



# Encontros Bibli

## UMA ESTRATÉGIA PARA A IDENTIFICAÇÃO E ANÁLISE DE PUBLICAÇÕES CIENTÍFICAS EM PERIÓDICOS DE ACESSO ABERTO

A strategy for the identification and analysis of scientific publications in open access journals

**Patrícia Mascarenhas Dias**

Mestre em Modelagem Matemática e Computacional  
Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, Brasil  
patriciamdias@gmail.com

<https://orcid.org/0000-0002-8448-6874> 

**Thiago Magela Rodrigues Dias**


Doutor em Modelagem Matemática e Computacional  
Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, Brasil  
thiagomagela@cefetmg.br

<https://orcid.org/0000-0001-5057-9936> 

**Gray Farias Moita**

Doutor em Aeronáutica  
Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, Brasil  
gray@dppg.cefetmg.br

<https://orcid.org/0000-0002-6510-1019> 

A lista completa com informações dos autores está no final do artigo 

### RESUMO

**Objetivo:** propor uma estratégia para a identificação e análise de publicações científicas realizadas em periódicos de acesso aberto realizadas pelos indivíduos com currículos cadastrados na Plataforma Lattes.

**Método:** no intuito de identificar as publicações em periódicos de acesso aberto, dentre todo o conjunto de publicações registradas em todos os currículos cadastrados na Plataforma Lattes, inicialmente, todos os currículos foram extraídos e armazenados localmente. Após, como validador para classificar o periódico que está classificado como acesso aberto, a relação de periódicos do Directory of Open Access Journals (DOAJ) foi utilizada. Logo, com a verificação dos ISSN e e-ISSN de cada publicação com o conjunto obtido do DOAJ, as publicações foram categorizadas.

**Resultado:** com a técnica adotada foi possível analisar, dentre os aproximadamente sete milhões de artigos em periódicos dos indivíduos com currículos cadastrados na Plataforma Lattes, aqueles que foram publicados em periódicos de acesso aberto.

**Conclusões:** conclui-se que aproximadamente 27% das publicações foram realizadas em periódicos de acesso aberto. Além disso, também foi possível identificar com o conjunto caracterizado, quais são os principais autores tendo em vista a sua produção neste meio de divulgação, bem como, como algumas áreas do conhecimento tendem a ser mais representativas.

**PALAVRAS-CHAVE:** Plataforma Lattes. Acesso Aberto. Produção Científica. Periódicos Científicos.

### ABSTRACT

**Objective:** propose a strategy for the identification and analysis of scientific publications carried out in open access journals carried out by individuals with curricula registered in the Lattes Platform.

**Methods:** in order to identify publications in open access journals, among the whole set of publications registered in all resumes registered in the Lattes Platform, initially all resumes were extracted and stored locally. Then, as a validator to classify the journal that is classified as open access, the Directory of Open Access Journals (DOAJ) list of journals was used. Therefore, with the verification of the ISSN and e-ISSN of each publication with the set obtained from the DOAJ, the publications were categorized.

**Results:** with the technique adopted, it was possible to analyze, among the approximately 7 million articles in journals of individuals with curricula registered in the Lattes Platform, those that were published in open access journals.

**Conclusions:** it is concluded that approximately 27% of the publications were made in open access journals. In addition, it was also possible to identify with the characterized set, who are the main authors in view of their production in this means of dissemination, as well as how some areas of knowledge tend to be more representative. **KEYWORDS:** Lattes Platform. Open Access. Scientific production. Scientific Journals.

## 1 INTRODUÇÃO

O tradicional formato impresso de comunicação científica vem, aos poucos, dando espaço para os novos formatos eletrônicos, devido à ascensão da tecnologia de informação e comunicação. No contexto das pesquisas e estudos bibliométricos, a comunicação científica surge, nos dias atuais, como um elemento central em diversos níveis de discussão. Logo, o periódico científico surge como um importante mecanismo para a comunicação de resultados de pesquisas.

Mueller (1999) afirma que o periódico científico desempenha pelo menos quatro funções essenciais: certificação da ciência com o aval da comunidade científica, canal de comunicação entre os cientistas e de divulgação mais ampla da ciência, arquivo ou memória científica e registro da autoria da descoberta científica.

Já Mueller (2006) destaca que o periódico científico se caracteriza como o canal mais desejado e eficiente para a divulgação de resultados de pesquisas, além de ser aquele que dá maior visibilidade ao pesquisador que divulga seus resultados, tendo em vista que tal meio é aceito e validado pela comunidade científica. Para o autor, embora o periódico científico não seja, mesmo nos dias atuais, o veículo preponderante em todas as áreas do conhecimento, as revistas científicas indexadas e arbitradas pelos pares tornaram-se o centro do sistema tradicional de comunicação científica. Seu *status* de canal preferencial foi concedido pela comunidade científica, que lhe delega a atribuição de confirmar a autoria da descoberta científica e conferir prestígio aos cientistas e aos periódicos científicos por meio de um sistema de avaliação baseado em indicadores, como quantidade de publicações, índices de citação e visibilidade internacional.

De acordo com vários estudos, os periódicos – principalmente os disponíveis em formato eletrônico – estão em crescimento desde a última década. Pode-se afirmar que os periódicos, em todas as áreas do conhecimento, têm o papel de ser um filtro para o reconhecimento dos trabalhos que foram aceitos. Para Rodrigues e Oliveira (2012), a publicação em uma revista reconhecida pela área é a forma mais aceita para registrar a originalidade do trabalho e para confirmar que os trabalhos foram confiáveis o suficiente para superar o ceticismo da comunidade científica.

Nesse contexto, nos primeiros anos do século XXI, o Movimento de Acesso Aberto, cuja definição é “tornar disponível a qualquer internauta ler, fazer download, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral de artigos ou utilizá-los para outros fins legais sem qualquer barreira, desde que o trabalho seja devidamente reconhecido e citado”, estimulou o surgimento de periódicos nesse formato (LETA; COSTA; MENA-CHALCO, 2017).

Neubert, Rodrigues e Goulart (2012) afirmam que o acesso aberto assume um importante papel em todo o contexto da atividade científica, pois permite ao pesquisador ter acesso aos resultados de outros estudos sem as barreiras de custo e as dificuldades de acesso, além de promover a visibilidade e a divulgação dos resultados das atividades científicas de cada pesquisador e de cada universidade.

A publicação científica em acesso aberto faz parte de um cenário mais amplo em prol da abertura do conhecimento em geral – acesso aberto, dados abertos, recursos educacionais abertos, *software* livre, licenças abertas – e constitui, essencialmente, um movimento em direção à concepção da informação e do conhecimento como bens públicos (FURNIVAL; SILVA-JEREZ, 2017).

Segundo Freire (2011), limitar a reprodução e a disseminação de um material na sociedade traz consequências negativas para ambos, uma vez que, se não houver contato com a obra, o autor não é reconhecido pela sociedade e ela não compartilha dos potenciais benefícios advindos das novas descobertas.

Tendo em vista que grande parte das pesquisas científicas no país é financiada com recursos públicos, geralmente, em instituições de ensino ou centros de pesquisa públicos, é de se esperar que os resultados de tais estudos sejam divulgados sem nenhum tipo de barreira, principalmente financeira.

Diante disso, compreender como as publicações de um determinado conjunto de pesquisadores vêm sendo realizadas em periódicos de acesso aberto possibilita identificar um panorama do estágio atual desse tipo de comunicação no Brasil. Permite, ainda, verificar se em determinadas áreas do conhecimento esse tipo de publicação tende a ser mais frequente.

Estudos como este se caracterizam como um importante mecanismo para avaliar a evolução das publicações em periódicos de acesso aberto pelos pesquisadores brasileiros, permitindo verificar se as políticas de incentivo à publicação de pesquisas nesse formato de comunicação têm alcançado resultados satisfatórios.

## 2 REVISÃO DE LITERATURA

No trabalho de Silva e Alcará (2008), os autores analisaram as políticas de acesso aberto à informação científica e as propostas de ação, com ênfase nas iniciativas governamentais em diferentes países. Foi identificado que o movimento de acesso livre à informação científica já era preocupação oficialmente registrada em vários países, embora com diferentes graus de desenvolvimento. Entre tais diferenças, estão as próprias determinações das políticas, já que algumas obrigam instituições públicas e pesquisadores a disponibilizarem em acesso aberto os resultados de suas pesquisas, enquanto outras apenas sugerem o envolvimento e a participação desses pesquisadores e instituições no movimento.

O trabalho de Freire (2011) propôs um sistema de gestão da produção científica no Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio de repositório digital de acesso aberto. Para tanto, recorre-se à literatura científica pertinente para a compreensão do processo de comunicação científica e sua formalização em canais competentes. Recorre-se, também, aos documentos institucionais do CNPq e da Coordenação de Aperfeiçoamento de Pessoal em Nível Superior (Capes) para identificar elementos de reconhecimento dessas agências quanto à disseminação e ao acesso à informação. Além disso, são utilizados os bancos de dados da Thomson Reuters (ISI), do Directory of Open Access Journal (DOAJ) e do Qualis da Capes, visando comparar a qualidade das revistas de acesso aberto, considerando o fator de impacto da base ISI e dos estratos do Qualis da Capes como critérios para essa análise. Finalmente, aplica-se um questionário aos pesquisadores bolsistas de produtividade em pesquisa do CNPq, visando conhecer suas razões para publicar artigos científicos, o nível de conhecimento sobre acesso aberto à informação científica e sua disposição em adotar os canais abertos para disseminação dos resultados de suas pesquisas.

Já em Chalhub e Pinheiro (2011), são identificados os principais canais de comunicação científica de acesso aberto utilizados por pesquisadores e analisados os fatores intervenientes na adesão ao autoarquivamento de sua produção científica. O trabalho tinha como objetivo identificar os principais canais de comunicação científica em acesso aberto utilizados por pesquisadores de universidades públicas do estado do Rio de Janeiro. Foi utilizada a listagem de 47 comitês de assessoramento do CNPq para bolsas de produtividade em pesquisa e efetuada uma amostragem probabilística

estratificada por área de conhecimento, seguindo a divisão por comitê de assessoramento (Ciências Agrárias, Ciências Biológicas, Ciências Exatas e da Terra, Ciências da Saúde, Ciências Humanas, Ciências Sociais Aplicadas, Engenharias, e Linguística Letras e Artes). A partir da seleção dos pesquisadores contemplados pelo programa de bolsa de produtividade em pesquisa do CNPq no ano de 2010, cuja relação está disponível no site desse órgão federal, foram identificados aqueles vinculados a universidades públicas do estado do Rio de Janeiro com cursos de pós-graduação *stricto sensu*. Após a identificação dos endereços eletrônicos dos selecionados, foi enviada correspondência contendo em anexo o formulário com questões fechadas e abertas sobre as seguintes categorias: comportamento informacional, publicação de acesso aberto e adesão a repositório institucional.

No trabalho de Pavan e Barbosa (2017), foi analisada a existência de uma política nacional de financiamento público para a publicação de artigos em acesso aberto, examinando 29 agências de fomento à pesquisa. Para isso, com auxílio de questionários e análises dos sites institucionais, foi possível identificar que só uma minoria das agências possui uma política para financiar custos com as taxas de publicação dos artigos. Isso sugere que é necessária uma reavaliação do sistema de fomento brasileiro e a criação de políticas estratégicas.

Visando explorar o cenário nacional e internacional e, dessa forma, apresentar uma investigação que busque uma solução tecnológica para efetivar o acesso aberto a dados de pesquisa, Pavão, Rocha e Gabriel Junior (2018) propõem uma metodologia dividida em cinco etapas:

1. Identificação de práticas de acesso aberto a dados de pesquisa em instituições brasileiras;
2. Mapeamento de seus usuários e suas necessidades;
3. Proposta de um portal *web* para reunir a comunidade nacional;
4. Levantamento dos serviços e soluções tecnológicas existentes no cenário internacional para o compartilhamento de dados de pesquisa;
5. Proposição de recomendações para o apoio à criação de repositórios de dados de pesquisa em instituições nacionais e à sua agregação a uma rede de pesquisa em acesso aberto a dados de pesquisa.

Como resultado, são propostas iniciativas e estratégias internacionais para a criação de repositório de dados de pesquisa e para a criação de comunidades de práticas em torno do assunto.

No trabalho de Pavan e Barbosa (2018), as autoras analisam documentos publicados em acesso aberto entre os anos de 2012 e 2016 por autores que tenham afiliação brasileira e identificam o perfil dessas publicações. Para isso, dados de 930 periódicos e 63.847 documentos foram coletados da WoS. Destaca-se, ainda, que a produção científica brasileira em acesso aberto é caracterizada por um perfil endógeno, e que ainda se fazem necessárias políticas para estimular a publicação de artigos em acesso aberto, principalmente em periódicos internacionais.

Além desses trabalhos, diversos outros estudos têm analisado a produção científica em canais de acesso aberto, como Swan (2008), Rodrigues e Oliveira (2012), Leta, Costa e Mena-Chalco (2017) e Zhang e Watson (2018).

Importante destacar que a grande maioria dos trabalhos que avaliam o movimento de acesso aberto não tem como objeto principal de estudo as publicações nesse tipo de formato, e sim os repositórios ou os periódicos em acesso aberto. Logo, apesar de os trabalhos apresentados neste texto serem importantes para compreender as iniciativas existentes e as opiniões dos pesquisadores brasileiros, bem como os principais repositórios em acesso aberto no Brasil, um estudo abrangente sobre as publicações em periódicos de acesso aberto e os pesquisadores brasileiros que têm publicado trabalhos neste meio de comunicação de forma ampla se faz necessário.

### **3 METODOLOGIA**

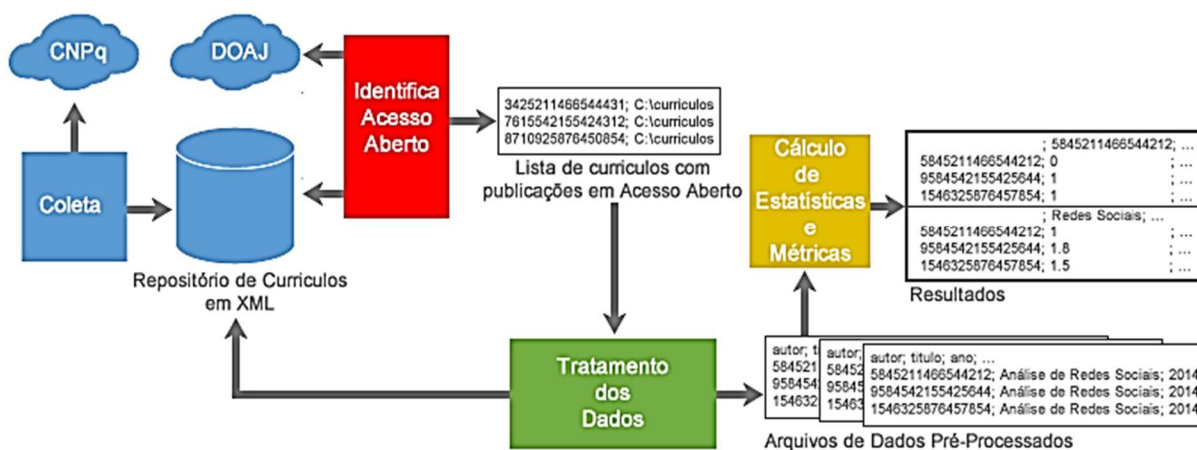
Para o processo de extração dos dados para as análises a serem realizadas no contexto deste trabalho, foram utilizados os dados curriculares da Plataforma Lattes do CNPq. Grande parte dos editais de financiamento de projetos de pesquisa, realizados por diversos órgãos de fomento, utiliza dados cadastrados nos currículos dos proponentes como uma das formas de avaliação das propostas. Logo, há um grande incentivo para que os pesquisadores mantenham as informações de seus currículos atualizadas. Tal fato torna os currículos da Plataforma Lattes uma excelente fonte de dados para análises. Por essa mesma razão, vários trabalhos têm utilizado a Plataforma Lattes como fonte de dados para diversos estudos sobre temas diversos,

como redes de colaborações científicas, análises sobre produtividade, genealogia acadêmica, entre outros.

Considerando que a maioria dos trabalhos correlatos analisou apenas grupos específicos de indivíduos, e tendo em vista que a manipulação de grandes quantidades de currículos da Plataforma Lattes não é uma tarefa trivial, já que existem problemas que envolvem recuperação de informação e algoritmos eficientes para manipulação de grandes volumes de dados, utilizamos o LattesDataXplorer (DIAS, 2016), um *framework* para extração e tratamento dos dados.

Conforme já explanado, um currículo cadastrado na Plataforma Lattes pode conter diversas informações capazes de auxiliar na compreensão da evolução da ciência brasileira sob diversas perspectivas. No entanto, para atender aos propósitos deste trabalho, somente dados de publicações de artigos em periódicos de acesso aberto foram considerados. Diante disso, foi proposta uma extensão do LattesDataXplorer com a inclusão de componentes, *a priori*, inexistentes, que avaliasse, para cada artigo publicado em periódico (a saber, 6.985.179), de cada um dos indivíduos (a saber, 5.901.161),<sup>1</sup> se o periódico no qual aquele artigo havia sido publicado era de acesso aberto (Figura 1). Logo, com a proposta dessa extensão, somente os autores e as publicações em periódicos de acesso aberto puderam ser analisados.

Figura 1 – LattesDataXplorer estendido



Fonte: elaborada pelos autores.

<sup>1</sup> Dados coletados em outubro de 2018.

Inicialmente, utilizando o LattesDataXplorer, foram coletados todos os currículos cadastrados na Plataforma Lattes em outubro de 2018, que foram armazenados no repositório local. Depois, o componente desenvolvido e denominado “Identifica acesso aberto” foi utilizado para recuperar todos os periódicos de acesso aberto registrados no portal do DOAJ, diretório *on-line* que indexa e fornece acesso a periódicos com acesso aberto. Em fevereiro de 2019, o DOAJ indexava 12.324 periódicos e 3.513.782 artigos. O DOAJ tem sido fonte de dados e referencial sobre periódicos de acesso aberto para diversos estudos, como em Swan (2008), Bjork e Salomon (2015), Costa (2018) e Sarvo (2018).

Logo, coletados os dados dos periódicos no portal do DOAJ em outubro de 2018, mesmo período da coleta dos currículos para as análises apresentadas no presente trabalho, foram recuperados 12.171 títulos de periódicos de acesso aberto, contendo dados como título, ISSN e eISSN, entre outras informações.

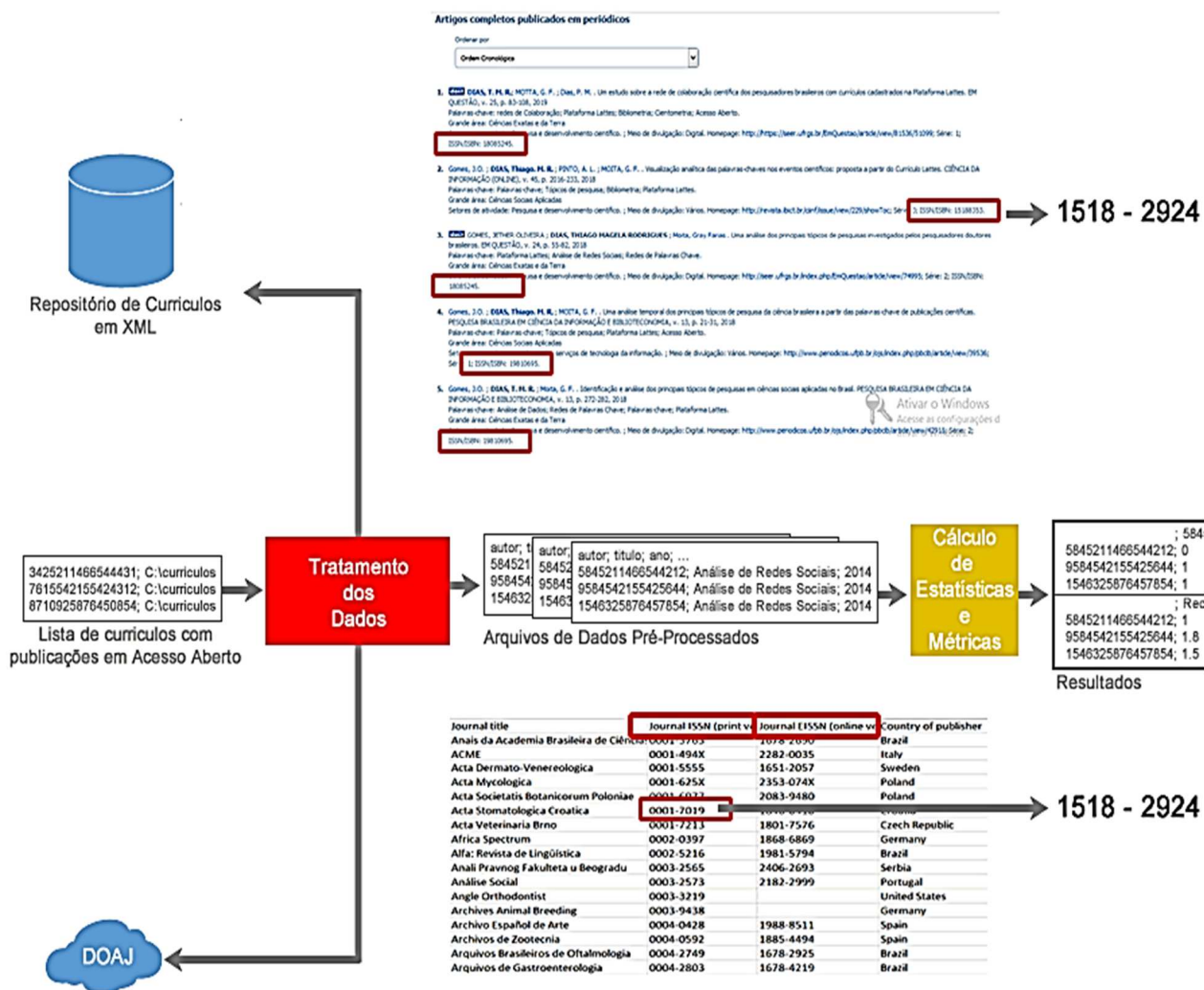
Com todo o conjunto de currículos cadastrados na Plataforma Lattes armazenado no “Repositório de currículos em XML”, foram aplicados os métodos propostos e implementados no componente “Identifica acesso aberto” para identificar os pesquisadores que têm publicado em periódicos de acesso aberto no Brasil.

Visando otimizar o máximo possível o processamento computacional dos currículos, toda vez que uma publicação cujo ISSN ou eISSN do periódico estivesse contido na relação de periódicos em acesso aberto extraída do DOAJ, imediatamente o identificador do currículo em análise era inserido na lista de currículos em acesso aberto, e o próximo currículo do conjunto em análise era avaliado.

Após a análise de todos os currículos que compõem o repositório local, uma lista contendo todos os currículos com publicações em acesso aberto é gerada, e ela se torna a base para o componente “Tratamento de dados”, que passou a incorporar os métodos propostos neste trabalho (Figura 2).



Figura 2 – Método para identificação de publicações em periódicos de acesso aberto



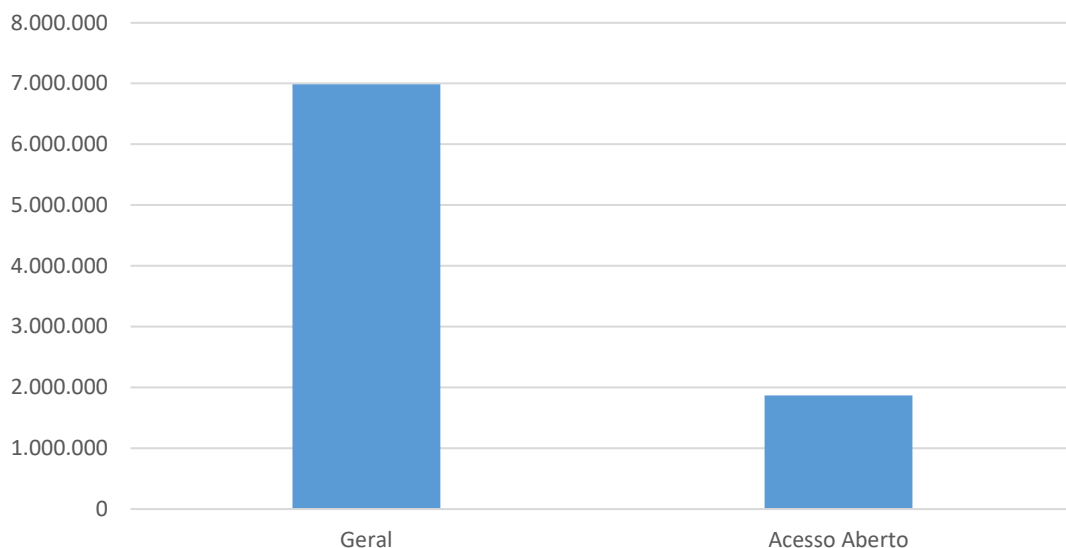
Fonte: elaborada pelos autores.

De posse da relação dos currículos que possuem artigos em acesso aberto, a identificação das publicações é realizada com o processamento dos currículos, utilizando o módulo de “Tratamento dos Dados” do LattesDataXplorer, no intuito de gerar os arquivos de dados pré-processados que sumarizam informações de interesse e que irão servir de base para o cálculo das métricas.

Além de dados gerais sobre os pesquisadores com publicações em acesso aberto, que compõem alguns dos arquivos, como dados sobre formação acadêmica, áreas de atuação, orientações e atuação profissional, foi analisado cada um dos artigos registrados na seção “Artigos completos publicados em periódicos” de cada currículo contido na “Lista de currículos com publicações em acesso aberto”. Para cada artigo de cada currículo, foi verificado e analisado se o ISSN ou o eISSN da

publicação estava presente na relação de periódicos recuperados do DOAJ, utilizando como meio de comparação os mesmos campos. Dessa forma, foi possível identificar todo o quantitativo de artigos em periódicos em acesso aberto (Figura 3).

Figura 3 – Quantitativo das publicações em periódicos registradas nos currículos



Fonte: elaborada pelos autores.

Como pode ser observado, do conjunto total de artigos publicados em periódicos, considerando todo o histórico de publicações registrado em todos os currículos cadastrados na Plataforma Lattes (6.985.179 publicações), um percentual de 26,76% (1.869.585) foi publicado em periódicos de acesso aberto, levando em conta a relação dos periódicos recuperados do DOAJ. Esse percentual de publicações em acesso aberto é relevante, acima de tudo, por considerar todo o histórico de publicações de cada pesquisador. Percebe-se que as publicações em periódicos de acesso aberto vêm recebendo atenção e adesão por parte dos pesquisadores ano a ano, apresentando-se como uma tendência na disseminação e na comunicação científica, sobretudo nos últimos anos.

## 4 RESULTADOS

Para os propósitos deste trabalho, utilizando a extensão proposta para o LattesDataXplorer, foram identificados todos os autores com currículos cadastrados na Plataforma Lattes que publicaram pelo menos um artigo em periódico de acesso aberto (370.431). Esses autores, apesar de serem uma pequena quantidade de

indivíduos em relação a todo o conjunto cadastrado na Plataforma Lattes (6,27%), possuem uma grande representatividade ao considerar o total de publicações de artigos em periódicos (aproximadamente 76%), conforme a Figura 4.

Figura 4 – Quantitativo das publicações em periódicos registradas



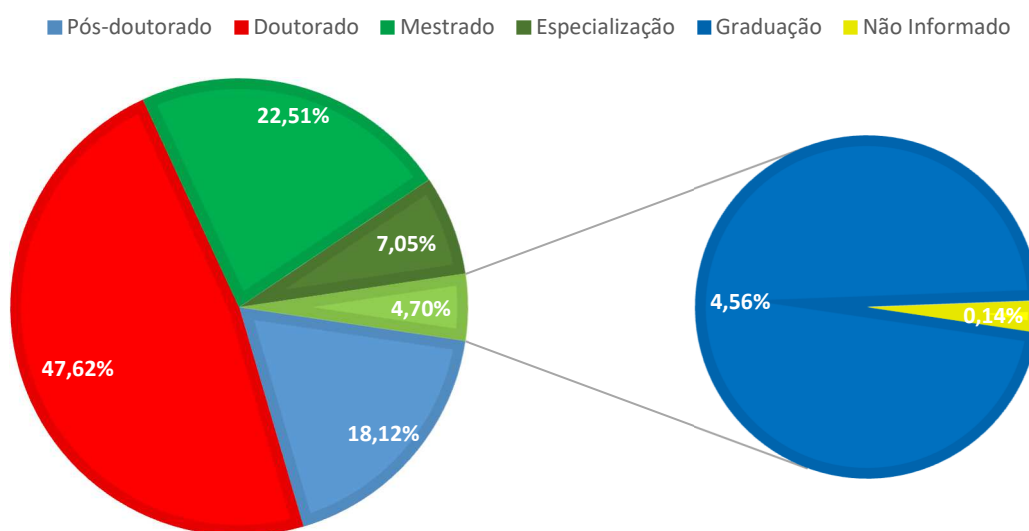
Fonte: elaborada pelos autores.

Ressalta-se que esse percentual de autores (6,27%) é bem inferior ao quantitativo de artigos em periódicos de acesso aberto, que representam, aproximadamente, 27% do total geral de publicações de artigos em periódicos de todos os indivíduos. Esse percentual de publicação está muito próximo do apresentado por Furnival e Silva-Jerez (2017), que destacam que apenas cerca de 30% do total de artigos científicos publicados no mundo anualmente estão disponível em canais de acesso aberto. Logo, identifica-se, aqui, que o percentual de publicações em periódicos de acesso aberto no Brasil está ligeiramente menor do que a média mundial de publicações nesse formato.

Uma informação importante para compreender o perfil dos autores que já publicaram em periódicos de acesso aberto diz respeito a seu nível de formação acadêmica. Tal informação é importante, pois possibilita compreender o perfil

acadêmico dos autores que têm publicado em periódicos desse formato no país. Tendo em vista que, em um currículo, é possível registrar todo o histórico de formação concluída e em andamento de um indivíduo, do ensino fundamental até o pós-doutorado, na Figura 5, é apresentada a distribuição dos autores com publicações em periódicos de acesso aberto de acordo com seu maior nível de formação acadêmica concluído.

Figura 5 – Distribuição dos autores por níveis de formação



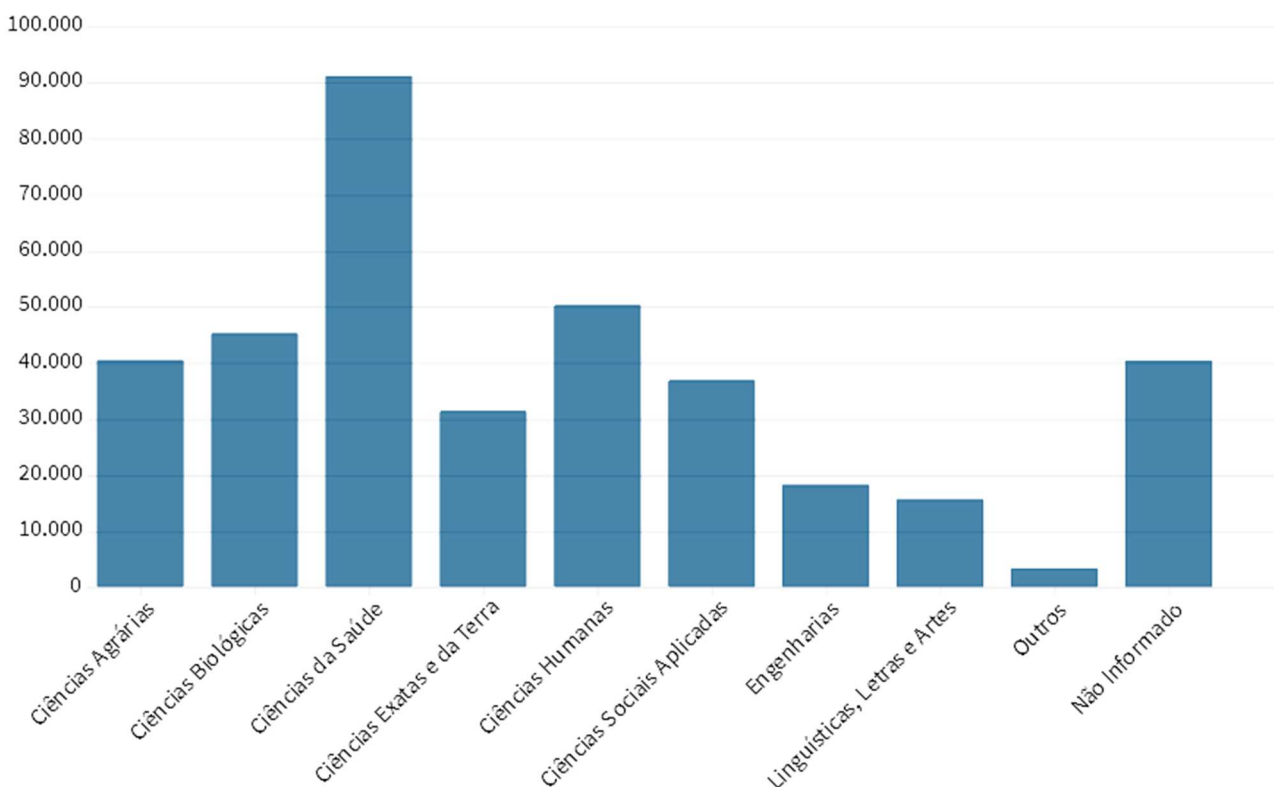
Fonte: elaborada pelos autores.

Como pode ser observado, apesar de o pós-doutorado não ser um nível de formação, os autores que o fizeram foram separados do conjunto de doutores para fins de estatística e melhor visualização. Percebe-se que a maioria dos autores possui doutorado concluído (47,62%). Em segundo lugar, estão os autores com mestrado (22,51%), seguidos dos que concluíram o pós-doutorado (18,12%), de modo que somente esses três grupos totalizam 88,25% do conjunto analisado. Além disso, observa-se que apenas 0,14% dos autores não informaram níveis de formação concluídos em seus currículos, e que 4,56% possuem apenas graduação concluída. A ausência e baixa representatividade de autores nos níveis mais inferiores de formação acadêmica pode ser justificada pelo fato de que as publicações científicas, principalmente aquelas realizadas em periódicos, são, em geral, resultado de pesquisas realizadas na pós-graduação.

Diante disso, fica evidente que a maioria do conjunto de autores identificados que já publicou trabalhos em periódicos de acesso aberto é composta por indivíduos com alto nível de formação acadêmica. Isso justifica sua significativa representatividade no conjunto total de publicações de artigos em periódicos, mesmo que seja um pequeno percentual do conjunto total de indivíduos com currículos cadastrados na Plataforma Lattes.

Ao analisar os autores das publicações realizadas em periódicos de acesso aberto pelas suas grandes áreas de atuação, é possível verificar sua distribuição por essas áreas, bem como quais delas têm maior representatividade, considerando a quantidade de indivíduos que já publicaram pelo menos um trabalho nesse meio de publicação (Figura 6).

Figura 6 – Distribuição dos autores pelas suas grandes áreas de atuação



Fonte: elaborada pelos autores.

Ressalta-se que o primeiro registro identificado em cada um dos currículos foi utilizado para determinar a principal grande área de atuação dos autores, tendo em vista que, nos currículos da Plataforma Lattes, é possível realizar o registro de até três

grandes áreas de atuação. Como pode ser observado, a grande área de Ciências da Saúde se destaca com a maior quantidade de indivíduos (24,53%), seguida pelas grandes áreas de Ciências Humanas (13,49%), Ciências Biológicas (12,14%), Ciências Agrárias (10,85%) e Ciências Sociais Aplicadas (9,89%). Já a menor quantidade de autores pertence à grande área de Linguística, Letras e Artes (4,17%), com quantidade próxima às Engenharias (4,87%).

Destaca-se, ainda, uma grande quantidade de indivíduos (10,81%) que não informou grande área de atuação em seus currículos, categorizada como “não Informado”. Na análise desse conjunto de indivíduos que não informou grande área, nota-se que a maioria deles é composta por autores que ainda estão em processo de formação e, possivelmente, ainda não definiram suas áreas de atuação. Além disso, também foi possível identificar um pequeno conjunto de indivíduos (0,81%) que informaram a grande área “outros”. Em geral, são indivíduos que não têm atuado em áreas tradicionais ou que têm realizado trabalhos interdisciplinares e, por esse motivo, não se definem como sendo de uma única área.

Considerando os currículos que possuem endereço profissional informado, é possível observar que a grande maioria dos indivíduos está concentrada nas regiões Sudeste e Sul, com destaque para o estado de São Paulo, que concentra aproximadamente 16% do conjunto de indivíduos, percentual bem superior ao de diversos outros estados. Em seguida, destacam-se os estados do Rio de Janeiro, Minas Gerais e Rio Grande do Sul. O somatório do quantitativo desses quatro estados é maior do que o somatório de todos os outros estados juntos. Esse cenário, não por acaso, também coincide com índices nacionais, que apresentam desigualdades semelhantes, como em níveis de escolaridade e de distribuição de renda. Além disso, uma hipótese para explicar tal distribuição pode ser a grande concentração de universidades e centros de pesquisas nesses estados. Além desses autores, foi possível identificar que 52 indivíduos do conjunto analisado declararam em seus currículos que estão atuando em instituições estrangeiras, sendo Colômbia – sete autores – e Portugal – cinco autores – as localidades mais representativas.

No intuito de melhor compreender o perfil das publicações em periódicos de acesso aberto realizadas pela comunidade científica brasileira, um importante elemento de análise é o idioma em que certas publicações foram realizadas. A informação referente ao idioma de uma determinada publicação é inserida pelo indivíduo no momento de registro do artigo, selecionando um idioma na lista

previamente cadastrada. Logo, erros ortográficos no registro dessa informação não existem, porém, por não ser um registro obrigatório, não é raro cadastros em que o autor não informa esse campo, deixando o artigo sem essa informação. Do conjunto total de artigos em periódicos de acesso aberto analisado neste trabalho, 85 artigos não possuem informação sobre o idioma da publicação.

Posteriormente, como era de se esperar, o idioma de maior representatividade é o português (63,76%), seguido, em menor quantidade, do inglês (35,22%). Em geral, as publicações em português são de artigos publicados em periódicos nacionais, sendo a maioria, justificando, dessa forma, a sua representatividade. Já os artigos escritos em inglês, em sua maioria, são de publicações em periódicos internacionais, sendo o inglês o idioma aceito pela grande maioria dos periódicos internacionais.

Em quantidade bem inferior, surgem o espanhol, francês e italiano. Ressalta-se, ainda, um total de 1.808 publicações que possuem outros idiomas informados, como alemão, chinês e japonês. Tais publicações sofrem, em parte, influência de colaborações de pesquisadores fluentes nesses idiomas.

Logo, devido à predominância de publicações em português, faz-se importante destacar as limitações que tais publicações podem sofrer quanto à visibilidade pela comunidade científica internacional, possibilitando análises específicas sobre esse tema para diversos outros estudos.

## 5 CONCLUSÕES

Com o intuito de traçar um retrato sobre a produção de artigos em periódicos de acesso aberto pelos pesquisadores brasileiros, foi necessário desenvolver componentes que, incorporados ao *LattesDataExplorer*, pudessem viabilizar as análises realizadas neste projeto. Assim, todo o repositório de dados curriculares da Plataforma Lattes foi analisado, possibilitando um estudo sobre a produção brasileira de artigos em periódicos de acesso aberto utilizando também dados do DOAJ, que tem sido objeto de estudo de diversos outros trabalhos, conforme apresentado.

Conforme apresentado, os métodos computacionais são eficientes e possibilitam a identificação de autores que já publicaram artigos em periódicos de acesso aberto, bem como as publicações realizadas nesse meio de divulgação. Destaca-se que tal estratégia pode ser aplicada a outros conjuntos de publicações sem a necessidade de grandes adaptações na estratégia.

Considerando os dados analisados, o conjunto de artigos publicados em periódicos de acesso aberto tem como autores um total de 370.431 indivíduos, o que representa, aproximadamente, 6% do conjunto total de indivíduos com currículos cadastrados na Plataforma Lattes. Ressalta-se que esse percentual de autores é bem inferior ao percentual de artigos em periódicos de acesso aberto, que representam, aproximadamente, 27% do total geral de publicações em periódicos, considerando todos os indivíduos. Esse percentual está muito próximo do apresentado por Furnival e Silva-Jerez (2017), que destacam que apenas em torno de 30% do total de artigos científicos publicados no mundo anualmente estão disponíveis em canais de acesso aberto.

Além disso, em uma análise inicial dos dados, foi possível perceber que o conjunto de indivíduos com publicações em periódicos de acesso aberto possuem, em geral, alto nível de formação acadêmica e estão atuando principalmente nos estados da região sudeste do país. Ao considerar suas grandes áreas de atuação, destacam-se algumas grandes áreas, como a de Ciência das Saúde, a de Ciências Humanas e a de Ciências Biológicas, grandes áreas essas que são influenciadas por algumas áreas de atuação.

Diante do exposto, percebe-se a importância deste tipo de estudo, já que possibilita classificar um grande conjunto de publicações em periódicos de acesso aberto de forma automática e, ainda, poderá proporcionar diversas novas pesquisas que visam analisar, de forma ampla, a produção de artigos em periódicos de acesso aberto no Brasil.

## REFERÊNCIAS

BJÖRK, Bo-christer; SOLOMON, David. Article processing charges in OA journals: relationship between price and quality. **Scientometrics**, [s. l.], v. 103, n. 2, p. 373-385, mar. 2015.

CHALHUB, Tânia; PINHEIRO, Lena Vânia Ribeiro. **Acesso aberto à informação científica no Brasil**: um estudo das universidades públicas do estado do Rio de Janeiro. Rio de Janeiro: 2011. Relatório Final de Atividades.

COSTA, Elaine Hipólito dos Santos. **Periódicos de acesso aberto: um canal de disseminação dos pesquisadores bolsistas PQ1A do CNPq?** 2018. Tese (Doutorado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.



DIAS, Thiago. Magela. Rodrigues. **Um estudo da produção científica brasileira a partir de dados da Plataforma Lattes**. 2016. Doutorado em Modelagem Matemática e Computacional) – Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, 2016.

FREIRE, José Donizetti. **CNPq e o acesso aberto à informação científica**. 2011. 275 f. Tese (Doutorado) – Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2011.

FURNIVAL, Ariadne Chloe Mary; SILVA-JEREZ, Nelson Sebastian. Percepções de pesquisadores brasileiros sobre o acesso aberto à literatura científica. **Informação & Sociedade**, João Pessoa, v. 27, n. 2, 2017.

LETA, Jacqueline; COSTA, Elaine Hipólito dos Santos; MENA-CHALCO, Jesús Pascual. Artigos em Periódicos de Acesso Aberto: um Estudo com Pesquisadores Bolsistas de Produtividade do CNPq. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [s. l.], v. 11, p.1-6, nov. 2017.

MUELLER, Suzana P. M. O círculo vicioso que prende os periódicos nacionais. **DataGramZero-Revista de Ciência da Informação**, n. dez/99, 1999.

MUELLER, Suzana P. M.. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, ago. 2006.

NEUBERT, Patricia da Silva; RODRIGUES, Rosângela Schwarz; GOULART, Luiza Helena. Periódicos da Ciência da Informação em acesso aberto: uma análise dos títulos listados no DOAJ e indexados na Scopus | Open access journals in information Science. **Liinc em Revista**, [s. l.], v. 8, n. 2, p. 389-401, dez. 2012.

PAVAN, Cleusa; BARBOSA, Márcia Cristina Bernardes. Financiamento público no Brasil para a publicação de artigos em acesso aberto: alguns apontamentos. **Em Questão**, Porto Alegre, v. 23, n. 2, p.120-145, abr. 2017.

PAVÃO, Caterina Groposo; ROCHA, Rafael Porte da; GABRIEL JUNIOR, Rene Faustino. Proposta de criação de uma rede de dados abertos da pesquisa brasileira. **Rdbci: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 16, n. 2, p.329-343, abr. 2018.

RODRIGUES, Rosângela Schwarz; OLIVEIRA, Aline Borges. Periódicos científicos na America Latina: títulos em Acesso Aberto indexados no ISI e SCOPUS. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 4, p.76-99, dez. 2012.

SARVO, Denilson D. Oliveira. **Avaliação da Plataforma Lattes como fonte de coleta de metadados para povoamento automatizado em repositórios institucionais**. 2018. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos, São Carlos, 2018.

SILVA, Terezinha Elizabeth da; ALCARÁ, Adriana Rosecler. Políticas de acesso aberto à informação científica: iniciativas governamentais. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais [...]**. São Paulo: Ancib, 2008.

SWAN, Alma. Why Open Access for Brazil? | Por que acesso livre no Brazil? **Liinc em Revista**, [s. l.], v. 1, n. 1, p.158-171, set. 2008.

ZHANG, Li; WATSON, Erin. The prevalence of green and grey open access: Where do physical science researchers archive their publications?. **Scientometrics**, [s. l.], v. 117, n. 3, p. 2021-2035, out. 2018.

## NOTAS

### CONTRIBUIÇÃO DE AUTORIA

**Concepção e elaboração do manuscrito:** P. M. Dias, T. M. R. Dias, G. F. Moita.

**Coleta de dados:** P. M. Dias, T. M. R. Dias.

**Análise de dados:** P. M. Dias, T. M. R. Dias, G. F. Moita.

**Discussão dos resultados:** P. M. Dias, T. M. R. Dias, G. F. Moita.

**Revisão e aprovação:** P. M. Dias, T. M. R. Dias, G. F. Moita.

### CONJUNTO DE DADOS DE PESQUISA

O conjunto de dados que dá suporte aos resultados deste estudo não está disponível publicamente.

### LICENÇA DE USO

Os autores cedem à **Encontros Bibli** os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a [Licença Creative Commons Attribution](#) (CC BY) 4.0 International. Esta licença permite que **terceiros** remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os **autores** têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

### PUBLISHER

Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Informação. Publicação no [Portal de Periódicos UFSC](#). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

### EDITORES

Enrique Muriel-Torrado, Raymundo N. Machado, Valdineia Barreto Ferreira, Kátia de Oliveira Rodrigues, Susane Barros e Genilson Geraldo.

### HISTÓRICO

Recebido em: 16-12-2020 – Aprovado em: 30-08-2021 – Publicado em: 08-10-2021.