

FONTES DE DADOS NA ALTMETRIA: um estudo exploratório

Vildeane da Rocha Borba

Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
vildeane.borba@gmail.com

Sônia Elisa Caregnato

Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
sonia.caregnato@ufrgs.br

1 INTRODUÇÃO

As ferramentas da internet e a evolução da web 2.0 têm modificado as formas de socialização, interação e principalmente a comunicação humana. Maresh-Fueher e Smith (2016) acreditam que o recurso da Internet que possui a maior contribuição para a mudança na comunicação são as mídias sociais, que permitem uma nova forma de comunicação em que os usuários, em vez de organizações ou guardiões de mídia, controlam a criação, o uso e a distribuição de informações.

A utilização das mídias sociais, em especial para difusão, discussão e compartilhamento de informações de cunho científico e acadêmico, pode ser comparado ao que Gladwell (2013, p. 15) chama de epidemia em uma mudança moderna, quando se percebe “[...] a possibilidade de contágio, o fato de que pequenas causas podem ter grandes efeitos e de que a mudança acontece não gradualmente, mas num momento decisivo”.

Esta “Epidemia” é percebida quando a curva de usuários ativos de mídias sociais cresce de forma ascendente no decorrer dos anos e a utilização de mídias sociais no compartilhamento, difusão e visibilidade de conteúdos científicos inaugurou uma nova área de estudos, chamada altmetria. A altmetria tem sido discutida como uma nova métrica advinda destas novas fontes emergentes e visa contribuir para a avaliação



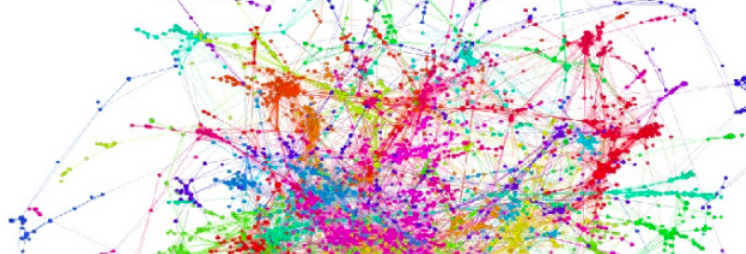
da ciência, juntamente com os indicadores tradicionais (KEMP, 2018). Neste sentido, este trabalho tem o objetivo de levantar as principais fontes de dados utilizadas nos estudos altmétricos, no sentido de entender suas características e possíveis contribuições que podem oferecer para a geração de indicadores de pesquisa.

2 ALTMETRIA

O desenvolvimento e utilização de mídias sociais como meio de expressão e comunicação tem se difundido no mundo. A utilização das mídias sociais na produção, discussão e difusão de conteúdos científicos afeta o sistema de comunicação da ciência vigente de diversas formas. Souza e Almeida (2013, p. 8) observam que uma dessas mudanças “[...] é a criação de novas métricas destinadas a acompanhar a disseminação de produtos de pesquisa na Web Social, as quais podem ser utilizadas como mecanismo de avaliação do impacto da ciência em complemento aos tradicionais estudos bibliométricos”.

Os estudos métricos tradicionais utilizam principalmente as bases de dados de citação como fonte de dados. A web social tem possibilitado novas formas de expressão e avaliação de conteúdos científicos com o intuito de complementar os indicadores tradicionais e para isso utiliza fontes alternativas de dados. Neste sentido, as fontes de dados na altmetria são aquelas encontradas em eventos e/ou atividades de usuários em plataformas, aplicativos e ferramentas *on-line* de mídias sociais com base na web social. Tais fontes são rastreadas pelos agregadores cuja função é monitorar as diversas fontes de dados advindas da web social, fornecendo as informações sobre os eventos e/ou atividades de menção ou compartilhamento de conteúdos científicos.

Erdt e colaboradores (2016) identificaram 40 diferentes fontes de dados que são monitoradas e coletadas pelos agregadores altmétricos, organizadas em oito categorias: A) Bibliotecas, Repositórios e Sistemas de informação digitais *on-line*, B) Blogging e Microblogging, C) Compartilhamento de vídeos, fotos e slides, D) Enciclopédias *on-line*, E) Marcadores sociais/Gerenciadores de Referências, F) Perguntas



e Respostas e Fóruns, G) Recomendações e Sistemas de Revisões e H) Redes Sociais. As publicações sobre tais fontes de dados, organizadas nessas oito categorias foram o objeto de estudo deste trabalho.

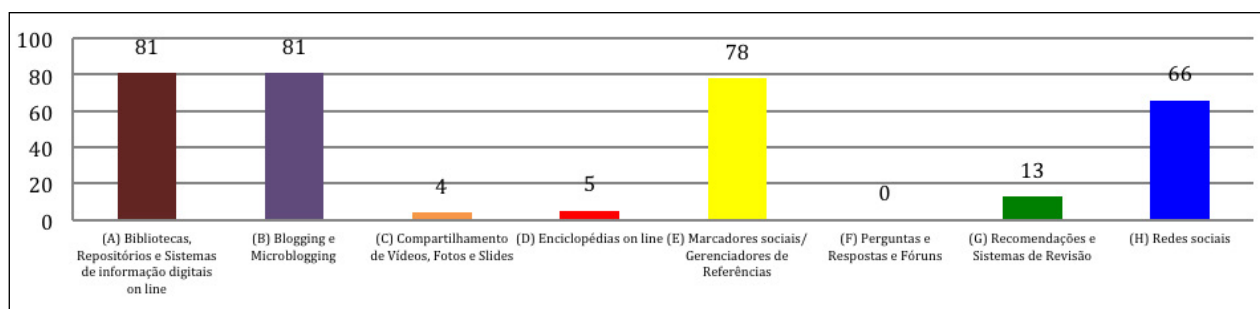
3 METODOLOGIA

Trata-se de pesquisa de caráter quali-quantitativo, do tipo descritiva e exploratória, que utiliza enquanto método a bibliometria e como técnica de coleta de dados a análise documental. A pesquisa foi realizada na base de dados *Scopus*, pois apresentou uma maior quantidade de registros sobre a temática, utilizando como estratégia de busca o termo “altmetrics”, o operador booleano AND, e o nome de cada fonte de dados, pesquisadas nos campos título do artigo, resumo e palavras-chave, realizada no dia 02 de janeiro 2018, abrangendo o período de 2012 a 2017. Fizeram parte do *corpus* de estudo 163 publicações, em que foram pesquisadas os estudos de 40 diferentes fontes de dados contempladas nas oito categorias apresentada por Erdt e colaboradores (2016).

4 RESULTADOS

As fontes de dados incluídas nas categorias A e B figuraram em 81 estudos cada. Foi percebida também um destaque para o grupamento E e H, com 78 e 66 estudos respectivamente, como se observa no Gráfico 1.

GRÁFICO 1 - QUANTITATIVO DOS ESTUDOS DE FONTES DE DADOS EM ALTMETRIA POR CATEGORIAS



Fonte: Dados de pesquisa, 2018.

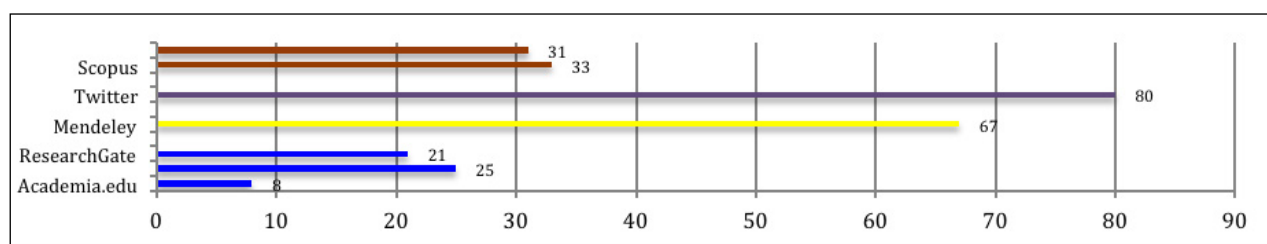
A categoria C teve um baixo número de estudos, que contemplaram as fontes de dados, *Dailymotion*, *Youtube*, *Vimeo*, *Pinterest*, *Flickr* e o *SlideShare*.



O grupo D, que inclui a *Wikipedia*, também apresentou um baixo número de publicação, enquanto que o F não contou com ocorrência. Fazem parte da categoria G, com 13 estudos, a *Amazon reviews*, *F1000*, *Goodreads*, *Publons*, *PubPeer* e *Reddit*.

O Gráfico 2 apresenta as fontes de dados das quatro categorias que apresentaram o maior número de publicações. No grupo A, sobressaem a *Web of Science* e *Scopus*. Tais bases de dados são utilizadas tradicionalmente nos trabalhos métricos, como fonte de coleta de informações bibliográficas e/ou para estudos de relação entre os dados altmétricos e contagem de citações.

GRÁFICO 2 - QUANTITATIVO DAS FONTES DE DADOS COM O MAIOR NÚMERO DE ESTUDOS



Fonte: Dados de pesquisa, 2018.

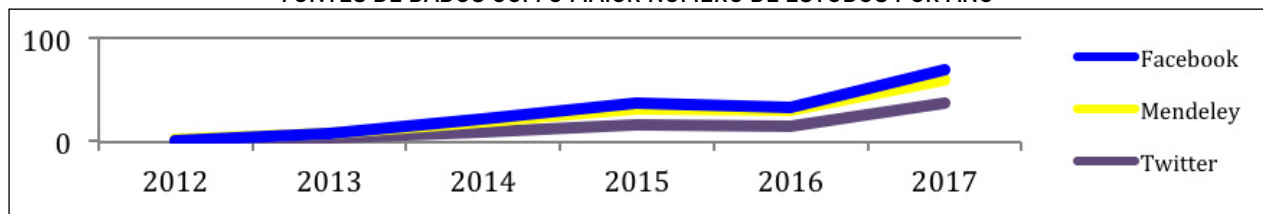
O Twitter aparece como a principal e maior fonte de dados para os estudos altmétricos, tanto para a categoria B, como para todas as categorias sociais pesquisadas, com 80 publicações. O *Mendeley* é a fonte de dados mais estudada para o grupo E, com 67 publicações, pois ele tem sido uma das fontes mais importantes de dados para a altmetria, refletindo o impacto das publicações em termos de público (BAR-ILAN, 2014; HAUNSCHILD; BORNMANN, 2016).

Na categoria H, o *Facebook*, apesar de ser a plataforma social com o maior número de usuários, mais de 1.871 milhões (KEMP, 2018), tem uma baixa ocorrência de estudos, com apenas 25 publicações. O *ResearchGate* e a *Academia.edu* aparecem com 21 e oito trabalhos, respectivamente. Interessante observar que elas são fontes de dados da web social desenvolvidas especificamente para fins acadêmicos, porém os dados demonstram uma fraca incidência, possivelmente pela dificuldade de coleta e análise de dados dessas fontes.



O Gráfico 3 apresenta a evolução do número de publicações por ano das três fontes de dados com o maior número de estudos. O *Twitter*, que é uma das plataformas sociais com o maior número de usuários, isto é, 317 milhões de acordo com Kemp (2018), apresenta estudos desde o ano de 2012. De 2012 a 2015 apresenta o número de publicações de forma ascendente, em 2016 tem uma pequena queda de estudos, tendo alcançado 37 publicações no ano de 2017.

GRÁFICO 3 - EVOLUÇÃO DO NÚMERO DE PUBLICAÇÕES DAS TRÊS FONTES DE DADOS COM O MAIOR NÚMERO DE ESTUDOS POR ANO



Fonte: Dados de pesquisa, 2018.

A curva para a fonte de dados *Mendeley* se apresentou de forma crescente, tendo seu primeiro estudo em 2013 e o maior número de publicações em 2017. O *Mendeley* foi a única fonte que apresentou sua curva crescente em todo o período estudado. Erdt e colaboradores (2016) afirmam que o *Mendeley* e *Twitter* tem sido as fontes de dados mais predominantes para estudos altmétricos, complementando que o *Twitter* tem recebido uma quantidade mais ou menos constante ao longo dos anos, mas parece que o interesse está se deslocando para o *Mendeley*.

O *Facebook* tem se mantido como uma fonte de dados de interesse para dados altmétricos, tendo visto estudos desde 2013 a 2015 de forma ascendente, em 2016 percebe-se uma queda no número de publicações, apresentando o maior número em 2017.

5 CONSIDERAÇÕES FINAIS

Das 40 fontes de dados analisadas, o *Twitter*, apesar de não ter sido desenvolvido para fins científicos e/ou acadêmicos, apresentou o maior número de estudos desde 2012 e é uma das plataformas sociais com o maior número de usuários, portanto merece atenção dos pesquisado-



res. Em contrapartida, o *Mendeley*, desenvolvido para fins acadêmicos, se apresenta como uma das fontes de dados mais promissoras para dados altmétricos, principalmente por sua abrangência, quantidade de usuários e difusão de compartilhamento de conteúdos científicos.

O *ResearchGate* e *Academia.edu* são fontes de dados desenvolvidas para a comunidade acadêmica, mas não apresentaram índices altos de estudos altmétricos. Acredita-se que tal fato deva-se a dificuldade de coleta e análise de informações das plataformas.

A web social tem demonstrado como um novo lócus para avaliação da ciência e um novo paradigma tem surgido para os estudos altmétricos. Diversos estudos ainda precisam ser desenvolvidos para entender e aprofundar dados sobre outras fontes e em especial aquelas desenvolvidos com o cunho acadêmico e científico.

REFERÊNCIAS

BAR-ILAN, J. Astrophysics publications on arXiv, Scopus and Mendeley: a case study. **Scientometrics**, v. 100, n. 1, p. 217-225, Jul. 2014.

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

GLADWELL, M. **O ponto da virada**. Rio de Janeiro: Sextante, 2013.

HAUNSCHILD, R.; BORNMANN, L. Normalization of Mendeley reader counts for impact assessment. **Journal of Informetrics**, v. 10, p. 62-73, 2016.

KEMP, S. **Digital in 2017 Global Overview**: a collection of internet, social media, and mobile data from around the world. London: we are social, 2018.

MARESH-FUEHRER, M. M.; SMITH, R. Social media mapping innovations for crisis prevention, response, and evaluation. **Computers in Human Behavior**, v. 54, p. 620-629, 2016.

SOUZA, I. V. P. de; ALMEIDA, C. H. M. de. Introdução à altmetria: métricas alternativas de comunicação científica. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 14., 2013, Florianópolis. **Anais...** Florianópolis: ANCIB, 2013.