

# Práticas de Ciência Aberta da comunidade acadêmica brasileira: estudo a partir da produção científica

## *Open Science Practices of the Brazilian academic community: A study based on scientific production*

Alejandro CABALLERO-RIVERO<sup>1</sup>  0000-0003-1061-0534

Nancy SÁNCHEZ-TARRAGÓ<sup>2</sup>  0000-0002-5114-6072

Raimundo Nonato Macedo dos SANTOS<sup>1</sup>  0000-0002-9208-3266

### Resumo

A Ciência Aberta representa um novo enfoque para o trabalho científico, resultante do rápido desenvolvimento de modos interativos e colaborativos de aquisição, produção e disseminação de conhecimento, facilitados pelas tecnologias da informação. O termo é utilizado como “guarda-chuva” que engloba um conjunto de iniciativas, entre elas Acesso Aberto e Dados Abertos. A pesquisa, com enfoque cientométrico, utiliza a produção científica brasileira de artigos publicados em periódicos indexados na *Web of Science Core Collection*, no período 2015-2018, visando explorar as práticas da comunidade acadêmica nacional relacionadas com Acesso Aberto e Dados Abertos. Constata o crescimento estável dos artigos em Acesso Aberto, os quais representam mais de um terço da produção brasileira na *Web of Science* no período. A publicação em *mega journals* internacionais de Acesso Aberto constitui uma prática destacada, com ênfase em comunidades que pesquisam as áreas de Agricultura e Ciência & Tecnologia, bem como o depósito em repositório de artigos previamente publicados em periódicos de acesso aberto. Já as práticas relacionadas a Dados Abertos parecem ainda incipientes, e os resultados sugerem que seu uso predomina nas áreas de Ciência & Tecnologia, Parasitologia e Genética & Hereditariedade.

**Palavras-chave:** Acesso Aberto. Brasil. Ciência Aberta. Dados Abertos. Práticas científicas.

### Abstract

*Open Science represents a new approach to scientific work, resulting from the fast development of interactive and collaborative modes of knowledge acquisition, production, and dissemination facilitated by information technologies. The term is used as an “umbrella” encompassing a set of initiatives, including Open Access and Open Data. This paper, with a scientometric approach, uses the Brazilian scientific production of articles published in journals indexed in the Web of Science Core Collection during the 2015-2018 period, to investigate the practices of the Brazilian academic community related to Open Access and Open Data. A stable growth of Open Access articles, representing more than a third of the Brazilian scientific production in the Web of Science, is observed during that*

<sup>1</sup> Universidade Federal de Pernambuco, Centro de Artes e Comunicação, Programa de Pós-Graduação em Ciência da Informação. Av. Prof. Moraes Rego, 1235, Cidade Universitária, 50670-901, Recife, PE, Brasil. Correspondência para/Correspondence to: A. CABALLERO-RIVERO. E-mail: <caballero.alecaba@gmail.com>.

<sup>2</sup> Universidade Federal de Rio Grande do Norte, Centro de Ciências Sociais Aplicadas, Departamento de Ciência da Informação. Natal, RN, Brasil.

Apoio: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) (Código de financiamento 001) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Bolsa de produtividade).

Recebido em 12 de março de 2019, versão final reapresentada em 11 de julho de 2019 e aprovado em 3 de outubro de 2019.

Como citar este artigo/How to cite this article

Caballero-Rivero, A.; Sánchez-Tarragó, N.; Santos, R.N.M. Práticas de Ciência Aberta da comunidade acadêmica brasileira: estudo a partir da produção científica. *Transinformação*, v.31, e190029, 2019. <http://dx.doi.org/10.1590/2318-0889201931e190029>



period. Publication in Open Access international mega journals is a prominent practice, particularly, in the Agriculture and Science & Technologies research communities, as well as the deposit in repositories of previously published articles in Open Access journals. Practices related to Open Data still seem incipient; the results suggest that the deposit/use of Open Data repositories predominates in the areas of Science & Technology, Parasitology and Genetics & Heredity.

**Keywords:** Open access. Brazil. Open science. Open data. Scientific practices.

## Introdução

Os princípios de abertura, integridade e reprodutibilidade têm sido considerados como essenciais na ciência desde a fundação das primeiras sociedades e revistas científicas no século XVII. Unicamente o acesso da comunidade acadêmica aos resultados de pesquisa permite que os dados, as ferramentas e as metodologias utilizadas sejam examinados, analisados, discutidos, validados ou rejeitados, definindo quais dessas reivindicações cumprem as exigências para passar a constituir parte do corpo comum de conhecimento. Assim, incrementa-se a confiabilidade (grau de fidelidade) e a confiança na reutilização desses resultados; logo, a troca aberta e livre de ideias e evidências garantem e aceleram o progresso científico.

Os rápidos avanços nas Tecnologias de Informação e Comunicação (TIC) alteraram e melhoraram radicalmente as práticas pelas quais os dados, as metodologias, as ferramentas e os resultados de pesquisa podem ser produzidos, disseminados e utilizados. Essas mudanças proporcionaram um crescimento exponencial da quantidade de dados e publicações disponíveis, criando oportunidades sem precedentes para acelerar e dinamizar a ciência (US National Science Foundation, 2018).

O conjunto de iniciativas que conformam o que se conhece como Ciência Aberta, tais como Acesso Aberto, Dados Abertos, plataformas colaborativas, dentre outras, constitui uma abordagem para a prática científica que aplica as novas tecnologias digitais e as ferramentas colaborativas para promover esforços conjuntos de investigação e avaliação, bem como para disseminar e usar o conhecimento produzido da forma mais rápida e ampla possível. Assim, promove a reprodutibilidade dos resultados de pesquisa, a transparência metodológica, o acesso global aos resultados de investigação e a economia de recursos (tempo, dinheiro, recursos humanos, equipamento), buscando incrementar o impacto social da ciência.

Os elementos indicados são reconhecidos como essenciais para a ciência, e muitos pesquisadores declaram estar interessados em incorporar práticas de Ciência Aberta em suas pesquisas (Scheliga; Friesike, 2014); logo, espera-se que estas se manifestem de forma rotineira. No entanto, diversos estudos evidenciam que isso nem sempre acontece; os sistemas de recompensa acadêmica (Nosek; Speis; Motyl, 2012), os requisitos de publicação (reais ou percebidos) (Ioannidis *et al.*, 2014), junto com outro conjunto de obstáculos individuais e sistêmicos, fazem das práticas de Ciência Aberta um dilema social, pois o que é considerado bom para a comunidade científica nem sempre é percebido como bom para o pesquisador individual (Scheliga; Friesike, 2014), que pode até se sentir desencorajado a participar dessas práticas.

Como argumentam Pontika *et al.* (2015), nesse contexto, a Ciência Aberta enfrenta vários desafios. Por se tratar de uma abordagem relativamente nova e complexa, a adoção dessas práticas exige uma mudança no comportamento dos pesquisadores no que diz respeito à forma de realizar as pesquisas e compartilhar dados, resultados ou metodologias. Por outro lado, a informação referente às novas práticas está dispersa em diversas fontes, dificultando sua localização; em decorrência, muitos pesquisadores, por não estarem cientes das diversas modalidades relativas a essas novas práticas, não as utilizam.

Apesar de existirem trabalhos teóricos nacionais e internacionais sobre o que é Ciência Aberta, suas práticas, sua importância e suas diretrizes (Albagli; Maciel; Abdo, 2015; Chan; Loizides, 2017; Santos; Almeida; Henning, 2017), bem como estudos sobre a implementação prática de algumas dessas modalidades no Brasil (Appel; Albagli, 2019;

Packer, 2014), não foram identificados estudos empíricos que tratassem sobre o assunto no contexto brasileiro, especificamente sobre Acesso Aberto e Dados Abertos.

A indagação acerca das práticas dos pesquisadores é fundamental para a implementação de políticas que sustentem o desenvolvimento e promoção de iniciativas de Ciência Aberta. Isso é particularmente importante no Brasil, onde, como argumenta Abdo (2015), as universidades, institutos e agências de pesquisa, vivem um estado de contradição. Por um lado, iniciativas de Ciência Aberta promovem o compartilhamento e a colaboração na produção de conhecimento e no uso dos recursos de pesquisa disponíveis, enquanto, por outro lado, nessas instituições persistem atitudes baseadas em privilégios, prestígios e interesses específicos, que perpetuam o modelo anterior e obstaculizam essas novas práticas.

Frente a essa constatação, formula-se aqui a seguinte pergunta de pesquisa: quais práticas de Ciência Aberta são utilizadas pela comunidade acadêmica brasileira? Consequentemente, o presente trabalho pretende explorar práticas de Ciência Aberta que estão sendo empregadas atualmente pela comunidade acadêmica brasileira, especificamente, aquelas relacionadas com as iniciativas de Acesso Aberto e Dados Abertos. Como objetivos específicos pretende-se identificar e descrever a existência dessas práticas, bem como seu crescimento (ou não), uso e alcance, em prol da formulação de políticas de apoio e promoção dessas iniciativas.

## Ciência Aberta e suas principais práticas

A Ciência Aberta representa um novo enfoque para o trabalho científico, resultante do rápido desenvolvimento de modos interativos e colaborativos de aquisição, produção e disseminação de conhecimento, facilitados pelas TIC. Como indicam Albagli, Maciel e Abdo (2015), esse é um campo emergente de pesquisa, de modo que não existe ainda um quadro teórico consensual nem uma definição formal de Ciência Aberta. O termo é utilizado como “guarda-chuva” que engloba um conjunto de iniciativas vinculadas ao uso de tecnologias digitais, ferramentas colaborativas e licenças alternativas de propriedade intelectual, que visam potencializar a pesquisa colaborativa, a transparência na avaliação e a rápida e ampla disseminação e uso dos resultados.

O estudo de Vicente-Saez e Martinez-Fuentes (2018) é considerado pelos autores deste trabalho como o esforço mais rigoroso, integral e atual para definir Ciência Aberta. Apresenta uma revisão sistemática, abrangente e detalhada da literatura publicada sobre essa temática, indexada em bases de dados internacionais (acadêmicas e intergovernamentais) no período 1985-2016. Essa sistematização levou aos autores em referência a definir a Ciência Aberta como “[...] o conhecimento transparente e acessível que é compartilhado e desenvolvido por meio de redes colaborativas” (Vicente-Saez; Martinez-Fuentes, 2018, p.434, tradução nossa<sup>3</sup>). Contempla, portanto, a construção, organização e disseminação do conhecimento científico de forma transparente, acessível e colaborativa, para garantir a integridade, a reprodutibilidade e o reuso das explicações, previsões e leis formuladas sobre o universo e a sociedade. Nesse sentido, as práticas a serem identificadas deverão corresponder a essas características.

No presente trabalho, utiliza-se a noção de práticas de MacIntyre (2007), que as define como:

[...] qualquer forma coerente e complexa de atividade humana cooperativa socialmente estabelecida, por meio da qual bens internos a essa forma de atividade são realizados na tentativa de alcançar padrões de excelência apropriados e parcialmente definitivos para essa forma de atividade (MacIntyre, 2007, p.187, tradução nossa<sup>4</sup>).

<sup>3</sup> [...] *transparent and accessible knowledge that is shared and developed through collaborative networks.*

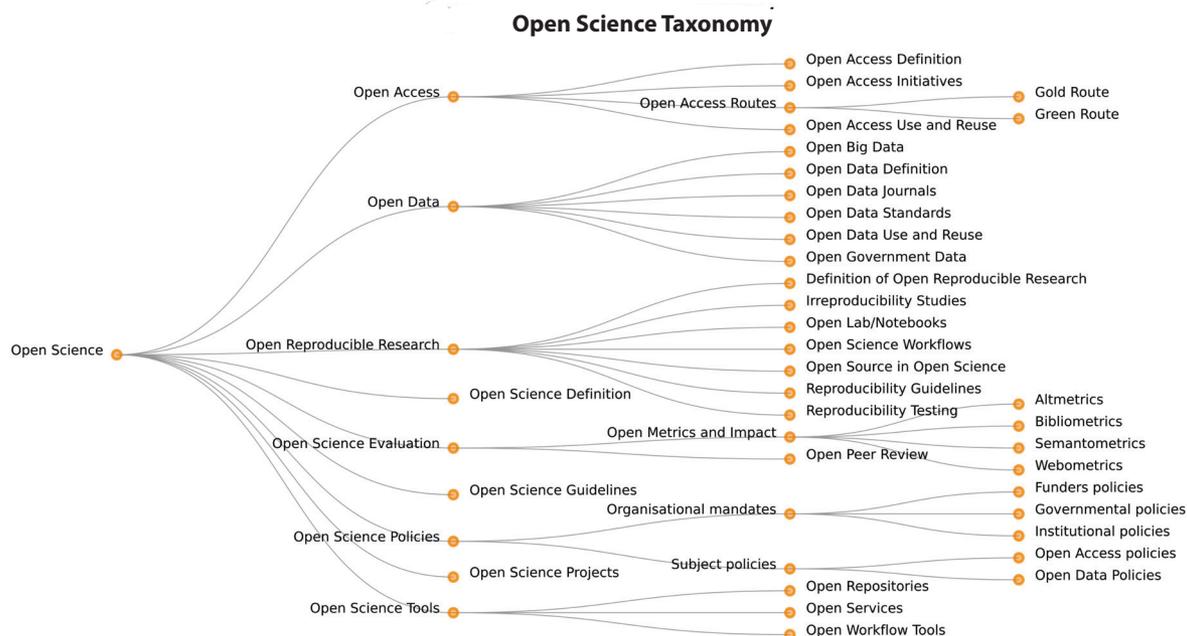
<sup>4</sup> [...] *any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity.*

Portanto, as práticas implicam a colaboração entre os seres humanos para alcançar objetivos compartilhados que emergem no contexto de um sistema maior; vão além de pessoas colaborando numa tarefa pontual; é a colaboração em tarefas inseridas numa prática coletiva. Envolve padrões de excelência, obediência às regras, bem como a realização de bens para o bem comum; logo, a participação nas práticas exige aceitar esses padrões e submeter a eles as atitudes, escolhas ou preferências individuais.

Como apresentado em Albagli, Maciel e Abdo (2015), a Ciência Aberta reflete a tensão existente entre as novas práticas de colaboração, interação e abertura na produção, disseminação e uso do conhecimento científico promovido pela comunidade acadêmica, por um lado a expansão e o endurecimento dos mecanismos de privatização do conhecimento (leis de propriedade intelectual, *etc.*), resultantes do aprofundamento das relações capitalistas de mercado em escala global por outro lado. Apesar de que o desenvolvimento das TIC tenha eliminado as barreiras técnicas para a produção colaborativa, a disseminação e o acesso ao conhecimento científico, os mecanismos de privatização têm impactado diretamente as formas institucionais e organizacionais de fazer ciência, estabelecendo barreiras econômicas ou jurídicas.

Nesse sentido, no presente trabalho, consideram-se práticas de Ciência Aberta as iniciativas colaborativas de produção, disseminação e uso do conhecimento científico, já estabelecidas no contexto da comunidade acadêmica global, as quais funcionam por meio de padrões de excelência científica e regras de comportamento que, utilizando as novas tecnologias digitais e ferramentas de colaboração, tentam contornar os mecanismos de privatização e garantir os princípios de abertura, integridade e reprodutibilidade.

Até o presente, não há consenso formado sobre quais são essas práticas (Albagli; Maciel; Abdo, 2015; Haddaway, 2018; Vicente Saez; Martinez-Fuentes, 2018). Conforme critério dos autores deste trabalho, o estudo de Pontika et al. (2015) apresenta uma taxonomia útil para visualizar o domínio da Ciência Aberta (Figura 1), fornecendo uma terminologia estruturada sobre as principais iniciativas e práticas associadas. Como o foco do presente trabalho está voltado para as práticas relacionadas com o Acesso Aberto e os Dados Abertos, utiliza-se a taxonomia



**Figura 1.** Taxonomia do domínio da Ciência Aberta.

Fonte: Pontika *et al.* (2015, p. 3).

para ilustrar as questões essenciais envolvidas. No entanto, esclarece-se que essa taxonomia foi elaborada para sistematizar os recursos educacionais sobre Ciência Aberta hospedados no portal de capacitação FOSTER (<https://www.fosteropenscience.eu/>); portanto, não deve ser tomada como definitiva, mas como um ponto de partida para uma exploração mais aprofundada sobre a temática.

## Acesso Aberto

As práticas de Acesso Aberto são centrais para a Ciência Aberta. Os artigos são os produtos das pesquisas; apresentam tanto os resultados, quanto uma descrição básica dos procedimentos metodológicos utilizados. Logo, constituem o ponto de partida para que os pesquisadores consigam acessar, analisar, avaliar, reproduzir e validar ou rejeitar esses resultados e, a partir deles, construir suas próprias pesquisas.

Na literatura, identificam-se duas práticas principais: publicação em periódicos de Acesso Aberto (rota dourada) e depósito em repositórios de Acesso Aberto (rota verde) (Budapest Open Access Initiative, 2002). No primeiro caso, basicamente, os pesquisadores publicam seus resultados de pesquisa em revistas que disponibilizam gratuitamente todos os artigos, mediante o uso de licenças que expressam as condições sob as quais o acesso ao conteúdo e seu reuso serão permitidos. Tipicamente, são licenças *Creative*. Além dessas, também existem outros tipos, como licenças de conteúdo aberto e declarações de permissão, por meio das quais os detentores dos direitos autorais consentem que os usuários reutilizem os artigos para leitura, descarga ou cópia.

Conforme os dados do *Directory of Open Access Journals* (DOAJ) (<http://doaj.org>), em julho de 2019 existiam 13.491 revistas desse tipo no mundo, publicadas em 131 países, sobre 20 áreas, num total de 4.100.680 artigos. O estudo de Alperin (2014) indica que os periódicos latino-americanos representam 19%-20% do total de periódicos em DOAJ. A maior parte dos periódicos *online* da região são de Acesso Aberto, bem como 74% de todos os periódicos latino-americanos em Scopus, proporção mais de oito vezes superior à do resto do mundo (9%).

No entanto, os artigos em Acesso Aberto não se limitam àqueles que aparecem em revistas reconhecidas como tal pelo DOAJ. Há editoras que utilizam modelos híbridos, como periódicos por assinatura em que se disponibilizam alguns dos artigos em Acesso Aberto. Tanto em um caso, como em outro, há editores que exigem que os autores (ou suas instituições) paguem uma taxa de processamento de artigos (APC, *Article Processing Charge*) para que o artigo seja disponibilizado em Acesso Aberto.

A segunda prática implica o depósito de artigos em plataformas digitais que fornecem acesso livre a esses resultados (tipicamente, os denominados repositórios de Acesso Aberto). O depósito depende do contrato assinado entre os autores e os editores no momento em que o artigo foi submetido para publicação. Assim, os pesquisadores poderiam depositar a versão final publicada pela revista ou a versão aceita para publicação. Os pesquisadores também têm a possibilidade de depositar os pre-prints, ou seja, uma versão de um artigo que precede à revisão por pares e à publicação em um periódico. Conforme o *Directory of Open Access Repositories* (OpenDOAR) - (<http://v2.sherpa.ac.uk/opensoar/>), o número de repositórios no mundo aumentou de 88, em dezembro de 2005, para 4.156, em julho de 2019. Na América Latina e Caribe destacam Brasil (110 repositórios), Colômbia (n=69), Peru (n=64) e Argentina (n=54).

Finalmente, existem outros veículos de comunicação utilizados pelos pesquisadores para disseminar os artigos em Acesso Aberto, tais como, os *websites* pessoais, *blogs*, *wikis*, fóruns de discussão, redes P2P, dentre outros. No entanto, a disseminação dos artigos por essas vias apresenta desvantagens com relação às revistas e repositórios, tais como problemas jurídicos, ausência de URL duradouras e dificuldade de preservação a longo prazo.

## Dados Abertos

Os dados resultantes das observações científicas, isto é, fatos, números, letras e símbolos que descrevem objetos, ideias, condições, situações ou outros fenômenos do universo, são fundamentais para o progresso da

ciência. No entanto, há fortes indícios de uma tendência progressiva à sua privatização e comercialização, utilizando mecanismos legais extrínsecos à ciência, tais como novos direitos autorais sobre os dados no entorno digital e restrições contratuais aplicadas tecnologicamente, dentre outros. Couldry e Mejias (2018) têm se referido a esse fenômeno como um novo colonialismo, sendo os dados o recurso explorado.

As iniciativas de Dados Abertos encontram-se no meio dessa tensão entre a socialização e a exploração comercial dos dados, pois buscam a disponibilidade destes para uma reutilização mais ampla. Segundo o *Open Data Handbook da Open Knowledge Foundation* (<http://opendatahandbook.org/>), os Dados Abertos são aqueles que podem ser usados livremente, reutilizados e redistribuídos por qualquer pessoa, sujeitos unicamente aos requisitos de atribuição de autoria e de compartilhamento na mesma forma em que foram obtidos. Isso implica, primeiramente, a disponibilidade integral dos conjuntos de dados a um “custo razoável de reprodução” (o provedor não deve assumir custos pelo fornecimento), preferivelmente, por meio da descarga gratuita pela *Internet*, em um formato legível por máquina. Em segundo lugar, os dados devem ser disponibilizados de forma a permitir sua reutilização e redistribuição, incluindo sua mistura com outros conjuntos de dados, e sem discriminação quanto ao tipo de usuários; portanto, deverão ser oferecidos com uma licença que garanta esses direitos (exemplo, *Data Base License*).

No contexto dos Dados Abertos, identificam-se várias práticas de Ciência Aberta. As abordagens mais tradicionais são a publicação dos dados como parte integrante dos artigos publicados, bem como sua publicação como arquivos suplementares. Desde 2009, a maioria dos periódicos aceita esse tipo de arquivo (Reilly *et al.*, 2011). Os artigos devem ser de Acesso Aberto e fornecer os conjuntos completos de dados em formato que garanta seu reuso. Deve-se considerar que, quando os dados são publicados como parte integral dos artigos, às vezes são listados como anexo, às vezes no texto, e nem sempre com informação para sua reutilização. Adicionalmente, há revistas que implementam restrições relativas ao número total de arquivos suplementares.

Uma abordagem relativamente recente é a publicação dos dados em *data papers*, ou seja, artigos que descrevem um ou vários conjuntos de dados acessíveis *online*; são publicados em periódicos conhecidos como *data journals*, cuja finalidade é descrever os dados de uma pesquisa, não a própria pesquisa (Chavan; Penev, 2011). É considerada uma prática de Ciência Aberta, pois, embora os *data journals* possam adotar modelos híbridos, via de regra exigem o depósito dos conjuntos de dados em repositórios de Dados Abertos. Em 2015 existiam 116 *data journals* de 15 editoras, majoritariamente nas áreas de Ciências da Saúde, da Vida e Física, publicando tanto *data papers* quanto artigos tradicionais (Candela *et al.*, 2015).

O depósito dos dados em repositórios de Dados Abertos também tem sido uma prática promovida por agências de fomento, formuladores de políticas, instituições de pesquisa e periódicos. Geralmente, o depósito é feito em repositórios de dados “qualificados”, reconhecidos internacional ou institucionalmente, que garantem a disponibilidade a longo prazo e o acesso permanente aos conjuntos de dados. Exemplos desses repositórios são *GenBank*, *Dryad*, *figshare* e *Zenodo*.

O depósito de Dados Abertos constitui uma prática já estabelecida em algumas áreas científicas, sobretudo naquelas que utilizam uma abordagem de pesquisa baseada em uso intensivo de dados. Nessas áreas já existem padrões estipulados para as estruturas dos conjuntos de dados, assim como repositórios bem estabelecidos com critérios e regras claras para descrevê-los e compartilhá-los (ex. astronomia, física de alta energia, sismologia, genômica, ciência da computação), algo que, em geral, não acontece em outras áreas (Wallis; Rolando; Borgman, 2013). Em muitas dessas áreas também existem políticas e mandatos institucionais e de financiadoras que reforçam essas práticas. Elas apresentam um grande crescimento, como evidenciam, por exemplo, as estatísticas de *GenBank* (2019), em que o número de bases de dados aumentou de 680.338, em dezembro de 1982, para 329.835.282.370, em julho de 2019, enquanto os registros de sequências genéticas se incrementaram de 606 para 213.383.758 em igual período.

Outra prática exercida pelos pesquisadores das áreas cujas pesquisas são baseadas em dados é o uso dos dados depositados nesses repositórios. Essa prática permite a avaliação de diferentes hipóteses a partir do mesmo conjunto de dados, além de possibilitar a reprodutibilidade dos resultados de investigação, ou seja, os dados podem ser reutilizados sem a necessidade de realizar novas coletas e podem ser agregados a partir de diferentes conjuntos para obter novos *insights*.

## Procedimentos Metodológicos

Para dar cumprimento ao objetivo de explorar práticas de Ciência Aberta relacionadas às iniciativas de Acesso Aberto e Dados Abertos na comunidade de pesquisadores brasileiros, decidiu-se realizar um estudo da produção científica publicada em periódicos, considerando que tais modalidades têm um reflexo importante nos canais de comunicação da ciência. Trata-se, portanto, de uma pesquisa exploratória que utiliza um enfoque cientométrico, definido como a medição quantitativa das atividades científicas (Price, 1969; Santos; Kobashi, 2009). Segundo esses autores, os dados quantitativos sobre revistas e artigos científicos obedecem a certas regras estáveis, configurando-se como indicadores do estado da ciência. Também Hjørland (2016) aponta o uso da cientometria e outros estudos métricos como contribuições aos estudos da sociologia da ciência sobre as atividades científicas.

Considerou-se como população a produção científica brasileira de todas as áreas do conhecimento. Como amostra, foi selecionada a produção científica brasileira de artigos baseados em pesquisa original, que contasse, entre os autores, com ao menos um afiliado a uma instituição sediada no país; e fosse publicada em periódico indexado nas bases de dados *Science Citation Index Expanded*, *Social Sciences Citation Index* e *Arts & Humanities Citation Index* da *Web of Science Core Collection* (WoS) da *Clarivate Analytics*, no período 2015-2018. A escolha dessas bases justifica-se por sua facilidade para realizar análises sobre a produção científica, na medida em que possuem registros bibliográficos detalhados e identificação de coautores e instituições nas publicações. Além disso, têm também uma cobertura multidisciplinar de periódicos e artigos de alcance global e apresentam uma quantidade representativa de textos de pesquisadores brasileiros, segundo a Fundação de Amparo à Pesquisa do Estado de São Paulo (2011). Adicionalmente, permitem a identificação direta dos artigos em Acesso Aberto publicados em revistas ou depositados em repositórios nas suas diversas modalidades, por meio de um campo específico nas bases de dados.

Para a análise das práticas de Acesso Aberto foram realizados os seguintes procedimentos:

a) Consulta às bases da WoS (7 de fevereiro de 2019), utilizando como estratégia de busca o termo *brazil* no campo “Endereço”, limitado ao período 2015-2018 e refinado pelo campo “Tipo de Documento” (Artigo) para um total de 181.283 registros. Os resultados foram filtrados utilizando o campo “*open access*”, totalizando 71.314 registros. Desse total, foram selecionados apenas os artigos cujo primeiro autor estivesse afiliado a uma instituição brasileira. O *corpus* final para a análise das práticas de Acesso Aberto foi conformado por 70.983 registros.

b) Análise das variáveis – quantidade de documentos por ano, por categoria de Acesso Aberto segundo WoS, por instituição, por título de revista e por área de conhecimento.

Já para a análise das práticas relacionadas com Dados Abertos foram adotados os seguintes procedimentos:

a) A partir da mesma consulta geral indicada no procedimento anterior, em que foram recuperados 181.283 registros, com tipologia de documento artigo e *brazil* no campo “Endereço”, realizou-se logo uma busca no campo “Tópico” (títulos, resumos ou palavras-chave) pelos termos *open data*, *data paper*, *Dryad*, *figshare*, *Zenodo* e *GenBank*, recuperando-se 305 registros.

b) Leitura dos resumos desses trabalhos, descartando-se aqueles que mencionassem os termos como objetos de estudo, mas não como práticas dos autores dos trabalhos. Desse total, selecionaram-se apenas os artigos cujo primeiro autor estivesse afiliado a uma instituição brasileira, obtendo-se um total de 249 artigos.

Levando em consideração que a busca pelos termos *Dryad* e *Zenodo* não recuperou resultados, decidiu-se complementar a coleta de dados procurando diretamente nesses repositórios a existência de registros associados a publicações disponíveis na WoS.

Para isso, realizou-se a consulta aos repositórios no dia 24 de junho de 2019, limitando a busca aos conjuntos de dados sob o termo *brazil*. No caso de *Zenodo*, ainda foi possível restringir a busca a arquivos com formatos típicos de dados, especificamente, .zip, .xml, .json e .xlsx, recuperando-se 98 registros. Em *Dryad* a busca restringiu-se aos *Data packages*, recuperando-se 360 registros. Desses totais, verificou-se um a um cada registro, selecionando-se apenas aqueles que cumprissem os seguintes critérios: corresponder a um artigo disponível na WoS, com data de publicação entre 2015-2018, cujo primeiro autor estivesse afiliado a uma instituição brasileira e que não tivesse sido identificado já nos repositórios de *GenBank* e *figshare*.

Os 110 registros resultantes se fundiram com os registros recuperados na primeira etapa da busca na WoS, num total de 339 registros sinalizando práticas relacionadas com Dados Abertos. As variáveis analisadas foram quantidade de documentos por ano, por instituição, por repositório de dados e por área de conhecimento.

Os processamentos foram realizados com o *software* VantagePoint (v.11Pro 64, Wesley Chapel, Florida, USA), e as tabulações em *Microsoft Excel 2016 Plus*.

## Resultados e Discussão

### Práticas de Acesso Aberto

Os artigos publicados em Acesso Aberto por autores de instituições brasileiras representam 39% da produção brasileira na WoS, no período 2015-2018, com um comportamento estável, nesses quatro anos: 20.032 artigos em 2015; 18.685 em 2016; 19.373 em 2017 e 12.983 em 2018 – este último ano, provavelmente, com uma cobertura incompleta devido à data de coleta dos dados. O estudo de Minniti, Santoro e Belli (2018) constata o crescimento significativo da produção em Acesso Aberto indexada pela WoS, especialmente entre 2014 e 2016, produzida no âmbito da América Latina e Caribe e liderada pelo Brasil, cuja produção científica, em números absolutos, representa 69,7% do total de documentos da região.

Essa produção está classificada na WoS em cinco categorias: DOAJ Ouro: artigos publicados em revistas reconhecidas pelo DOAJ como de Acesso Aberto; Outro Ouro: artigos em Acesso Aberto publicados em revistas por assinatura (híbridas); Bronze: artigos em Acesso Aberto localizados nos sites das editoras sem licenciamento *Creative Commons*; Verde publicado: versões finais de artigos depositadas em repositórios de Acesso Aberto; e Verde aceito: versões aceitas de artigos, ainda não publicadas (Tabela 1).

Os dados mostram que predomina a produção em periódicos identificados com licença *Creative Commons* e cadastrados no DOAJ. O total de artigos publicados em periódicos de Acesso Aberto é muito alto (91%), restando 9% que, presumivelmente, não foram publicados nesse tipo de revista, mas depositados em repositórios. Por outro lado, 15% dos artigos foram categorizados como Bronze, sinalizando que foram publicados sem uma licença *Creative Commons*. Pode-se inferir que esses artigos tenham sido publicados em periódicos que, embora autodenominados de Acesso Aberto, não utilizam essas licenças para respaldar o modelo vigente, tal como detectou a pesquisa de Sánchez-Tarragó *et al.* (2016). Outra possível interpretação é que esses artigos possam ter sido publicados em periódicos que não utilizam o modelo de Acesso Aberto, mas que permitiram ou permitem o Acesso Aberto aos artigos como parte de uma promoção ou outra condição temporal.

**Tabela 1.** Quantidade de artigos segundo categoria de Acesso Aberto (periódicos e repositórios; 2015-2018).

Categoria	Quantidade (n)	% do total de artigos em Acesso Aberto
DOAJ Ouro	50.221	70,7
Outro Ouro	3.877	5,5
Bronze	10.742	15,1
Total em periódicos	64.840	91,3
Verde Publicado	24.072	33,9
Verde Aceito	1.298	1,8
Total em repositórios	25.370	35,7

Fonte: Elaborado pelos autores (2019).

Nota: DOAJ: *Directory of Open Access Journals*.

Aproximadamente 36% da produção foi depositada em repositórios de Acesso Aberto. Predomina o depósito das versões finais publicadas em periódicos. O cruzamento dos dados das categorias de periódicos e repositórios sugere que quase 26% dos artigos publicados em periódicos de Acesso Aberto também estão acessíveis em repositórios. Se essa duplicidade está em conformidade com o interesse do pesquisador de diversificar as vias de disseminação das suas pesquisas ou se ela se deve a políticas institucionais, é algo a ser investigado em outros estudos. Uma outra prática interessante e amplamente estendida em algumas áreas, como a Física, é o depósito de *pre-prints*, porém, não foi possível analisá-la devido à inexistência dessa categoria no WoS.

No entanto, considerar o depósito em repositórios como uma prática de Ciência Aberta levanta algumas questões. Primeiro, alguns estudos indicam que apenas uma pequena porcentagem dos pesquisadores deposita “espontaneamente” seus artigos em repositórios (Creaser *et al.*, 2010); também, as taxas de depósito de artigos e *pre-prints* são variáveis, dependendo da área do conhecimento. Diferentemente do que acontece na publicação em revistas, em que os pesquisadores mostram um grau relativamente alto de independência na decisão sobre o veículo de comunicação, no caso do depósito nos repositórios, essa independência é menor devido ao envolvimento significativo de outros atores, além da questão dos direitos autorais. Por exemplo, a falta de clareza na política editorial das revistas em relação ao depósito em repositórios faz com que, muitas vezes, os autores não tenham certeza sobre seus direitos autorais e fiquem preocupados sobre sua possível violação. Como mostra o estudo de Ramirez e Hanlon (2011), apenas 10% dos autores depositam seus artigos sem assistência da sua instituição, ou seja, na maior parte das vezes o depósito é realizado conjuntamente com as pessoas responsáveis pelos repositórios. Considerando essas questões, os autores deste trabalho assumem que o depósito em repositórios de Acesso Aberto pode ser considerado uma prática de Ciência Aberta, quando realizado regularmente pelos pesquisadores como parte do fluxo de comunicação científica.

De fato, a presente pesquisa mostra que as práticas de publicação e depósito em Acesso Aberto diferem segundo a área de estudo. Foram identificadas 145 áreas com produção em Acesso Aberto. A Tabela 2 registra uma lista ordenada, decrescentemente, das dez áreas de pesquisa com maior quantidade (acima de 1.900), de artigos indexados no período, assim como a proporção correspondente ao total de artigos em Acesso Aberto (70.983) para cada categoria. Desse total, 12,84% dos artigos publicados em periódicos de Acesso Aberto foram classificados na área de Agricultura, enquanto 8,90% foram classificados em Ciência & Tecnologia. É necessário ressaltar que essas são as duas áreas de melhor representação na WoS, com 9.324 e 10.929 artigos, respectivamente. Também é interessante notar que, enquanto a produção científica em Agricultura dissemina-se fundamentalmente em periódicos de Acesso Aberto, com uma presença pouco expressiva em repositórios (0,30%), os artigos de Ciência & Tecnologia usam ambas as vias com proporções equilibradas (periódicos 8,90%; repositórios 6,50%).

**Tabela 2.** Proporção de artigos segundo categorias de Acesso Aberto por área de pesquisa (2015-2018).

Áreas de pesquisa WoS	DOAJ	Outro	Bronze	Total	Verde	Verde	Total
	Ouro	Ouro		Periódicos	publicado	aceito	Repositórios
%							
<i>Science &amp; Technology - other topics</i>	8,43	0,16	0,31	8,90	6,40	0,10	6,50
<i>Agriculture</i>	11,84	0,17	0,83	12,84	0,30	0	0,30
<i>Physics</i>	2,54	0,86	0,77	4,17	2,32	0,03	2,35
<i>Public, Environmental &amp; Occupational Health</i>	4,37	0,10	0,44	4,91	1,30	0,06	1,36
<i>Chemistry</i>	3,12	0,21	0,61	3,94	1,49	0,08	1,57
<i>Tropical medicine</i>	2,54	0,04	0,18	2,76	1,94	0	1,94
<i>Parasitology</i>	2,56	0,04	0,08	2,68	1,60	0,03	1,63
<i>Veterinary Sciences</i>	3,41	0,06	0,33	3,80	0,31	0	0,31
<i>Biochemistry &amp; Molecular Biology</i>	1,16	0,18	1,39	2,73	1,15	0,19	1,34
<i>Microbiology</i>	1,26	0,10	0,90	2,26	1,54	0,05	1,59

Fonte: Elaborado pelos autores (2019).

Nota: WoS: *Web of Science*; DOAJ: *Directory of Open Access Journals*.

**Tabela 3.** Produção científica brasileira de Acesso Aberto em periódicos (2015-2018).

Periódicos	Quantidade de artigos
<i>PLoS One</i>	4.661
<i>Scientific Reports</i>	1.477
Semina: Ciências Agrárias	1.397
<i>Journal of the Brazilian Chemical Society</i>	1.237
Ciência Rural	1.182
Ciência & Saúde Coletiva	1.136
Anais Brasileiros de Dermatologia	880
Arquivo Brasileiro de Medicina Veterinária e Zootecnia	852
Pesquisa Veterinária Brasileira	826
Revista de Saúde Pública	732

Fonte: Elaborado pelos autores (2019).

Ao analisar o crescimento por área no período 2015-2017<sup>5</sup>, observou-se que, das 23 áreas que publicaram mais de 1 mil trabalhos em Acesso Aberto, os maiores incrementos (acima de 10%) aconteceram em Ciência dos Materiais, Medicina Geral & Interna, Silvicultura, Física, Ciência & Tecnologia, Química e Ciências Veterinárias, nessa ordem. Por categoria, os crescimentos mais significativos das áreas (acima de 20%) se manifestaram nas publicações DOAJ Ouro, Verde publicado e Bronze; nas primeiras, destacaram Ciência dos Materiais, Microbiologia, Medicina Geral & Interna e Química; nas segundas, Ciência dos Materiais, Medicina Geral & Interna e Química; nas terceiras, unicamente Silvicultura. Nas outras categorias não foi observado crescimento expressivo em nenhuma área.

A produção científica em Acesso Aberto foi publicada em 2.261 periódicos. A Tabela 3 apresenta as dez primeiras posições dos periódicos com mais de 700 artigos no período.

Os títulos *PloS One* e *Scientific Reports*, no topo da lista, são periódicos de Acesso Aberto, publicados pela *Public Library of Science e Nature Research*, respectivamente. Ambos são considerados *mega journals*, tipologia de

<sup>5</sup> Não se considerou o ano de 2018, pois, provavelmente, a quantidade de artigos está incompleta devido à data de coleta dos dados.

periódico que se caracteriza por cobrir uma ampla diversidade de disciplinas (ciências, tecnologias, medicina), grande quantidade de artigos publicados anualmente (mais de 20 mil), níveis relativamente baixos de rejeição e cobrança de taxas por processamento de artigos. Ambos os periódicos estão registrados no DOAJ. Os demais são periódicos brasileiros de Acesso Aberto também cadastrados no DOAJ. É notável que, embora quase 10% dos artigos tenham sido publicados nesses *mega journals*, as publicações nacionais de Acesso Aberto constituem um canal muito importante de disseminação dos resultados de pesquisa, com 8% deles publicados nos quatro periódicos nacionais que ocupam o topo da lista, cada um com mais de mil artigos no período.

No que diz respeito às instituições, cabe observar que em algumas predomina amplamente a publicação em periódicos de Acesso Aberto, com destaque para os periódicos DOAJ Ouro, enquanto, aparentemente, é menor o depósito desses artigos em repositórios. Esse é o caso, por exemplo, da Universidade Federal de Santa Maria (77,0% DOAJ Ouro e 11,1% em repositórios), da Universidade Federal de Lavras (76,7% e 7,7%), da Universidade Federal de Viçosa (71,2% e 16,0%) e da Universidade Estadual de Londrina (71,6% e 16,3%). Outras instituições, além de utilizarem os periódicos de Acesso Aberto, servem-se da via Verde em maior proporção, destacando a Fundação Oswaldo Cruz (58,3% DOAJ Ouro e 34,0% em repositórios), a Universidade de São Paulo (52,7% e 33,1%), a Universidade Federal de Rio de Janeiro (52,5% e 31,1%), a Universidade Estadual de Campinas (53,0% e 31,0%) e a Universidade Federal de São Paulo (55,7% e 31,7%).

## Práticas relacionadas com Dados Abertos

Os resultados obtidos neste estudo mostraram a criação, depósito ou uso de conjuntos de Dados Abertos associados a pesquisas publicadas em 339 artigos, bem como evidenciaram uma tendência ascendente no período estudado: 76 artigos em 2015, 83 em 2016, 86 em 2017 e 94 em 2018. Adicionalmente, 10 artigos foram publicados como *data papers* em *data journals*. Esses resultados sugerem ser essa uma prática ainda incipiente, na medida em que esses 339 artigos representam, aproximadamente, apenas 0,48% da produção total em Acesso Aberto, no período.

No entanto, é necessário tomar esses resultados com cautela. A metodologia utilizada na pesquisa tem duas limitações que podem ter impactado os resultados. Primeiro, a coleta de dados foi realizada tomando como referência somente os artigos publicados na WoS, que, como bem se sabe, não garante uma cobertura equilibrada da produção científica mundial e apresenta vieses no que diz respeito ao tipo de publicação, idioma, área de conhecimento, dentre outros. Adicionalmente, como parecem confirmar os resultados, é muito provável que, nas áreas em que o uso/depósito de Dados Abertos já esteja imbricado nas rotinas de pesquisa, tais práticas não sejam explicitadas nos artigos. Dos 339 artigos, 205 (60%) declararam explicitamente o uso de repositórios de Dados Abertos, especificamente *GenBank* (202) e *figshare* (3), enquanto em 134 deles isso não aconteceu. Foi preciso realizar consultas diretas nos repositórios para identificar que outros 110 registros tinham conjuntos de dados depositados em *Dryad* ou em *Zenodo*. Consequentemente, a quantidade de conjuntos de Dados Abertos criados ou utilizados por pesquisadores brasileiros deve ser maior, caso se realizem consultas em fontes adicionais.

Um olhar sobre as áreas de pesquisa mostra 52 delas envolvidas com Dados Abertos. A Tabela 4 apresenta um recorte com as primeiras vinte áreas e os comportamentos quantitativos nas categorias identificadas. Os resultados sugerem que o depósito/uso de repositórios de Dados Abertos predomina nas áreas de Ciência & Tecnologia, Parasitologia e Genética & Hereditariedade. Particularmente, *GenBank*, base pública de sequências de Ácido Desoxirribonucleico (ADN) mantida pelo Instituto Nacional de Saúde dos Estados Unidos, destaca nas áreas de Parasitologia, Microbiologia, Ciências Veterinárias e Genética & Hereditariedade. Já *Dryad*, um repositório disciplinar

internacional com foco em publicações científicas e médicas, é priorizado nas áreas de Ciência & Tecnologia, Ciências Ambientais & Ecologia e Biologia Evolucionária. Por sua vez, *Zenodo*, um repositório de Acesso Aberto de uso geral (conjuntos de dados, *software*, relatórios) operado pela Organização Europeia para Pesquisa Nuclear (CERN), foi mais utilizado em Zoologia, enquanto *figshare*, repositório de dados e outros materiais, foi reportado apenas na área de Ciência & Tecnologia.

A publicação de *data papers* foi observada, principalmente, na área de Biodiversidade & Conservação. Chama atenção que algumas áreas de pesquisa intensivas em dados, por exemplo, Astronomia ou Física, que se destacam no contexto da produção científica brasileira indexada na WoS, não aparecem nessa lista, ficando explícita a necessidade de outros estudos que permitam um aprofundamento sobre as práticas de Dados Abertos nessas áreas.

Por outro lado, identificaram-se 126 instituições brasileiras que implementaram práticas de Dados Abertos, destacando-se aí a Universidade de São Paulo, que aparece em 14,75% dos artigos que utilizam GenBank. Com práticas similares, mas em menor proporção, aparecem a Universidade Federal do Rio de Janeiro (8,85%), o Instituto Oswaldo Cruz - Fiocruz e a Universidade Federal de Rio Grande do Sul (5,60% cada uma). Já a Universidade Estadual Paulista "Júlio de Mesquita Filho" (6,49%) utiliza de forma equilibrada *GenBank* e *Dryad*.

**Tabela 4.** Comportamento de práticas de Dados Abertos por áreas de pesquisa (2015-2018).

Áreas de pesquisa	Criação, depósito ou uso de conjuntos de Dados Abertos	Uso de Repositórios de Dados Abertos					<i>Data papers</i>
		Total	<i>GenBank</i>	<i>figshare</i>	<i>Dryad</i>	<i>Zenodo</i>	
<i>Science &amp; Technology - Other Topics</i>	47	47	8	5	30	4	1
<i>Parasitology</i>	47	40	40	0	0	0	0
<i>Genetics &amp; Heredity</i>	41	35	24	0	10	1	0
<i>Microbiology</i>	29	27	27	0	0	0	0
<i>Tropical Medicine</i>	28	23	23	0	0	0	0
<i>Veterinary Sciences</i>	28	25	25	0	0	0	0
<i>Environmental Sciences &amp; Ecology</i>	26	24	3	0	18	3	2
<i>Zoology</i>	25	23	13	0	4	6	0
<i>Plant Sciences</i>	24	22	18	0	3	1	0
<i>Evolutionary Biology</i>	24	24	6	0	18	0	0
<i>Biochemistry &amp; Molecular Biology</i>	20	14	14	0	0	0	0
<i>Infectious Diseases</i>	18	16	16	0	0	0	0
<i>Agriculture</i>	15	13	13	0	0	0	0
<i>Biotechnology &amp; Applied Microbiology</i>	11	8	8	0	0	0	0
<i>Biodiversity &amp; Conservation</i>	10	10	4	0	2	4	7
<i>Life Sciences &amp; Biomedicine - Other Topics</i>	10	11	5	0	6	0	0
<i>Virology</i>	9	8	8	0	0	0	0
<i>Computer Science</i>	7	1	1	0	0	0	0
<i>Marine &amp; Freshwater Biology</i>	7	6	6	0	0	0	0
<i>Entomology</i>	7	6	5	0	0	1	0

Fonte: Elaborado pelos autores (2019).

Nota: Há dupla contagem nos totais obtidos por soma, a qual se explica pelo fato de alguns artigos terem sido publicados em periódicos indexados em mais de uma área e, portanto, serem computados uma vez em cada uma delas.

## Considerações Finais

O objetivo deste estudo foi explorar algumas práticas de Ciência Aberta da comunidade acadêmica brasileira, a partir do exame dos artigos publicados e indexados na WoS no período 2015-2018, especificamente aquelas relacionadas à publicação e depósito de artigos em canais de Acesso Aberto (periódicos e repositórios), bem como aquelas associadas ao uso e depósito de Dados Abertos.

Constatou-se o crescimento estável da produção em Acesso Aberto, representando mais de um terço das publicações brasileiras indexadas na WoS, no período. Predomina a utilização de periódicos de Acesso Aberto cadastrados em DOAJ, que oferecem acesso imediato e são distribuídos com licença *Creative Commons*. O estudo permitiu observar o peso importante de dois *mega journals* (*PloS One* e *Scientific Reports*) e também a relevância dos periódicos nacionais de Acesso Aberto.

Outrossim, o estudo permitiu constatar o depósito em repositórios de Acesso Aberto, fundamentalmente, de versões previamente publicadas em revistas de Acesso Aberto. Seria interessante analisar em outros estudos os benefícios (ou potenciais malefícios) dessa duplicidade. Diferenças disciplinares e institucionais também foram constatadas no que diz respeito à publicação em periódicos de Acesso Aberto e ao depósito em repositórios, por exemplo, o predomínio do uso de periódicos de Acesso Aberto na área de Agricultura, enquanto em Ciência & Tecnologia o uso de periódicos e repositórios é equilibrado. Embora o período investigado tenha sido curto, foi possível observar uma tendência de crescimento dessas práticas em várias áreas.

Com relação às práticas que envolvem a utilização/produção de Dados Abertos, o enfoque metodológico adotado tem como principais limitações o fato de que a identificação dos artigos dependeu de uma única fonte (WoS) e da explicitação dos termos vinculados a Dados Abertos no título, nas palavras-chave e na leitura dos resumos, verificados manualmente pelos autores. Essa limitação foi contornada parcialmente com a busca direta em dois dos repositórios, o que permitiu identificar outros artigos com conjuntos de dados associados. Sem dúvida, a interpretação dos resultados deverá levar em consideração as fontes selecionadas e os documentos coletados. Contudo, os resultados constituem uma primeira aproximação dessas práticas, tendo sido possível observar a sua tendência de crescimento, com a publicação de dados em revistas especializadas nesse tipo de material.

Futuros estudos poderão explorar outras práticas incluídas na taxonomia de Ciência Aberta, assim como aprofundar a pesquisa aqui realizada, utilizando outros métodos, como a análise detida dos artigos (por exemplo, da seção “metodologia”) por meio da interrogação direta dos pesquisadores, ou o uso de outras fontes para a coleta de dados.

## Agradecimentos

Ao Dr. Piotr Trzesniak, por seus comentários críticos durante a etapa inicial da coleta e análise dos dados; e aos Pareceristas anônimos, por suas observações e recomendações quando da avaliação do artigo.

## Colaboradores

A. CABALLERO-RIVERO e N. SÁNCHEZ-TARRAGÓ contribuíram na concepção e desenho da pesquisa, análise dos dados e redação do artigo. R.N.M. SANTOS colaborou na coleta, tabulação e análise dos dados. Todos os autores participaram da revisão do artigo e aprovaram a versão final.

## Referências

- Abdo, A.H. Guidelines for a contemporary, open academy. In: Albagli, S.; Maciel, M.L.; Abdo, A.H. (Ed.). *Open Science, open issues*. Brasília: IBICT, 2015. p.269-287.
- Albagli, S.; Maciel, M.L.; Abdo, A.H. *Open science, open issues*. Brasília: IBICT, 2015.
- Alperin, J.P. Open Access indicators: Assessing growth and use of Open Access resources from developing regions: The case of Latin America. In: Alperin, J.P.; Babini, D.; Fischman, G. (Ed.). *Open Access Indicators and Scholarly Communications in Latin America*. Buenos Aires: Unesco, 2014. p.15-80.
- Appel, A.L.; Albagli, S. The adoption of Article Processing Charges as a business model by Brazilian Open Access journals. *Transinformação*, v.31, e180045, 2019. Doi: <http://dx.doi.org/10.1590/2318-0889201931e180045>
- Budapest Open Access Initiative. *Read the Budapest Open Access Initiative*. Budapest: BOAI, 2002. Available from: <https://www.budapestopenaccessinitiative.org/read>. Cited: Mar. 8, 2019.
- Candela, L. et al. Data journals: A survey. *Journal of the Association for Information Science and Technology*, v.66, n.9, p.1747-1762, 2015.
- Chan, L.; Loizides, F. (Ed.). *Expanding perspectives on Open Science: Communities, cultures and diversity in concepts and practices*. Amsterdam: IOS Press, 2017.
- Chavan, V.; Penev, L. The data paper: A mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics*, v.12, n.15, p.1-12, 2011.
- Couldry, N.; Mejias, U. Data colonialism: Rethinking big data's relation to the contemporary subject. *Television and New Media*, v.20, n.4, p.336-349, 2018.
- Creaser, C. et al. Authors' awareness and attitudes toward Open Access repositories. *New Review of Academic Librarianship*, v.16, S1, p.145-161, 2010.
- Fundação de Amparo à Pesquisa do Estado de São Paulo. Análise da produção científica a partir de indicadores bibliométricos. In: Fundação de Amparo à Pesquisa do Estado de São Paulo. *Indicadores de ciência, tecnologia e inovação em São Paulo 2010*. São Paulo: Fapesp, 2011. v.1, Cap. 4.
- Genbank. *GenBank and WGS statistics*. Available from: <https://www.ncbi.nlm.nih.gov/genbank/statistics/>. Cited: Jul. 5, 2019.
- Haddaway, N.R. Open Synthesis: On the need for evidence synthesis to embrace Open Science. *Environmental evidence*, v.7, n.26, p.1-5, 2018.
- Hjørland, B. Informetrics needs a foundation in the Theory of Science. In: Sugimoto, C.R. (Ed.) *Theories of Informetrics and Scholarly Communication*. Berlin: Walter de Gruyter, 2016. p.20-46.
- Ioannidis, J.P.A. et al. Publication and other reporting biases in cognitive sciences: Detection, prevalence and prevention. *Trends in Cognitive Sciences*, v.18, n.5, p.235-241, 2014.
- MacIntyre, A. *After virtue: A study in moral theory*. Notre Dame, Indiana: University of Notre Dame Press, 2007. p.187
- Minniti, S.; Santoro, V.; Belli, S. Mapping the development of Open Access in Latin America and Caribbean countries: An analysis of Web of Science Core Collection and SciELO Citation Index (2005-2017). *Scientometrics*, v.177, n.3, p.1905-1930, 2018.
- Nosek, B.A.; Spies, J.R.; Motyl, M. Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, v.7, n.6, p.615-631, 2012.
- Packer, A.L. The metrics used in the development of the SciELO network model. In: Alperin, J.P.; Babini, D.; Fischman, G. (Ed.). *Open Access Indicators and Scholarly Communications in Latin America*. Buenos Aires: Unesco, 2014. p.81-96.
- Pontika, N. et al. Fostering Open Science to research using a Taxonomy and an eLearning Portal. In: IKnow: International Conference on Knowledge Technologies and Data Driven Business, October, 2015, Graz, Austria. *Proceedings [...]*. Walton Hall, UK: Open Research Online, 2015 Available from: [http://oro.open.ac.uk/44719/2/kmi\\_foster\\_iknow.pdf](http://oro.open.ac.uk/44719/2/kmi_foster_iknow.pdf). Acesso: 6 fev. 2019.
- Price, D.J.S. The structures of publication in science and technology. In: Gruber, H.; Marquis, D.G. (Ed.). *Factors in the transfer of technology*. Cambridge, MA: The MIT Press, 1969. p.91-104.
- Ramirez, M.; Hanlon, A. Asking for permission: A survey of copyright workflows for institutional repositories. *PORTAL: Libraries and the Academy*, v.11, n.2, p.683-702, 2011.
- Reilly, S. et al. *ODE Report on integration of data and publications: Opportunities for Data Exchange*. [S.l.]: ODE, 2011. Available from: [https://www.stm-assoc.org/2011\\_12\\_5\\_ODE\\_Report\\_On\\_Integration\\_of\\_Data\\_and\\_Publications.pdf](https://www.stm-assoc.org/2011_12_5_ODE_Report_On_Integration_of_Data_and_Publications.pdf). Cited: Feb. 10, 2019.
- Sánchez-Tarragó, N. et al. Las revistas científicas en América Latina hacia el camino del acceso abierto: un diagnóstico de políticas y estrategias editoriales. *Transinformação*, v.28, n.2, p.159-172, 2016. Doi: <http://dx.doi.org/10.1590/2318-08892016000200003>
- Santos, P.X.; Almeida, B.A.; Henning, P. (Org.). *Livro Verde - ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional*. Rio de Janeiro: Fiocruz, 2017.
- Santos, R.N.M.; Kobashi, N.Y. Bibliometria, cientometria, infometria: conceitos e aplicações. *Pesquisa Brasileira Ciência da Informação*, v.2, n.1, p.155-172, 2009.
- Scheliga, K.; Friesike, S. Putting open science into practice: A social dilemma? *First Monday*, v.19, n.9, 2014. Doi: <http://dx.doi.org/10.5210/fm.v19i9.5381>.
- US National Science Foundation. *Science and Engineering Indicators 2018*. Alexandria, VA: National Science Foundation, 2018. Available from: <https://www.nsf.gov/statistics/2018/nsb20181/assets/nsb20181.pdf>. Cited: Feb 6, 2019.
- Vicente-Saez, R.; Martinez-Fuentes, C. Open science now: A systematic review for an integrated definition. *Journal of Business Research*, v.88, p.428-436, 2018.
- Wallis, J.C.; Rolando, E.; Borgman, C.L. If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *Plos One*, v.8, n.7, p.1-17, 2013.