

AGREGADORES DE DADOS ALTMÉTRICOS: ANALISANDO O ALTMETRIC.COM E O WEBOMETRIC ANALYST

Altmetric data aggregators: analyzing altmetric.com and webometric analyst

Vildeane da Rocha Borba

Doutora em Comunicação e Informação

Universidade Federal de Pernambuco, Departamento de Ciência da Informação, Recife, Brasil

vildeane.borba@ufpe.br

<https://orcid.org/0000-0002-2067-9802> 

Sônia Elisa Caregnato

Doutorado em Information Studies

Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

Professora Associada do Programa de Pós-Graduação em Comunicação e Informação

sonia.caregnato@ufrgs.br

<https://orcid.org/0000-0002-5676-2763> 

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

RESUMO

Objetivo: relata a experiência de utilização dos agregadores de dados altmétricos *Altmetric.com* e *do Webometric Analyst*, em um estudo sobre a presença dos artigos de revistas científicas brasileiras da Ciência da Informação nas plataformas de mídias sociais.

Método: o estudo foi caracterizado como descritivo, de natureza quali-quantitativa, cujo objeto de análise foram os agregadores *Altmetric.com* e *Webometric Analyst*. O corpus utilizado para testá-los constituiu-se de 13 periódicos brasileiros da área da Ciência da Informação, englobando 2.724 artigos científicos publicados no período de 2011 a 2018.

Resultado: a partir da utilização desses agregadores para coletar dados altmétricos sobre artigos de periódicos brasileiros em Ciência da Informação, pode-se constatar primariamente que distintas metodologias e ferramentas geram diferentes resultados, com discrepâncias e semelhanças pontuais. Percebeu-se também que os números de DOIs de alguns artigos de periódicos estavam incorretos, isto é, o link destes DOIs remetia a uma página de erro na *CrossRef*, presumindo-se que a baixa quantidade de resultados positivos de dados altmétricos no *Altmetric.com* via API deve-se a isso. Embora o *Webometric Analyst* tenha apresentado uma maior cobertura de dados altmétricos para os artigos de periódicos brasileiros em Ciência da Informação, possivelmente pela forma de busca de dados utilizada, neste caso autor, título e ano, foram percebidas algumas inconsistências no Mendeley, como um artigo com diferentes versões ou com datas de publicação distintas, mostrando divergentes contagens de leitores para um mesmo artigo.

Conclusões: considera-se que as ferramentas utilizadas para gerar dados altmétricos devem ser entendidas como facilitadores da coleta de dados em plataformas de mídias sociais, porém deve-se estar atento à abrangência e limitação de cada uma dessas ferramentas, assim como da coleção escolhida para teste e captura de dados. Como a natureza das plataformas é desigual, elas podem ser utilizadas de modo complementar em estudos altmétricos.

PALAVRAS-CHAVE: Altmetria. Agregadores de dados altmétricos. API do *Altmetric.com*. *Webometric Analyst*.

ABSTRACT

Objective: reports the experience of using altmetric data aggregators *Altmetric.com* and *Webometric Analyst*, in a study on the presence of articles from Brazilian scientific journals on Information Science on social media platforms.

Methods: the study was characterized as descriptive, of qualitative and quantitative nature, which analyzed the aggregators *Altmetric.com* and *Webometric Analyst*. The corpus used to test them consisted of 13 Brazilian journals in Information Science, encompassing 2,724 scientific articles published in the period from 2011 to 2018.

Result: from the use of the two aggregators to collect altmetric data about articles from Brazilian journals in Information Science, it was primarily verified that different methodologies and tools generate different results, with discrepancies and specific similarities. It was also noticed that the DOI numbers of some articles were incorrect, that is, the link of these DOIs pointed to an error page in CrossRef, assuming that was the reason for the small number of positive results in the *Altmetric.com* via API. Although the *Webometric Analyst* has shown greater coverage of altmetric data for articles from Brazilian journals in Information Science, possibly due to the search strategy used, in this case author, title and year, some

inconsistencies were noticed in Mendeley, such as an article with different versions or with different publication dates, showing divergent counts of readers for the same article.

Conclusions: it is considered that the tools used for generating altmetric data should be understood as facilitators of data collection on social media platforms, however one must be aware of the scope and limitations of each of these tools, as well as of the collection chosen for testing and capture. As the nature of the platforms is uneven, they can be employed as to complement one another.

KEYWORDS: Altmetric. Altmetric data aggregators. Altmetric.com API. Webometric Analyst.

1 INTRODUÇÃO

O desenvolvimento das tecnologias digitais, em especial da *internet*, tem modificado a forma como se cria, gerencia e interage com a informação. A *web*, em especial, influencia diferentes práticas comunicacionais, tais como quando se conecta, modifica, compartilha, discute, visualiza e/ou “curte” (*like*) em plataformas de mídias sociais, também chamadas de sites de rede social *on-line*, como Facebook e Twitter.

Dentre as diversas informações discutidas e disseminadas em plataformas de mídias sociais, a informação científica chama a atenção quando publicações acadêmicas, sejam elas artigos de revisão, originais, capítulos de livros, anais de eventos, teses e/ou dissertações, são divulgadas neste novo locus. Meschede e Siebenlist (2018) afirmam que o campo científico está operando no mundo conectado, na medida em que cientistas estão usando, apresentando e colaborando com outros por meio de portais e ferramentas de mídias sociais.

O efeito disso é que “[...] os resultados da pesquisa se tornaram mais diversificados e estão sendo cada vez mais comunicados e discutidos nas mídias sociais” (ERDT et al., 2016). Cabrera, Roy e Chisolm (2018) são ainda mais otimistas e afirmam que o novo uso de mídias sociais para fins acadêmicos revolucionou a academia em termos de seu alcance, usando novas ferramentas em meio digital como complemento das atividades tradicionais.

A utilização de mídias sociais para a difusão e visibilidade de conteúdos científicos fez surgir uma subárea de estudo no âmbito dos estudos métricos da informação, chamada altmetria. Ela utiliza as fontes de dados emergentes, isto é, as plataformas de mídias sociais, para contribuir na avaliação da pesquisa, complementando os indicadores tradicionais.

Percebe-se que essa nova subárea provocou mudanças na avaliação da ciência, sendo um complemento em relação à bibliometria, cientometria e webometria, notadamente por oferecer indicações em tempo real sobre a difusão de produtos de pesquisas que se diferenciam dos tradicionais, por sua cobertura abranger um público para além do ambiente acadêmico dito formal.

Nesse sentido, por ser uma área em desenvolvimento, a coleta e a análise de dados relativos à produção científica advindos de plataformas de mídias sociais têm sido um desafio. Um dos aspectos que necessita ser estudado em maior profundidade é o papel dos agregadores de dados, ferramentas que reúnem dados alométricos e das quais derivam indicadores.

Assim, este artigo tem o objetivo de relatar a experiência de utilização dos agregadores de dados alométricos Altmetric.com e do Webometric Analyst em um estudo sobre a presença dos artigos de revistas científicas brasileiras da Ciência da Informação (CI) nas plataformas de mídias sociais. Ressalta-se que o objeto deste artigo são os agregadores de dados acima descritos e não os dados alométricos das revistas, embora esses permitam avaliar aqueles.

Pretende-se, com este estudo, evidenciar questionamentos sobre quais os agregadores de dados alométricos seriam os mais indicados para levantamento de indicadores alométricos, levando em consideração as abrangências, limitações e variações de cobertura de dados.

2 ALTMETRIA

A origem do termo “altmetria” relaciona-se à utilização da *hashtag* “#altmetrics” em um tweet por Jason Priem, em 2010. Posteriormente, uma publicação dele com colaboradores, conhecida como Manifesto da Altmetria, estabeleceu o potencial da altmetria para melhorar a comunicação científica e a avaliação da pesquisa, diferenciando-a dos indicadores tradicionais e definindo-a como a descoberta de vestígios invisíveis de impacto acadêmico, por meio da observação de atividades em ferramentas e sistemas *online* (ERDT et al., 2016; HAUSTEIN; BOWMAN; COSTAS, 2016; PRIEM et al., 2010; SANKAR; KAVITHA, 2016).

A altmetria visa complementar os indicadores tradicionais a partir de dados abertos de fontes advindas de plataformas de mídias sociais. Tais mídias foram desenvolvidas a partir da *web* 2.0, que está inserida nos preceitos na ciência aberta. Segundo Bartling e Friesike (2014, p. 10), a ciência aberta “refere-se a uma cultura científica que se caracteriza por sua abertura, ao compartilhamento de resultados para um público amplo”.

Além disso, diferentemente das métricas tradicionais, a altmetria é capaz de quantificar a disseminação para um amplo grupo de interessados e demonstrar engajamento público, o que pode ser relevante para o impacto final de um produto,

mostrando a oportunidade de avaliar o impacto acadêmico em nível de usuário, o que tem sido reconhecida por várias organizações acadêmicas. Apesar de ser uma temática nova e emergente, a altmetria tem se mostrado de grande importância como complemento aos indicadores tradicionais e, principalmente, pelo aumento exponencial da erudição digital (CABRERA; ROY; CHISOLM, 2018).

Assim, a altmetria pressupõe que o impacto científico seja mensurado em um sentido mais amplo, menos erudito, abrangendo todos os produtos de pesquisa, não apenas a análise de citações tradicionais, incluindo, também, outros ambientes, que incorporam comunidades científicas e eventos fora do meio acadêmico (HAUSTEIN; BOWMAN; COSTAS, 2016).

Trata-se de um modelo de democratização do gerenciamento de informações, que pode aumentar a velocidade e o alcance da tradução do conhecimento, não apenas no contexto da academia, mas também entre o público em geral, tornando-se a característica mais importante das ferramentas de mídias sociais (CABRERA; ROY; CHISOLM, 2018).

Sugimoto e demais autores (2017) afirmam que, nos últimos anos, as discussões sobre a altmetria centraram-se no papel das mídias sociais para o aumento da visibilidade de estudiosos, que oferecem novos veículos para a disseminação, como também no aumento da demanda por entidades administradoras e financiadoras de indicadores de atividade científica e tecnológica, especialmente no que diz respeito ao valor da pesquisa para um público mais amplo.

3 AGREGADORES DE DADOS ALTMÉTRICOS

Os agregadores de dados altmétricos são definidos como “ferramentas e plataformas que agregam e oferecem eventos on-line como métricas derivadas de provedores de dados altmétricos” (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2016, p. 70). Alguns provedores de dados, isto é, plataformas de mídias sociais, como Twitter, Facebook e Mendeley, acompanham eventos *on-line* em suas plataformas através de métricas de uso, visualização ou *download*, mas não coletam esses dados de outras plataformas. Os agregadores de dados mais populares e que coletam dados de diversas fontes são: Impactstory, Plum Analytics, PLOS ALM, Altmetric.com e Webometric Analyst. Esses dois últimos são as ferramentas de análise de dados exploradas neste artigo.

As estruturas de coleta de dados dos agregadores altmétricos se apresentam de diferentes formas. Os agregadores primários são aqueles que rastreiam os dados

altmétricos diretamente dos provedores de dados, isto é, das plataformas de mídias sociais, tendo como exemplo o Altmetric.com e o PLOS ALM. Os agregadores secundários são aqueles que coletam os dados altmétricos de agregadores primários, como o Impactstory, Plum Analytics e Webometric Analyst e os agregadores terciários são os que recolhem tanto de agregadores primários quanto de secundários, como é o caso do Snowball Metrics (ERDT et al., 2016).

O Altmetric.com foi desenvolvido em 2011 por Euan Adie e lançado em fevereiro de 2012, sendo, a partir desse mesmo ano, financiado pela Digital Science, uma empresa de tecnologia que oferece soluções para a pesquisa. Pesquisadores afirmam que o Altmetric.com é um dos *frameworks* emergentes indiscutivelmente mais populares que rastreia, analisa e coleta atividades *on-line* contínuas em torno de resultados de pesquisa (ADIE; ROE, 2013; ERDT et al., 2016; MUKHERJEE; SUBOTIC; CHAUBEY, 2018).

O Altmetric.com monitora menções a livros, capítulos de livros, artigos de periódicos, apresentações, teses e dissertações, relatórios, anais de eventos, revisões, conjunto de dados, literatura cinzenta, entre outros, com o propósito principal de apresentar o texto original de cada menção e, em alguns casos, também os dados demográficos sobre o autor da menção.

Esse agregador tem interfaces específicas para atender a diferentes públicos, entre eles editores, instituições, pesquisadores, financiadores e para pesquisas em desenvolvimento, contribuindo para o monitoramento, navegação, pesquisa e avaliação em torno da pesquisa publicada em fontes não convencionais (DIGITAL SCIENCE, c2018). As ferramentas disponíveis abrangem o Altmetric explorer for publishers Institutions e Funders, Altmetric for books, além do Altmetric Badges, o Altmetric API e Free tools.

A ferramenta Altmetric API (Application Programming Interface) fornece acesso aos dados métricos dos resultados de pesquisa coletados pelo Altmetric.com via API. Tal interface pode ser utilizada através de identificadores digitais e tipos de licenças, entre elas a API livre, a API livre com chave e API comercial. Entre os produtos disponibilizados pelo Altmetric.com, alguns são livres, como o Altmetric Bookmarklet, Institutional Repository badges, Explorer for Academic Librarians, Badges for Individual Researchers e API for Research, este último utilizado nesta pesquisa.

Com a publicação da recomendação NISO RP-25-2016, *Outputs of the Alternative Assessment Metrics Project*, como resultado de projeto financiado pela Alfred P. Sloan Foundation, que buscou estabelecer um consenso sobre altmetria entre partes interessadas, foi possível verificar relatórios sobre os códigos de conduta dos agregadores

de dados altmétricos, e em específico do *Altmetric.com*. Da mesma forma, foi possível ampliar a compreensão sobre como cada métrica é definida, como os dados são gerados e mantidos, seu monitoramento, além de todas as limitações reportadas (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2016).

O Webometric Analyst é um *software* gratuito projetado para realizar análises automáticas variadas na *web* para fins de pesquisa em ciências sociais. Ele pode criar diagramas de rede de coleções de *sites*, estimar o impacto *on-line* de coleções e recuperar informações em larga escala sobre *blogs* e vídeos do Youtube (THELWALL, 2009). O programa é baseado em Windows e possibilita análises para altmetria, de citações, da *web social* e webometria, incluindo análise de *link*. Ele baixa dados da *web* por meio de APIs ou *download* direto e inclui uma ampla gama de opções de processamento (STATISTICAL CYBERMETRICS RESEARCH GROUP, [2019]).

Para as métricas alternativas e análise de citações, o Webometric Analyst processa dados do Mendeley, Altmetric.com, Google books e Worldcat, já para análise da *web social*, o programa processa dados do Youtube, Twitter, Goodreads e Flickr, entre outros, através de interface distinta para cada tipo de fonte de coleta a ser realizada. A estrutura de entrada de *query* no *software* é diferenciada de acordo com o provedor de dados escolhido para análise (THELWALL, 2009).

Os processos envolvidos em qualquer tarefa do Webometric Analyst abrangem as etapas de entrada de dados, que deve ser realizada a partir de um arquivo de texto simples, podendo ser criado usando o bloco de notas do Windows ou em um processador de texto, escolhendo a opção de salvar o documento como um arquivo de texto simples. Para análise de produções científicas no Mendeley, por exemplo, é necessária uma lista de artigos organizados em arquivo de texto simples e salvos no formato separado por tabulações, a saber: <Blank –tab- title – tab- authors –tab- year>.

O Webometric Analyst pode recorrer a motores de pesquisa, a *web*, Technorati ou YouTube para obter os seus dados, e a sua análise pode ser conduzida dependendo da tarefa, que pode requer a combinação de uma sequência de funções do Webometric Analyst, às vezes até exigindo a coleta de dados adicionais. O *software* finaliza com a apresentação dos resultados através de arquivos de dados brutos, tabelas de resumo e diagramas de rede (STATISTICAL CYBERMETRICS RESEARCH GROUP, [2019]).

Os arquivos de dados brutos são arquivos de texto simples que contêm os dados extraídos de uma ou mais fontes, apresentados em colunas para facilitar o carregamento em uma planilha ou banco de dados para processamento. As tabelas de resumo são

arquivos de texto simples de colunas de dados ou páginas da *web* contendo tabelas que resumem o resultado do processamento dos dados brutos. Os diagramas de rede são imagens tiradas do processamento de dados brutos. Esses diagramas podem ser salvos como arquivos – por exemplo, em formato Pajek – e podem ser editados na aparência e no conteúdo antes de imprimir e salvar (THELWALL, 2009).

4 METODOLOGIA

O estudo foi caracterizado como descritivo, de natureza quali-quantitativa, cujo objeto de análise foram os agregadores Altmetric.com e Webometric Analyst. O *corpus* utilizado para testá-los constituiu-se de 13 periódicos brasileiros na área da Ciência da Informação (CI), englobando 2.724 artigos científicos publicados no período de 2011 a 2018 que disponibilizavam seus DOIs.

A escolha dos periódicos da Ciência da Informação (CI) deveu-se em função de compreender como a altmetria está se desenvolvendo na CI brasileira e como as plataformas de mídias sociais estão sendo utilizadas pela comunidade da CI, percebendo a disseminação e interações de produtos de informação científica e seus atores como os principais elementos que compõe esse novo modelo de comunicação científica, ressaltando, mais uma vez, que o objeto deste artigo são os agregadores de dados, não os dados altmétricos das revistas.

O *corpus* de periódicos, organizado a partir do Qualis Capes (2013-2016) foi: para o extrato A1, *Informação e Sociedade: estudos (I&S)*; *Perspectivas em Ciência da Informação (PCI)* e *Transinformação*. Para o extrato A2, *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação (Encontros Bibli)*, *Em Questão e Informação & Informação (I&I)* e, para o extrato B1, *Brazilian Journal of Information Science (Brajis)*, *Ciência da Informação (RCI)*, *Revista de Ciência da Informação e Documentação (INCID)*, *Liinc em Revista (LIINC)*, *Perspectivas em Gestão & Conhecimento (PG&C)*, *Ponto de Acesso* e *Revista Digital de Biblioteconomia e Ciência da Informação (RDBCI)*.

A API aberta do Altmetric.com possibilita ao usuário fazer a pesquisa de dados altmétricos a partir da pesquisa pelo DOI do artigo. No entanto, para realização de estudos altmétricos, que necessitam buscar um quantitativo de DOIs relativamente alto, a coleta individual no Altmetric.com é inviável.

Nesse sentido, foi desenvolvido um aplicativo *web* que permitiu a pesquisa automática de uma coleção maior de DOIs, cujos resultados fossem apresentados de uma

forma melhor estruturada para posterior tabulação, comparação e análise. A aplicação web foi desenvolvida em angular e NodeJS com o banco de dados Cloud Firestore e foi ligado diretamente na API gratuita do Altmetric.com, possibilitando a pesquisa pela entrada de um arquivo de planilha com os números de DOIs organizados por linha.

A aplicação web possui uma arquitetura simples de informação, tendo como ferramenta principal para consulta a aba “selecionar arquivo”, que, a partir de uma planilha com cada DOI por linha, na primeira coluna da primeira planilha, sem cabeçalho, insere os dados. A aplicação web registra os números absolutos e percentuais de artigos que obtiveram registros altmétricos e o quantitativo dos que não foram localizados, mostrando os dados descritivos (número de DOI, título e nome do periódico). O resultado da consulta é exportado em arquivo de planilha, conforme mostra Figura 1.

Figura 1 – Tela de resultados da aplicação web de dados altmétricos, a partir da API aberta do Altmetric.com: teste com DOIs do periódico em questão

#	Nº DOI	Título	Periódico da Publicação
1	10.19132/1808-5245242.12-36	Aspectos éticos da coautoria em publicações científicas	Em Questão
2	10.19132/1808-5245241.267-295	Em nome da moral e dos bons costumes: censura a livros com temática de g...	Em Questão
3	10.19132/1808-5245240.258-277	Cenários da dinâmica de hiper crescimento e colapso das revistas científicas b...	Em Questão
4	10.19132/1808-5245241.296-302	A altmetria na prática e o papel dos bibliotecários no seu uso e aplicação	Em Questão
5	10.19132/1808-5245233.228-249	Análise bibliométrica da produção científica brasileira e do nordeste em Biote...	Em Questão
6	10.19132/1808-5245233.179-206	Uma proposta de taxonomia para a categorização das falhas no ambiente da...	Em Questão
7	10.19132/1808-5245233.276-299	Para cada universo colecionista, suas fontes: dinâmicas informacionais nas tr...	Em Questão
8	10.19132/1808-5245232.167-185	O catálogo da biblioteca e o linked data	Em Questão
9	10.19132/1808-5245231.129-153	Representação semântico-discursiva de cibercordéis	Em Questão
10	10.19132/1808-5245230.213-234	Incipiência da visualização de indicadores bibliométricos e altmétricos nos Re...	Em Questão

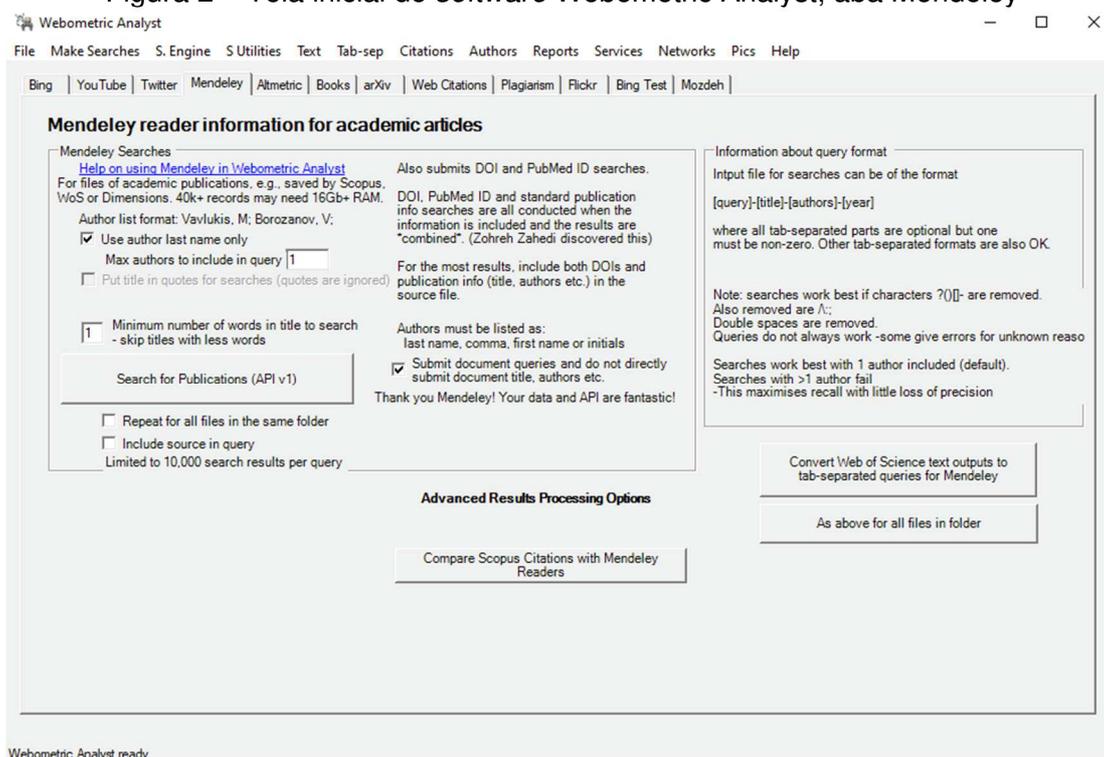
Fonte: dados de pesquisa (2019).

O Webometric Analyst foi baixado do site do grupo de pesquisa Statistical Cybermetrics, sendo executado no sistema operacional Windows. A entrada de dados foi realizada através de arquivo de notas (.txt), contendo as informações de título, autor e ano, em que pelo menos um desses elementos não fosse nulo, <Blank –tab- title – tab- -authors –tab-year>, conforme mostra a Figura 2, da tela inicial do software. Para esse software, não é obrigatória a existência do DOI do artigo.

Para a estrutura de entrada do formato do arquivo no Webometric Analyst, isto é, título, autor e ano, deve-se excluir os caracteres “?”, “[]”, “()”, “^” ou espaços duplos, no intuito de evitar erros. Também é necessária a retirada de todos os acentos, pois a codificação do *software* foi configurada para o inglês.

As escolhas de utilização dessas duas ferramentas deveram-se ao Mendeley ter sido a mídia com o maior número de registros altmétricos na maioria dos periódicos pesquisados oriundos do Altmetric.com. Porém, esse agregador apresenta dados quantitativos de leitores, sem demonstração de aspectos complementares (demográficos e geográficos). Nesse sentido, o Webometric Analyst possibilita conhecer, além da quantidade de artigos com marcação no Mendeley, a quantidade de marcações por usuário, o perfil ocupacional dos leitores, as áreas do conhecimento, as nacionalidades dos leitores e os artigos científicos que obtiveram o maior número de leitores.

Figura 2 – Tela inicial do *software* Webometric Analyst, aba Mendeley



Fonte: dados de pesquisa (2019).

5 RESULTADOS E ANÁLISES

A visualização dos dados pela ferramenta Altmetric.com é destacada a partir do conjunto de 73 metadados que descrevem diversas variáveis a serem analisadas, evidenciando 28 deles, como pode ser visto no Quadro 1. Tais dados permitem o

levantamento, verificação e investigação de informações dos artigos científicos compartilhados em plataformas de mídias sociais.

Quadro 1 – Metadados utilizados para análise dos dados altmétricos extraídos do *Altmetric.com*

METADADOS	EXPLICAÇÃO DOS DADOS
title	Título da publicação
doi:	DOI
tq: [Título da menção
uri:	URL do identificador capturado
issns: []	ISSN
journal:	Nome da Revista
cohorts: {	Grupos de usuários que mencionaram a publicação no Twitter
abstract:	Resumo do artigo
authors: []	Nome dos autores
publisher_subjects: []	Assuntos da publicação pelo editor
context_journal_count	Número de publicações na Revista
cited_by_fbwalls_count:,	Número de páginas compartilhadas no Facebook
cited_by_feeds_count:	Número de blogs que mencionaram a publicação
cited_by_gplus_count:	Número de contas que compartilharam no Google+
cited_by_msm_count:	Número de fontes de notícias que mencionaram a publicação
cited_by_posts_count:	Número de menções
cited_by_tweeters_count:	Número de contas do twitter que twitaram a publicação
cited_by_videos_count:	Número de canais do Youtube/Vimeo
cited_by_accounts_count:	Número de contas que mencionaram a publicação
score:	Pontuação do Altmetric Attention Score (AAS)
at:	AAS corrente
url:	URL da publicação
added_on:	Data quando foi capturada a primeira atenção altmétrica
published_on:	Data de publicação
10endeley:	Leitores no Mendeley
readers_count:	Total de leitores
images: {	Link para a versão png do donut
details_url:	URL para a página de detalhes relevantes do <i>Altmetric.com</i>

Fonte: Altmetric Support (c2018).

O Altmetric.com coleta dados do Mendeley, Twitter, Facebook, *blogs*, YouTube/Vimeo, Google+, CiteULike, Reddit, Conotea e de outras fontes de dados, como *sites* de notícias jornalísticas, documentos governamentais, entre outros, além de contabilizar diferentes eventos por plataformas de mídias sociais e de disponibilizar o *hiperlink* individual de cada artigo científico com os detalhes dos registros altmétricos no Altmetric.

O Webometric Analyst permite a pesquisa em diversas plataformas de mídias sociais, sendo utilizada especificamente a do Mendeley, por ter sido a mídia social que

apresentou o maior número de registros altmétricos no *corpus* de periódicos em Ciência da Informação (CI) aqui pesquisados. O Twitter foi a segunda mídia social que apresentou maior número de registros altmétricos, porém, o Webometric Analyst só realiza a consulta nos últimos 3.200 tweets mais recentes na plataforma do Twitter, tornando esse aprofundamento inviável para o *corpus* pesquisado.

Nesse sentido, o Webometric Analyst permite a visualização de detalhes obtidos do Mendeley, como número de usuários, local e disciplina, ao passo em que sua estrutura de busca permite uma abrangência maior de registros. Enfatiza-se que o Altmetric disponibiliza apenas o número de leitores por artigo no Mendeley, tendo o Webometric Analyst apresentado mais detalhes para além de número de leitores.

Em relação à quantidade de dados altmétricos encontrados nas duas ferramentas, na análise realizada com a API da Altmetric em abril de 2019, dos 13 periódicos brasileiros em Ciência da Informação (CI), foram encontrados registros altmétricos para oito deles. Dos 2.724 artigos com DOI pesquisados, apenas 209 artigos apresentaram registros altmétricos, o equivalente a 7,7% do conjunto pesquisado, enquanto 92,3% não indicaram registros na Altmetric, conforme mostra o Quadro 2. Enfatiza-se que só foi possível realizar a busca no *software* a partir de um identificador digital, ou seja, artigos sem identificadores digitais – sejam eles, DOI, PMID, ArXiv, ISBN – não são passíveis de recuperação utilizando a API da Altmetric, mesmo que potencialmente tenham registro altmétricos.

Quadro 2 – Quantidade de artigos com registros altmétricos coletados pelo agregador Altmetric via API e Webometric Analyst, por periódico brasileiro em Ciência da Informação (2011-2018)

QUALIS CAPES	PERIÓDICOS	Nº DE DOIs PESQUISADOS	Nº DE ARTIGOS COM REGISTROS NO ALTMETRIC	Nº DE ARTIGOS COM REGISTROS NO WEBOMETRIC ANALYST
A1	Informação & Sociedade: estudos	119	7	69
	Perspectivas em Ciência da Informação	369	34	358
	Transinformação	158	39	156
A2	Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação	277	10	219
	Em Questão	178	19	148
	Informação & Informação	354	21	257
B1	BRAJIS: Brazilian Journal of Information Science: research trends	140	0	53
	Ciência da Informação	212	0	136
	InCID: Revista de Ciência da Informação e Documentação	165	20	133
	LiINC em Revista	280	0	163
	Perspectivas em Gestão & Conhecimento	68	0	34

PontodeAcesso	157	0	96
RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação	247	59	200
TOTAL	2.724	209	2.022

Fonte: dados de pesquisa (2019).

Já a porcentagem de dados altmétricos para artigos marcados no Mendeley pelo Webometric Analyst atingiu 74,2%, isto é, nove vezes a mais que o resultado apresentado pela Altmetric, destacando-se que o primeiro apresenta dados apenas da mídia social Mendeley e o último abrange um conjunto maior de mídias sociais. Dos 2.724 artigos científicos publicados em periódicos brasileiros da área da Ciência da Informação com DOI, 2.022 artigos apresentaram registros altmétricos no Webometric Analyst. Acredita-se que a consulta ao Mendeley é feita pela Altmetric apenas se o artigo já tiver tido alguma repercussão em plataformas de mídia social. Assim, foram encontrados registros altmétricos no Mendeley pelo Webometric Analyst para todos os periódicos brasileiros da área da Ciência da Informação que fizeram parte desta pesquisa.

A partir da utilização das duas ferramentas acima relatadas em artigos de periódicos brasileiros em Ciência da Informação (CI), pode-se inferir, primariamente, que distintas metodologias e ferramentas em altmetria podem gerar diferentes resultados, com discrepâncias e semelhanças pontuais. Alguns trabalhos brasileiros mostraram esses diferentes resultados em periódicos brasileiros em CI devido à utilização de diferentes ferramentas.

Os estudos altmétricos tendo como objeto os artigos de periódicos em Ciência da Informação (CI) no Brasil utilizaram diferentes ferramentas e apresentaram resultados diferentes, por exemplo, o trabalho de Nascimento e Oddone (2014, 2015), realizado em 25 periódicos brasileiros da CI, utilizou como ferramenta o Altmetric explorer e empregou o título do periódico e ISSN para a busca. O resultado do trabalho mostrou registros altmétricos para os periódicos *PCI*, *Transinformação* e *RCI* nas plataformas de mídias sociais Mendeley, Twitter, Facebook e blogs, nessa ordem (NASCIMENTO; ODDONE, 2014, 2015). Enfatiza-se que, neste estudo, não foram encontrados registros altmétricos para a revista *Ciência da Informação* utilizando a API da Altmetric, apenas pelo Webometric Analyst no Mendeley.

Manhique e demais autores (2019) analisaram 12 periódicos da Ciência da Informação pelo DOI, usando a ferramenta Bookmarklet for Researchers – Altmetric no período de 2012 a 2017. Eles observaram dados altmétricos relativos a seis periódicos, dos

quais apenas dois coincidem com os periódicos deste estudo – *Em Questão* e *I&I* –, não encontrando registros para a *PCI*, *LiiNC*, *PontodeAcesso*, *Brajis* e *RCI*. Os provedores de dados que mostraram registros altmétricos foram o Mendeley, Twitter, Facebook e blogs, nessa ordem de grandeza.

Araújo (2015) apresentou uma análise de artigos de periódicos A1 em Ciência da Informação utilizando a ferramenta Altmetric Bookmarklet. O autor encontrou dados altmétricos para a *PCI* – n=5 artigos –, *Transinformação* – n=19 artigos – e *Ciência da Informação*, com n=21 artigos. A última, que na época era Qualis A1, apresentou o maior número de dados altmétricos, não retornando dados para a *I&S*.

Alguns estudos estrangeiros também relatam divergências de resultados e, conseqüentemente, dos agregadores de dados altmétricos. Bar-Ilan, Halevi e Milojević (2019) afirmam que há diferenças nas contagens altmétricas relatadas pelos agregadores de dados altmétricos Altmetric.com e PlumX, justificando que os agregadores podem cobrir diferentes fontes e podem usar diferentes maneiras de identificar menções, como, por exemplo, por DOI, PMID, arXivID ou ano, assim como podem ter cronogramas diferentes para atualização das fontes de dados primárias.

Os autores reiteram, indicando que a diferença do Altmetric.com em relação ao PlumX para o Mendeley pode ser devido ao fato de que Altmetric.com registra contagens de leitores do Mendeley somente se houver pelo menos um indicador altmétrico adicional para o documento. Portanto, se nada foi rastreado além de uma leitura de Mendeley, o Altmetric.com não o contará. Isso pode explicar a cobertura inferior do Altmetric.com em comparação com o PlumX (BAR-ILAN; HALEVI; MILOJEVIĆ, 2019).

Karmakar, Banshal e Singh (2021) ratificam os resultados quando declaram que a cobertura e média de menções em agregadores de dados altmétricos para o mesmo conjunto de artigos variam em diferentes plataformas. Nesse estudo, os autores informaram que o Altmetric.com registrou menções mais altas no Twitter e em *blogs*, e o agregador PlumX registrou menções mais altas no Facebook e Mendeley. Os autores apresentaram comparações entre os agregadores Altmetric.com e PlumX e enfatizam, no estudo, a limitação do Altmetric.com, por exemplo, na coleta de dados a partir do Mendeley.

Não obstante, Ortega (2018, 2020) assegura que a maioria dos estudos altmétricos utiliza o agregador de dados Altmetric.com, sendo esse agregador o serviço que mais captura menções de *sites* de mídia social, *blogs* e veículos de notícias. Já o PlumX é o que possui a melhor cobertura para leitores Mendeley, porém, é um agregador que abrange

menos eventos de mídias sociais. Tal resultado em relação ao PlumX pode ser comparado ao Webometric Analyst, neste artigo.

Zahedi, Fenner e Costas (2015) mencionam que o Altmetric.com oferece o melhor número de cobertura de dados no Twitter e Facebook comparado com o agregador Lagotto, que apresenta melhor cobertura para o CiteULike e Reddit. A cobertura para o Twitter pode ser devido ao fato que o Altmetric.com coleta menções *on-line* de artigos acadêmicos do Twitter em tempo real. Eles rastreiam *links* (URLs) para documentos dentro de tweets, o que inclui comentários públicos e retuítes, mas sem favoritos.

Os autores consideram que as principais razões para as diferentes contagens estão relacionadas aos diferentes métodos de coleta e processamento de métricas pelos diferentes fornecedores. Como cada provedor consulta de fontes, seja usando DOI, APIs ou outros metadados, podem existir possíveis atrasos na coleta de dados ou problemas de atualização (ZAHEDI; FENNER; COSTAS, 2015).

Constatou-se, também, que a ausência de identificadores digitais na totalidade de artigos de periódicos brasileiros em CI dificulta a coleta e padronização de dados altmétricos, levando em consideração que, atualmente, a ferramenta mais difundida para indicar dados altmétricos, o Altmetric, só pode ser realizar buscas através de identificadores digitais.

Nesse sentido, percebeu-se que os números de DOIs de alguns artigos de periódicos brasileiros em CI estavam incorretos, isto é, o *link* desses DOIs remetia a uma página de erro na CrossRef, presumindo-se que a baixa quantidade de resultados positivos de dados altmétricos no Altmetric via API deve-se a isso. Tal resultado foi verificado, principalmente, nos periódicos com Qualis Capes B1, isto é, no periódico *Brazilian Journal of Information Science: research trends* e *Ciência da Informação*, que obtiveram registros no Webometric Analyst oriundos do Mendeley, e não alcançaram registros no Altmetric, que também abrange essa plataforma de mídia social em seus resultados.

Embora o Webometric Analyst tenha apresentado uma maior cobertura de dados altmétricos para os artigos de periódicos brasileiros em CI, possivelmente pela forma de busca de dados utilizada – nesse caso, autor, título e ano –, foram percebidas algumas inconsistências no Mendeley, como um artigo com diferentes versões ou com datas de publicação distintas, mostrando divergentes contagens de leitores para um mesmo artigo. Os erros em DOIs e algumas inconsistências percebidas podem ter causado limitações e/ou discrepâncias nos resultados encontrados.

Bar-Ilan, Halevi e Milojević (2019, tradução nossa) afirmam que:

apesar de ser abrangente, o banco de dados do Mendeley, tem alguns problemas inerentes que afetam os dados que gera. Primeiro, os usuários conduzem o banco de dados em que cada usuário identifica uma publicação de seu interesse e a adiciona às suas bibliotecas pessoais. Isso, por sua vez, torna os dados mais sujeitos a erros do que bancos de dados com curadoria de terceiros. Em segundo lugar, muitas vezes existem erros que podem ser encontrados nos campos de metadados, que impedem o Mendeley de agregar corretamente as contagens de leitores, e muitas vezes há mais de um registro para um determinado item. Finalmente, ao usar a API Mendeley para recuperar contagens de leitores por DOI, apenas um único registro é recuperado (o que não é necessariamente o registro com o maior número de leitores), portanto, os números relatados pelos agregadores podem ser subestimados. A API Mendeley permite pesquisar por vários campos, como autor, título, fonte de publicação, ano de publicação e semelhantes, mas como os metadados são inseridos por usuários que não usam um formato predefinido, muitas vezes vários registros são criados para o mesmo item. Além disso, devido a possíveis erros de metadados, a consulta de pesquisa pode perder registros para o item fornecido. Ao consultar a API pelo título do artigo, Mendeley geralmente retorna vários registros para o mesmo artigo. Finalmente, os caracteres especiais nos campos de texto não funcionam bem no Mendeley.

Realizado esse levantamento e análise, percebe-se que o Webometric Analyst apresenta uma boa cobertura de dados altmétricos para publicações brasileiras em Ciência da Informação para a plataforma de mídia social Mendeley, especialmente pela forma de busca e coleta, porém, essa análise é restrita, levando em consideração que só foram obtidos registros do Mendeley.

Constatou-se, também, que a ausência de identificadores digitais na totalidade de artigos de periódicos brasileiros em CI dificulta a coleta e padronização de dados altmétricos, levando em consideração que, atualmente, a ferramenta mais difundida para indicar dados altmétricos, o Altmetric.com, só pode ser realizada através de identificadores digitais.

Percebe-se que, independentemente da área de conhecimento analisada, existem diferenças entre os agregadores de dados altmétricos pela forma de busca aos dados e da própria natureza dos agregadores. Zahedi e Costas (2018) afirmam que diferentes formas de acessar os dados de diversas plataformas de mídia social, junto com diferentes abordagens de coleta, processamento, resumo e atualização de métricas de mídia social causam diferenças substanciais nos dados e métricas oferecidos por agregadores de dados altmétricos.

6 CONCLUSÃO

O ambiente da comunicação científica tem sido reconfigurado, especialmente pelo uso, difusão e compartilhamento de produtos de pesquisa em plataformas de mídias sociais, e os estudos altmétricos têm se desenvolvido com o propósito de coletar os rastros invisíveis não mensurados pelos tradicionais sistemas de avaliação da pesquisa.

Constatou-se que existem divergências de resultados conforme o agregador ou ferramenta de dados altmétricos utilizados para a coleta e análise de estudos empíricos em altmetria, o que afeta as métricas derivadas deles. Além disso, pelos erros encontrados, observa-se que é necessário conscientizar os editores de periódicos sobre a necessidade de padronização dos metadados dos artigos e, especialmente, sobre a necessidade de alocar DOIs de forma apropriada. Deve-se também levar em consideração que as plataformas de mídias sociais não foram desenvolvidas com a finalidade de compartilhar e difundir resultados de pesquisa, motivo pelo qual a evolução e progressão da área ainda está sendo estabelecida.

As ferramentas utilizadas para coleta e análise de dados altmétricos devem ser entendidas como facilitadores de dados em plataformas de mídias sociais, porém, deve-se estar atento às abrangências e limitações dessas ferramentas, assim como a coleção escolhida para teste e captura, entendendo que a natureza das plataformas é divergente.

Espera-se que outros estudos sejam realizados, com dados de outras áreas do conhecimento, que abranjam outros agregadores de dados altmétricos, descrevendo e pontuando cada limitação e abrangência dos agregadores, contribuindo para a assertiva escolha de agregadores de dados e, conseqüentemente, de indicadores de dados altmétricos mais fidedignos para a comunidade científica.

REFERÊNCIAS

ADIE, E.; ROE; W. Altmetric: enriching scholarly content with article-level discussion and metrics. **Learned Publishing**, Bristol, v. 26, n. 1, p. 11-17, 2013.

ARAÚJO, R. F. Mídias sociais e comunicação científica: análise altmétrica em artigos de periódicos da ciência da informação. **Em Questão**, Porto Alegre, v. 21, n. 1, p. 96-109, jan./abr. 2015.

BAR-ILAN, J.; HALEVI, G.; MILOJEVIĆ, S. Differences between Altmetric Data Sources - A Case Study. **Journal of Altmetrics**, United Kingdom, v. 2, n. 1, p. 1-8, 2019.

BARTLING, S.; FRIESIKE, S. Towards Another Scientific Revolution. *In*: BATLING, S.; FRIESIKE, S. (ed.). **Opening Science**: the evolving guide on how the Internet is Changing Research, Collaboration and Scholarly Publishing. Heidelberg: Springer Open, 2014. p. 3-15.

CABRERA, D.; ROY, D; CHISOLM, M. S. Social Media Scholarship and Alternative Metrics for Academic Promotion and Tenure. **American College of Radiology**, Reston, v. 15, n. 1, p. 135-141, jan. 2018.

DIGITAL SCIENCE. Altmetric: Who's talking about your research? **Altmetric**, [s. l.], c2018. Disponível em: <https://www.altmetric.com>. Acesso em: 10 jun. 2021.

ERDT, M. *et al.* Altmetrics: an analysis of the state-of-the-art in measuring research impact on social media. **Scientometrics**, New York, v. 109, n. 2, p. 1117-1166, 2016.

HAUSTEIN, S.; BOWMAN, T. D.; COSTAS, R. Interpreting 'altmetrics': Viewing acts on social media through the lens of citation and social theories. *In*: SUGIMOTO, C. R. (ed.). **Theories of Informetrics and scholarly communication**. Frankfurt: De Gruyter, 2016. p. 372-406.

KARMAKAR, M.; BANSHAL, S. K.; SINGH, V. K. A large-scale comparison of coverage and mentions captured by the two altmetric aggregators: Altmetric.com and PlumX. **Scientometric**, New York, v. 126, p. 4465-4489, 2021.

MANHIQUE, I. L. E. *et al.* Indicadores altmétricos em periódicos brasileiros da Ciência da Informação no Brasil. **Revista Ibero-Americana de Ciência da Informação**, Brasília, DF, v. 12, n. 2, p. 515-532, maio/ago. 2019.

MESCHEDE, C.; SIEBENLIST, T. Cross-metric compatibility and inconsistencies of altmetrics. **Scientometrics**, New York, v. 115, n. 1, p. 283-297, 2018.

MUKHERJEE, B.; SUBOTIC, S.; CHAUBEY, A. K. And now for something completely different: the congruence of the Altmetric Attention Score's structure between different article groups. **Scientometrics**, New York, v. 114, n. 1, p. 253-275, 2018.

NASCIMENTO, A. G.; ODDONE, N. E. Uso de altmetrics para avaliação de periódicos científicos brasileiros em Ciência da Informação. **Ciência da Informação em Revista**, Maceió, v. 2, n. 1, p. 3-12, jan./abr. 2015.

NASCIMENTO, A. G. do; ODDONE, N. E. Uso de indicadores altmetrics na avaliação de periódicos científicos brasileiros em Ciência da Informação. *In*: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 4., 2014, Recife. **Anais [...]**. Recife: UFPE, 2014. p. 1-2.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **NISO RP-25-2016**: Outputs of the NISO Alternative Assessment Metrics Project. Baltimore: NISO, 2016. Disponível em: <https://www.niso.org/publications/rp-25-2016-altmetrics>. Acesso em: 10 jun. 2020.

ORTEGA, J.-L. Altmetrics data providers: A meta-analysis review of the coverage of metrics and publications. **El profesional de la información**, León, v. 29, n. 1, p. 1-23, 2020.

ORTEGA, J.-L. Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event Data. **Scientometrics**, New York, v. 116, p. 2123-2138, 2018.

PRIEM, J. *et al.* Altmetrics: A manifesto. **Altmetrics**, [s. l.], 2010. Disponível em: <http://altmetrics.org/manifesto>. Acesso em: 30 jul. 2017.

SANKAR, P.; KAVITHA, E. S. Altmetrics: new research tool in the growth social media. **E-Journal of Library and Information Science**, Madison, v. 3, n. 1, p. 80-91, Feb. 2016.

STATISTICAL CYBERMETRICS RESEARCH GROUP. **Webometric Analyst**. [2019]. Disponível em: <http://lexiurl.wlv.ac.uk/index.html>. Acesso em: 10 jun. 2020.

SUGIMOTO, C. R. *et al.* Scholarly use of social media and altmetrics: a review of the literature. **Journal of the Association for Information Science and Technology**, New York, v. 68, n. 9, p. 2037-2062, 2017.

THELWALL, M. **Introduction to Webometrics**: quantitative Web Research for the Social Sciences. San Rafael: Morgan & Claypool, 2009.

ZAHEDI, Z.; COSTAS, R. General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. **Plos One**, São Francisco, v. 13, n. 5, May 2018.

ZAHEDI, Z.; FENNER, M.; COSTAS, R. Consistency among altmetrics data provider/aggregators: What are the challenges? *In*: ALTMETRICS WORKSHOP, 9., 2015, Amsterdam. **Proceedings** [...]. Amsterdam: Amsterdam Science Park, 2015. Disponível em: <https://scholarlypublications.universiteitleiden.nl/handle/1887/48265>. Acesso em: 10 jun. 2021.

NOTAS

CONTRIBUIÇÃO DE AUTORIA

Concepção e elaboração do manuscrito: V. da R. Borba, S. E. Caregnato

Coleta de dados: V. da R. Borba

Análise de dados: V. da R. Borba

Discussão dos resultados: V. da R. Borba

Revisão e aprovação: S. E. Caregnato

CONJUNTO DE DADOS DE PESQUISA

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no próprio artigo.

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: 14-12-2020 – Aprovado em: 27-07-2021 – Publicado em: 08-10-2021.