35 (2,71%) publicaram três, 169 (13,12%) dois e 1.050 (81,52%) um artigo científico. Esse resultado foi semelhante ao de Guimarães e Bezerra (2019), corroborando a lei de Lotka, ao estabelecer que poucos autores produzem muito e muitos autores produzem pouco (Costa & Cunha, 2015; Guedes

& Borschiver, 2005).

<b>Autores</b>	<b>Organização/país</b>	<b>Artigos (%)</b>	
Andrew Martin Cox	<i>University of Sheffield</i> , Reino Unido	13	1,92
Stephen Pinfield	<i>University of Sheffield</i> , Reino Unido	7	1,03
Marta Teperek	<i>Delft University of Technology</i> , Holanda	7	1,03
Mary Anne Kennan	<i>Charles Sturt University</i> , Austrália	6	0,88
Amanda L. Whitmire	<i>Stanford University</i> , Estados Unidos	6	0,88
Georg Bareth	<i>University of Cologne</i> , Alemanha	5	0,73
Kevin B. Read	<i>New York University</i> , Estados Unidos	5	0,73
Martin Donnelly	<i>University of Edinburgh</i> , Reino Unido	5	0,73
Oliver Koepler	<i>German National Library of Science and Technology</i> , Alemanha	5	0,73
Elaine Russo Martin	<i>Harvard University</i> , Estados Unidos	5	0,73

**Tabela 2.** Autores com maiores quantidades de artigos publicados (Fonte: dados da pesquisa, 2021)

Percebe-se que com exceção do terceiro colocado, Marta Teperek da Holanda, e do quarto colocado, Mary Anne Kennan da Austrália, os demais autores que mais produziram são dos países/regiões com maiores destaques na temática (Estados Unidos, Alemanha e Reino Unido).

Levantou-se, também, o indicador de periódicos científicos, sendo os com mais publicações sobre o tema, com nove ou mais documentos, exibidos na Tabela 3.

<b>Periódicos científicos</b>	<b>Nº artigos</b>	<b>(%)</b>
<i>International Journal of Digital Curation</i>	74	10,93
<i>Journal of eScience Librarianship</i>	38	5,61
<i>Data Science Journal</i>	25	3,69
<i>Septentrio Conference Series</i>	15	2,21
<i>Bulletin of the Association for Information Science and Technology</i>	13	1,92
<i>IFLA Journal</i>	13	1,92
<i>BIBLIOTHEK Forschung und Praxis</i>	12	1,77
<i>Procedia Computer Science</i>	10	1,47
<i>Journal of Librarianship and Scholarly Communication</i>	10	1,47
<i>IASSIST Quarterly</i>	9	1,32
<i>The Journal of Academic Librarianship</i>	9	1,32
<i>LIBER Quarterly</i>	9	1,32
<i>Proceedings of the Association for Information Science and Technology</i>	9	1,32
<i>Research Ideas and Outcomes</i>	9	1,32

**Tabela 3.** Periódicos científicos com maiores quantidades de artigos publicados (Fonte: dados da pesquisa, 2021).

Destacam-se os periódicos *International Journal of Digital Curation*, que aborda aspectos relacionados a ferramentas de gestão e curadoria de dados digitais com 74 publicações, *Journal of eScience Librarianship* que traz a teoria e prática da Biblioteconomia em conformidade com pesquisas sobre dados com 38 artigos publicados, e com 25 publicações, *Data Science Journal*, que aborda temas relacionados com a gestão, disseminação, uso e reutilização de dados de pesquisa e banco de dados. Zhang e Eichmann-Kalwara (2019), encontraram mais publicações na *PLoS One*, seguida pela *Nucleic Acids Research* e pela *Nature*. No top 10 dos periódicos elencados por Setyowati e Heriyanto (2019), os mais produtivos foram: *Voeb-Mitteilungen*, *Communications In Computer And Information Science* e *Ifla Journal*. Gupta, Arora e Chakravarty (2021) tiveram como fontes mais produtivas *Zeitschrift fur Bibliothekswesen und Bibliographie*, *PLoS One* e *Journal of the American Medical Informatics Association* e, Peng e Ye (2021) tiveram: *IEEE Access*, *Future Generation Computer Systems: The International Journal of eScience* e *Journal of Medical Internet Research*. Todas essas pesquisas apontam para a multidisciplinaridade presente na gestão de dados científicos, com a ocorrência de publicações em periódicos de distintas áreas do conhecimento.

A rede de acoplamento bibliográfico de periódicos científicos (FIGURA 2), gerada no *VOSviewer*, é composta pelos 26 periódicos que apresentaram cinco ou mais artigos publicados com referências em comum. Nessa categoria procura-se “[...] aproximar unidades que compartilham referências em comum. No caso, quanto maior a coincidência de determinados veículos de publicação citados em comum, maior a proximidade entre as fontes” (Palludeto & Felipini, 2019, p. 322).

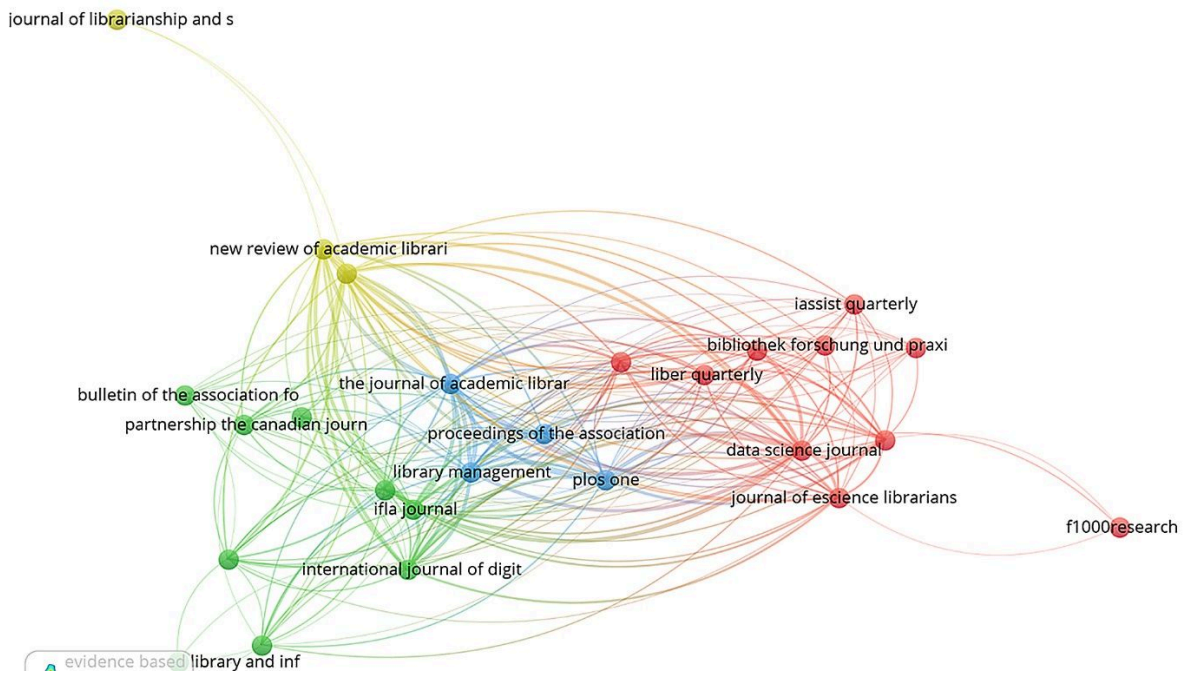


Figura 2. Rede de acoplamento bibliográfico de periódicos científicos (Fonte: dados da pesquisa, 2021).

Nota-se uma grande interação entre os *clusters* com proximidade, inferindo-se que os periódicos apresentados possuem referências compartilhadas entre si. Também sobre esse tipo de mapa, considera-se que “[...] publicações ‘acopladas’ tratam temáticas conceitualmente próximas, indicando núcleos referenciais de pesquisas” (Lima & Leocádio, 2018, p. 200), nesse sentido, percebe-se que os principais temas dos periódicos científicos são relacionados com as práticas da Biblioteconomia e da Ciência da Informação, presente



em todos os *clusters*, bem como periódicos que abordam a Ciência de Dados (*Data Science Journal*, *cluster* vermelho), Ciência e Medicina (*PLoS One*, *cluster* azul) e curadoria digital (*International Journal of Digital Curation*, *cluster* verde).

Para os campos de pesquisa (denominados pela base de dados *Dimensions*), foram levantados os 15 que mais se destacaram neste cenário com mais de 12 publicações (TABELA 4).

Campos de pesquisa	Nº artigos	(%)
<i>Information and Computing Sciences</i>	483	71,34
<i>Information Systems</i>	362	53,47
<i>Library and Information Studies</i>	113	16,69
<i>Medical and Health Sciences</i>	53	7,82
<i>History and Archaeology</i>	35	5,16
<i>Public Health and Health Services</i>	35	5,16
<i>Curatorial and Related Studies</i>	33	4,87
<i>Education</i>	29	4,28
<i>Studies in Human Society</i>	27	3,98
<i>Artificial Intelligence and Image Processing</i>	22	3,24
<i>Engineering</i>	16	2,36
<i>Specialist Studies In Education</i>	16	2,36
<i>Chemical Sciences</i>	15	2,21
<i>Biological Sciences</i>	12	1,77
<i>Sociology</i>	12	1,77

**Tabela 4.** Campos de pesquisa com maiores quantidades de artigos publicados (Fonte: dados da pesquisa, 2021).

Em comparação com os resultados de trabalhos correlatos, o estudo bibliométrico de Zhang e Eichmann-Kalwara (2019) identificou que os campos com mais publicações foram *Science*, *Physiology* e *Biology*. Enquanto no de Peng e Ye (2021) foram *Engineering*, *Computer Systems* e *Medical Research*. Os resultados sobre os periódicos científicos e os campos de pesquisa apontam para a multidisciplinaridade do tema gestão de dados científicos, em que diferentes campos o abordam. Além dos exemplos de pesquisa já mencionados, é possível verificar estudos em Química (Chen *et al.*, 2019), Ecologia (Zhang *et al.*, 2020), Ciência da Computação (Yang *et al.*, 2012) e também na Ciência da Informação (Sayão & Sales, 2012; Sayão & Sales, 2019).

Também foi gerado o mapa bibliométrico da rede de coocorrência de palavras retiradas dos títulos dos artigos (FIGURA 3), que permite compreender as temáticas com relações dentro do contexto da gestão de dados científicos, as áreas de estudo mais relevantes e possíveis assuntos de interesse (tendências de pesquisas). O critério de corte foram os termos que apresentaram no mínimo oito ocorrências nos títulos dos artigos, totalizando 28 termos. Nesse mapa o tamanho dos nós indica a quantidade de ocorrências que o termo obteve (Lima & Leocádio, 2018; Van Eck & Waltman, 2020).

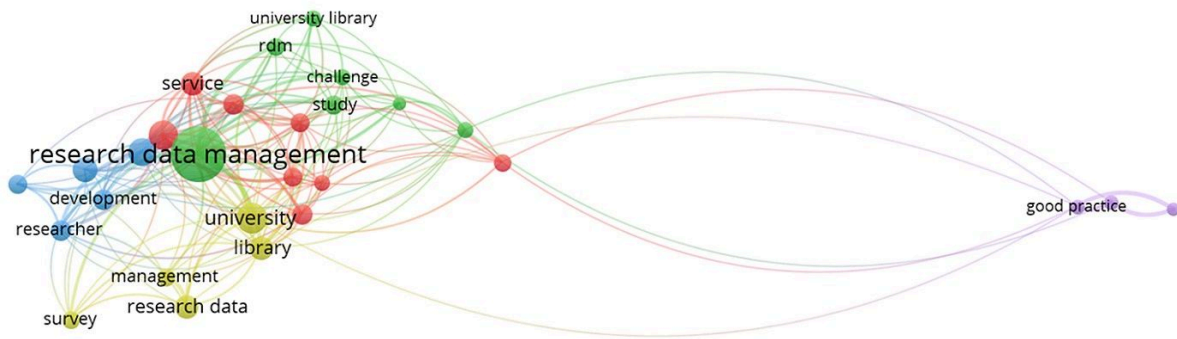


Figura 3. Rede de coocorrência de termos nos títulos dos artigos científicos (Fonte: dados da pesquisa, 2021).

Há presença de cinco *clusters* com ligações entre si e próximos, indicando grande interação entre os termos que ocorreram simultaneamente nos títulos dos artigos. O termo *research data management* ocorreu 181 vezes, sendo o mesmo termo utilizado para a busca bibliográfica na *Dimensions*, seguido pelo termo *university* com 51 ocorrências e *data* com 46. Pesquisas correlatas apontam resultados relacionados a estes, como a de Guimarães e Bezerra (2019) que ao buscar pelo termo “*data management*”, no período de 1963 a 2018, nas bases *Web of Science*, *Scopus*, *ScienceDirect* e *Emerald*, teve como palavras-chave mais indexadas *data management*, *information management* e *human*, seguidas por termos relacionados à gestão da informação, gestão de dados e às tecnologias da informação. Zhang e Eichmann-Kalwara (2019) identificaram: curadoria digital, processamento informacional, biblioteca digital, *Big Data*, compartilhamento de dados, metadados, aquisição de dados e preservação de dados. Enquanto Setyowati e Heriyanto (2019) trouxeram três novos termos que indicam tendências de pesquisa: *data science*, *digital humanities* e *open access*. Gupta, Arora e Chakravarty (2021) mapearam seis temáticas principais: *Data Management Plan*, *Data Management*, *Data Literacy and Lifecycle approach*, privacidade e qualidade de dados, *data science and literacy* e preservação de dados e opções de compartilhamento.

Foram levantados, também, dados sobre as citações do universo tratado. A citação é um importante indicador bibliométrico de atividade científica, que permite quantificar informações registradas para posterior análise, retirando informações de partes das publicações, como autores, títulos, origens geográficas, anos e idiomas das publicações (Foresti, 1990; Vanz & Caregnato, 2003). Dos 677 artigos que compõem o universo da pesquisa, 415 (61,29 %) receberam uma ou mais citações, havendo um total de 3.608 citações ao longo dos anos pesquisados. Entre 1970 e 2011, a maior quantidade de citações feitas por ano foi cinco, enquanto alguns anos não apresentaram nenhuma. A quantidade de citações aumentou a partir de 2012 com 49 citações, 2015 com 189, 2020 com 605 e 2021 com 660 (até agosto deste ano). O aumento no número de citações ocorreu à medida que houve o crescimento na quantidade de publicações, como também apontado por Guimarães e Bezerra (2019) e Cunha e Costa (2020).

Entre os 10 artigos mais citados, apresentados no Quadro 1, têm-se:

<b>Autores</b>	<b>Título do artigo</b>	<b>Citações</b>
Lowe <i>et al.</i> (2009)	STRIDE--An integrated standards-based translational research informatics platform	306
Sharma <i>et al.</i> (2014)	Panorama: A Targeted Proteomics Knowledge Base	144
Corrall, Kennan e Afzal (2013)	Bibliometrics and Research data management Services: Emerging Trends in Library Support for Research	104
Tenopir <i>et al.</i> (2014)	Research data management services in academic research libraries and perceptions of librarians	95
Cox e Pinfield (2013)	Research data management and libraries: Current activities and future priorities	94
Anderson <i>et al.</i> (2007)	Issues in Biomedical Research data management and Analysis: Needs and Barriers	89
Skripcak <i>et al.</i> (2014)	Creating a data exchange strategy for radiotherapy research: Towards federated databases and anonymised public datasets	76
Cox <i>et al.</i> (2017)	Developments in research data management in academic libraries: Towards an understanding of research data service maturity	66
Arend <i>et al.</i> (2016)	PGP repository: a plant phenomics and genomics data publication infrastructure	61
Arend <i>et al.</i> (2014)	e!DAL - a framework to store, share and publish research data	57

**Quadro 1.** Artigos científicos mais citados (Fonte: dados da pesquisa, 2021).

Percebe-se que os três primeiros colocados apresentam quantidades de citações com elevada diferença em relação ao restante. As temáticas dos artigos mais citados variam de plataformas, *frameworks* e repositórios que permitam o armazenamento, compartilhamento e reuso de dados até barreiras e serviços relacionados a gestão de dados científicos.

Em complementaridade às técnicas bibliométricas e ao indicador de citação, as análises altmétricas permitem analisar um universo mais amplo do que apenas o voltado à academia, ao medir a atenção *on-line* que a produção científica recebe (Costas, Zahedi & Wouters, 2015; Haustein *et al.*, 2013), que pode ser observada por meio das menções recebidas em ambiente virtual. Observou-se que dos 677 artigos recuperados, 666 possuem DOI (98,37%), desse total, 376 (55,53%) apresentaram algum tipo de atenção *on-line* de acordo com a consulta no sistema *Altmetric*, com pelo menos uma menção recebida. Assim, foram encontradas 6.027 menções distribuídas de acordo com as variadas fontes virtuais, apresentadas na Tabela 5.

<b>Tipo de fonte</b>	<b>Nº de menções</b>	<b>(%)</b>
Mídias sociais	5.686	94,34
Notícias e blogs	271	4,49
Políticas e patentes	48	0,79
Fontes acadêmicas	12	0,19
Outras fontes	10	0,16
	6.027	100

**Tabela 5.** Distribuição de menções pelas fontes virtuais (Fonte: dados da pesquisa, 2021).

Foi possível averiguar que as mídias sociais são as fontes nas quais os artigos sobre gestão de dados científicos mais apresentaram atenção *on-line*, sendo predominante o *Twitter* com 5.411 menções (89,77%), seguido pelo *Facebook* com 230 (3,81%). Esses resultados corroboram com os estudos de Haustein *et al.* (2013), Costas, Zahedi e Wouters (2015) e Araujo e Furnival (2016), que apontaram a preferência pelo *Twitter* no compartilhamento e divulgação de publicações, o que pode ser explicado pelo fato dessa rede social fornecer dados das postagens de todos os perfis abertos para o público em geral, enquanto o *Facebook* fornece apenas as mensagens postadas em páginas públicas, excluindo as páginas pessoais dos usuários.

Também foram levantados dados de atenção *on-line* das publicações de acordo com os países e regiões e as fontes utilizadas (*Twitter* por meio dos *tweets*, *Facebook* por meio das postagens, notícias e documentos de política). A Tabela 6 apresenta os 15 países com maior quantidade de postagens.

<b>Países/ regiões</b>	<b>Nº de postagens</b>	<b>Nº de <i>tweets</i></b>	<b>Nº de postagens do <i>Facebook</i></b>	<b>Nº de notícias</b>	<b>Nº de políticas</b>
EUA	1.207	1.160	40	6	1
Reino Unido	501	495	3	3	-
Alemanha	380	370	3	7	-
França	270	270	-	-	-
Austrália	195	152	-	1	42
Holanda	194	193	-	1	-
Canadá	186	185	-	-	1
Irlanda	172	172	-	-	-
Áustria	84	84	-	-	-
Espanha	73	69	1	3	-
África do Sul	52	43	9	-	-
Noruega	46	43	1	-	2
Bélgica	45	45	-	-	-
Japão	36	34	1	1	-
Suécia	34	34	-	-	-

**Tabela 6.** Países/regiões com maior quantidade de publicações com atenção *on-line*. (Fonte:

dados da pesquisa, 2021).

Os países com maiores índices de engajamento e atenção *on-line* das publicações são, em sua maioria, os mesmos que mais publicam sobre gestão de dados, como os EUA, Alemanha, Reino Unido, Austrália, Canadá, Holanda e África do Sul, o que pode ser um indicativo de que os próprios autores, assim como os periódicos científicos dos países que mais publicam e as instituições vinculadas a estes autores estão divulgando seus artigos com maior frequência nas fontes virtuais. A China não aparece entre esses resultados, devido a proibição do uso de redes sociais estrangeiras a sua população. Enquanto o Brasil aparece em 25º com 14 postagens, sendo todas por meio de *tweets*. Nota-se, também, que houve grande dispersão entre as postagens do *Facebook*, notícias e políticas, uma vez que poucos países apresentaram pouco engajamento nessas fontes, com destaque para os EUA com 40 postagens no *Facebook*, o maior índice, e a Austrália com 42 documentos de política.

Os 15 periódicos científicos que apresentaram mais menções de seus artigos são listados na Tabela 7.

Periódicos científicos	Nº de menções	Nº de publicações
<i>Journal of eScience Librarianship</i>	1.049	34
<i>Research Ideas and Outcomes</i>	762	9
<i>Data Science Journal</i>	635	24
<i>New Review of Academic Librarianship</i>	314	7
<i>Insights: the UKSG journal</i>	287	5
<i>PLOS ONE</i>	197	5
<i>Journal of Librarianship and Scholarly Communication</i>	189	10
<i>Liber Quarterly: The Journal of European Research Libraries</i>	167	8
<i>IFLA Journal</i>	109	12
<i>Giga Science</i>	99	4
<i>FACETS</i>	95	1
<i>Records Management Journal</i>	86	3
<i>Journal of the Association for Information Science and Technology</i>	80	3
<i>Journal of Librarianship &amp; Information Science</i>	75	4
<i>Advances in Archaeological Practice</i>	64	2

**Tabela 7.** Periódicos científicos com mais menções em suas publicações (Fonte: dados da pesquisa, 2021).

Nota-se que, à exceção do primeiro e terceiro colocados, poucos artigos publicados nesses periódicos apresentaram muitas menções, como exemplo, o segundo colocado, *Research Ideas and Outcomes*, que conta com nove publicações sendo mencionadas nas fontes virtuais 762 vezes, bem como o periódico *FACETS*, que conta com apenas um artigo publicado recebendo 95 menções. Além dos periódicos que mais publicaram sobre o tema já mencionados, também se destacam como periódicos com mais menções em suas publicações sobre gestão de dados: *Research Ideas and Outcomes*, que contempla todas as áreas de pesquisa, como ciência e tecnologia, humanidades e ciências sociais; *Insights: the UKSG journal*, que abrange publicações e notícias sobre o setor de informação acadêmica;

*Giga Science* que cobre publicações sobre ciências biomédicas e da vida, e *Advances in Archaeological Practice*, que trata sobre os desafios da prática da Arqueologia mundialmente. A diversidade de campos do conhecimento realizando menções a artigos sobre a temática aponta novamente para a multidisciplinaridade que a permeia, conforme evidenciado também por Peng & Ye (2021) e Zhang e Eichmann-Kalwara (2019).

Os 10 artigos com mais visibilidade e atenção *on-line* indicados pelo número de menções nas fontes virtuais foram:

<b>Autores</b>	<b>Título</b>	<b>Periódico</b>	<b>Ano</b>	<b>Nº de menções</b>
Higman, Rosie; Bangert, Daniel; Jones, Sarah	Three camps, one destination: the intersections of research data management, FAIR and Open	Insights: the UKSG journal	2019	195
Briney, Kristin; Coates, Heather; Goben, Abigail	Foundational Practices of Research Data Management	Research Ideas and Outcomes	2020	187
Altenhön, Reinhard <i>et al.</i>	NFDI4Culture - Consortium for research data on material and immaterial cultural heritage	Research Ideas and Outcomes	2020	165
Cox, John	Positioning the Academic Library within the Institution: A Literature Review	New Review of Academic Librarianship	2018	156
Briney, Kristin	The Problem with Dates: Applying ISO 8601 to Research Data Management	Journal of eScience Librarianship	2018	139
Borghi, John <i>et al.</i>	Support Your Data: A Research Data Management Guide for Researchers	Research Ideas and Outcomes	2018	127
Soyka, Heather <i>et al.</i>	Using Peer Review to Support Development of Community Resources for Research Data Management	Journal of eScience Librarianship	2017	106
Cruz, Maria <i>et al.</i>	Policy Needs to Go Hand in Hand with Practice: The Learning and Listening Approach to Data Management	Data Science Journal	2019	101
Pinfield, Stephen; Cox, Andrew M.; Smith, Jen	Research Data Management and Libraries: Relationships, Activities, Drivers and Influences	PLOS ONE	2014	100
Sewell, Claire; Kingsley, Danny	Developing the 21st Century Academic Librarian: The Research Support Ambassador Programme	New Review of Academic Librarianship	2017	99

**Tabela 8.** Artigos com maiores números de menções (**Fonte:** dados da pesquisa (2021)).

Os artigos com maior número de menções apresentam temáticas relacionadas aos princípios FAIR, possibilidades de serviços que bibliotecários e bibliotecas podem oferecer, além do uso de políticas, normas, guias e melhores práticas no contexto da gestão de dados científicos.

## 4 CONSIDERAÇÕES FINAIS

Objetivou-se analisar a produção científica sobre gestão de dados científicos indexada na *Dimensions* a partir de indicadores bibliométricos e altmétricos, utilizando como termo de busca “*research data management*”. A partir da análise das 677 pesquisas recuperadas, observou-se um crescimento da produção científica ao longo dos anos, com destaque a partir de 2012 e pico de publicações em 2019 e 2020 (116 e 102 artigos respectivamente). E que do total de artigos recuperados, 510 estão em acesso aberto (352 na via dourada, 55 na bronze, 54 na híbrida e 49 na verde), e 167 estão em acesso fechado. Os países/regiões com maiores índices de publicações foram os EUA (133 artigos), Alemanha (91) e Reino Unido (53).

Entre os autores mais produtivos se destacam Andrew Martin Cox (13 artigos), Stephen Pinfield e Marta Teperek (sete cada) e Mary Anne Kennan e Amanda L. Whitmire (seis cada). O universo da pesquisa teve 1.288 autores, destes 1.050 tiveram apenas uma publicação, estando em conformidade com a Lei de Lotka, em que poucos autores publicam muito e muitos autores publicam pouco. Os periódicos científicos que mais apresentaram artigos publicados foram *International Journal of Digital Curation*, *Journal of eScience Librarianship* e *Data Science Journal*, com 74, 38 e 25 publicações respectivamente, e entre as áreas de pesquisa: *Information and Computing Sciences* com 483, *Information Systems* com 362 e *Library and Information Studies* com 113 artigos. A multidisciplinaridade em gestão de dados científicos foi demonstrada pela variedade de áreas de periódicos científicos e campos do conhecimento que publicam sobre a temática.

Os mapas bibliométricos gerados pelo *VOSviewer* permitiram analisar redes de acoplamento bibliográfico de periódicos científicos e de coocorrência de termos nos títulos dos artigos. Neste último, notou-se que as temáticas com maiores proximidades e ocorrências na produção sobre gestão de dados científicos, além do termo “*research data management*”, foram “*university*” e “*data*”.

Foram encontradas, também, 3.608 citações distribuídas em 415 artigos, havendo aumento do impacto na produção científica sobre gestão de dados científicos em âmbito acadêmico, uma vez que tem sido crescente sua quantidade de citações, principalmente a partir de 2012. Como artigo mais citado do universo da pesquisa, tem-se I-Min A. Chen *et al* (2019) *IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes* com 214 citações, publicado no periódico *Nucleic Acids Research*.

Dos 677 artigos, 376 apresentaram 6.027 menções em sua totalidade, encontradas pela consulta ao sistema *Altmetric*, que permitiu averiguar a atenção *on-line* da produção científica analisada. As fontes com maiores índices de atenção foram as mídias sociais, entre elas o *Twitter* e o *Facebook*. Entre os países/regiões que mais apresentaram engajamento de artigos com atenção *on-line*, destacam-se: EUA, Reino Unido e Alemanha. E entre os periódicos científicos com maiores quantidades de menções em seus artigos destacam-se *Journal of eScience Librarianship*, *Research Ideas and Outcomes* e *Data Science Journal*, confirmando a multidisciplinaridade da gestão de dados científicos. O artigo com o maior

número de menções nas fontes virtuais foi o de Rosie Higman, Daniel Bangert e Sarah Jones (2019) *Three camps, one destination: the intersections of research data management, FAIR and Open* com 195 menções.

Ao pensar na era do *Big Data*, a gestão de dados científicos é um tema em desenvolvimento para garantir o compartilhamento e reuso destes, e conseqüentemente, o avanço da ciência. Este estudo bibliométrico e altmétrico permitiu acompanhar o desempenho da literatura sobre gestão de dados científicos, ao delinear o perfil das produções científicas, o aumento das publicações ao longo dos anos, além da verificação da visibilidade gerada por pesquisadores, periódicos, organizações e regiões.

Uma limitação da pesquisa foi o uso exclusivo da base de dados *Dimensions* para coleta de dados bibliométricos, o que dificulta a realização de generalizações a respeito dos resultados encontrados. Como pesquisa futura, sugere-se a ampliação das bases utilizadas no estudo bibliométrico e altmétrico sobre gestão de dados científicos.

## **AGRADECIMENTOS**

A primeira autora agradece à Capes pelo apoio financeiro concedido no desenvolvimento de sua pesquisa de doutorado. Os autores agradecem à Dimensions pelo acesso e uso não comercial dos dados da produção científica e de citação utilizados no estudo. De igual modo, agradecem à Altmetric.com por fornecer os dados altmétricos deste estudo gratuitamente para fins de pesquisa.

## **DECLARAÇÃO DE CONTRIBUIÇÃO**

Conceptualização, curadoria de dados, análise formal, pesquisa, metodologia, administração do projeto, validação, visualização, escrita do rascunho inicial, escrita da revisão e edição: Marília Catarina Andrade Gontijo, Raíssa Yuri Hamanaka e Ronaldo Ferreira de Araujo.

Supervisão: Ronaldo Ferreira de Araujo.

## **CONFLITOS DE INTERESSE**

Os autores declaram que não há conflitos de interesse.

## **DECLARAÇÃO DE CONSENTIMENTO DE DADOS**

Os dados gerados durante este artigo foram incluídos no manuscrito.

## **REFERÊNCIAS**

Adie, E., & Roe, W. (2013). Altmetric: enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, 26(1), 11-17. <https://doi.org/10.1087/20130103>

Araujo, R. F. de, & Furnival, A. C. M. (2016). Comunicação científica e atenção online: em busca de colégios virtuais que sustentam métricas alternativas. *Informação & Informação*, 21(2), 68-89. <http://dx.doi.org/10.5433/1981-8920.2016v21n2p68>



- Barros, M. (2015). Altmetrics: métricas alternativas de impacto científico com base em redes sociais. *Perspectivas em Ciência da Informação*, 20(2), 19-37. <https://doi.org/10.1590/1981-5344/1782>
- Bode, C. et al. (2019). *A Guide to the Dimensions Data Approach*. Cambridge: Digital Science.
- Chen, I. A. et al. (2019). IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Research*, 47(1). <https://doi.org/10.1093/nar/gky901>
- Codato, A. (2018). Utilizando citações para além do fator de impacto: uma alternativa para determinar topografias científicas. *SciELO 20 Years Repository*, 1-19. Recuperado de <https://repository.scielo20.org/index.php/documents/article/view/119/81>
- Costa, M. M., & Cunha, M. B. (2015). A literatura internacional sobre e-Science nas bases de dados LISA e LISTA. *Encontros Bibli: revista eletrônica de Biblioteconomia e Ciência da Informação*, 20(44), 127-144. <https://doi.org/10.5007/1518-2924.2015v20n44p127>
- Costa, M., & Leite, F. C. L. (2015). Repositórios institucionais de acesso aberto à informação científica: proposta de modelo de avaliação. *RECIIS – Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, 9(3), 1-20. <https://doi.org/10.29397/reciis.v9i3.996>
- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “Altmetrics” Correlate With Citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(1), 2003-2019. <https://doi.org/10.1002/asi.23309>
- Cunha, M. B., & Costa, M. M. (2020). Fontes de informação sobre gestão de dados científicos. *Informação & Sociedade: Estudos*, 30(4), 1-59. <https://doi.org/10.22478/ufpb.1809-4783.2020v30n4.57183>
- Digital Science & Research Solutions. (2021). *Dimensions*. Recuperado de <https://app.dimensions.ai/discover/publication>
- Di Martino, B., Aversa, R., Cretella, G., Esposito, A., & Kołodziej, J. (2014). Big data (lost) in the cloud. *International Journal of Big Data Intelligence*, 1(1-2), 3-17.
- Foresti, N. A. B. (1990). Contribuição das revistas brasileiras de biblioteconomia e ciência da informação enquanto fonte de referência para a pesquisa. *Ciência da Informação*, 19(1), 53-71. Recuperado de <http://revista.ibict.br/ciinf/article/view/375/375>
- Freitas, J. L., Rosas, F. S., & Miguel, S. E. (2017). Estudos métricos da informação em periódicos do portal SciELO: visibilidade e impacto na Scopus e Web of Science. *Palavra Chave*, 6(2), 1-12. <https://doi.org/10.24215/PCe021>
- Guedes, V. L. S., & Borschiver, S. (2005). *Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica*. Artigo apresentado no 6º Encontro Nacional de Ciência da Informação, Salvador, Brasil. Recuperado de [http://www.cinform-antiores.ufba.br/vi\\_anais/docs/VaniaLSGuedes.pdf](http://www.cinform-antiores.ufba.br/vi_anais/docs/VaniaLSGuedes.pdf)

- Guimarães, A. J. R., & Bezerra, C. A. (2019). Gestão de dados: uma abordagem bibliométrica. *Perspectivas em Ciência da Informação*, 24(4), 171-186. <https://doi.org/10.1590/1981-5344/4192>
- Gupta, N., Arora, S., & Chakravarty, R. (2021). Science Mapping and Visualization of Research Data Management (RDM): Bibliometric and Scientometric Study. *Library Philosophy and Practice*, 1-23.
- Hey, T., & Trefethen, A. (2003). The data deluge: an e-science perspective. In Berman, F., Fox, G., & Hey, A. J. G. (Eds.), *Grid computing: Making the global infrastructure a reality* (pp. 809-824). [S. l.]: Wiley.
- Haustein, S. *et al.* (2013). Tweeting Biomedicine: An Analysis of Tweets and Citations in the Biomedical Literature. *Journal of the Association for Information Science and Technology*, 65(4), 656-669. <https://doi.org/10.1002/asi.23101>
- Lima, S. H. de O., & Leocádio, Á. L. (2018). Mapeando a produção científica internacional sobre inovação aberta. *Revista Brasileira de Gestão e Inovação*, 5(2), 181-208. <https://doi.org/10.18226/23190639.v5n2.08>
- Ortega, J. L. (2018). Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event Data. *Scientometrics*, 116, 2123-2138. <https://doi.org/10.1007/s11192-018-2838-z>
- Palludeto, A. W. A., & Felipini, A. R. (2019). Panorama da literatura sobre a financeirização (1992-2017): uma abordagem bibliométrica. *Economia e Sociedade*, 28(2), 313-337. <http://dx.doi.org/10.1590/1982-3533.2019v28n2art02>
- Peng, Q., & Ye, X. (2021). Research trends in social media/big data with the emphasis on data collection and data management: A bibliometric analysis. In: Nara, A., & Tsou, H. (Ed.), *Empowering Human Dynamics Research with Social Media and Geospatial Data Analytics* (pp. 47-63). Switzerland: Springer Nature. <https://doi.org/10.1007/978-3-030-83010-6>
- Peters, I. *et al.* (2016). Research data explored: an extended analysis of citations and altmetrics. *Scientometrics*, 107, 723-744. <https://doi.org/10.1007/s11192-016-1887-4>
- Piwowar, H. *et al.* (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, 1-23. <https://doi.org/10.7717/peerj.4375>
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). *Altmetrics in the Wild: Using Social Media to Explore Scholarly Impact*. New York: Cornell University.
- Robinson-García, N. *et al.* (2014). New data, new possibilities: exploring the insides of Altmetric.com. *El profesional de la información*, 23(4), 359-366. <https://doi.org/10.3145/epi.2014.jul.03>
- Rodrigues, A. A., Nóbrega, E., & Dias, G. A. (2017). Desafios da gestão de dados na era do Big Data: perspectivas profissionais. *Informação & Tecnologia*, 4(2), 63-79. <https://doi.org/10.22478/ufpb.2358-3908.2017v4n2.40538>
- Sayão, L. F., & Sales, L. F. (2019). A ciência invisível: os dados da cauda longa da pesquisa científica. In Dias, G. A., & Oliveira, B. M. J. F. (Org.), *Dados científicos: perspectivas e*

*desafios* (pp. 33-52). João Pessoa: Ed. UFPB.

Sayão, L. F., & Sales, L. F. (2012). Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. *Informação & Sociedade: Estudos*, 22(3), 179-191.

Setyowati, L., & Heriyanto, H. (2019, December). *Research Data Management in Higher Educations: Knowledge Mapping Using Bibliometric Analysis*. Artigo apresentado no Proceeding of International Conference on Documentation and Information.

Sousa, E. da S. *et al.* (2019). Mapeamento da produção científica internacional sobre intenção empreendedora. *Revista Gestão e Secretariado*, 10(3), 114-139. <http://dx.doi.org/10.7769/gesec.v10i3.901>

Specht, A. *et al.* (2015). Data management challenges in analysis and synthesis in the ecosystem sciences. *Science of the Total Environment*, 534, 144-158. <https://doi.org/10.1016/j.scitotenv.2015.03.092>

Tattersall, A., & Carroll, C. (2018). What Can Altmetric.com Tell Us About Policy Citations of Research? An Analysis of Altmetric.com Data for Research Articles from the University of Sheffield. *Frontiers in Research Methods and Analytics*, 2(9). <https://doi.org/10.3389/frma.2017.00009>

Van Eck, N. J., & Waltman, L. (2020). *VOSviewer Manual*. Leiden, Países Baixos: Universiteit Leiden.

Vanti, N., & Sanz-Casado, E. (2016). Altmetria: a métrica social a serviço de uma ciência mais democrática. *Transinformação*, 28(3), 349-358. <https://doi.org/10.1590/2318-08892016000300009>

Vanz, S. A. de S., & Caregnato, S. E. (2003). Estudos de citação: uma ferramenta para entender a comunicação científica. *Em Questão*, 9(2), 295-307.

Yang, X. *et al.* (2012). An e-Science data infrastructure for simulations within Grid computing environment: methods, approaches and practice. *Concurrency and Computation: Practice and Experience*, 25, 385-409. <https://doi.org/10.1002/cpe.2849>

Zhang, D. *et al.* (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources*, 20(1), 348-355. <https://doi.org/10.1111/1755-0998.13096>

Zhang, L., & Eichmann-Kalwara, N. (2019). Mapping the Scholarly Literature Found in Scopus on Research data management: A Bibliometric and Data Visualization Approach. *Journal of Librarianship and Scholarly Communication*, 7(1). <https://doi.org/10.7710/2162-3309.2266>