

ALTMETRIA E ATENÇÃO *ON-LINE* DA PRODUÇÃO DE PESQUISADORES BRASILEIROS: *análise dos bolsistas de produtividade do CNPq*

Ronaldo Ferreria de Araújo
Universidade Federal de Alagoas, Maceió, AL, Brasil
ronaldfa@gmail.com

Marcelo Alves
Universidade Federal Fluminense, Niterói, RJ, Brasil
marceloalves.ufsj@hotmail.com

1 INTRODUÇÃO

A economia da atenção tem sido considerada uma abordagem adequada à gestão de informações que considera a atenção humana como uma mercadoria escassa, sendo necessária a aplicação da teoria econômica para resolver os problemas de gerenciamento informacional. Essa perspectiva coloca a atenção como recurso estratégico da ação humana, e frente ao crescimento abundante e imediato de conteúdos disponíveis, se torna um fator limitante no consumo de informações. Para Davenport e Beck (2001), a atenção é o engajamento mental focado em um determinado item de informação. Os itens entram em nossa consciência, dentre eles atendemos a um item específico e então decidimos agir. Assim, a atenção é usada para filtrar a informação mais importante a partir de um conjunto de informações que envolvem o ser humano, e que se proliferam cada vez mais nas redes e mídias sociais dos ambientes digitais.

É no contexto do âmbito da comunicação científica e do reconhecimento da necessidade de filtros que auxiliem pesquisadores na tomada de decisão de se dedicar a um item em detrimento de outro, uma vez que “ninguém pode ler tudo”, que Priem et al. (2010) ressaltam a necessidade de mais ferramentas e pesquisas baseadas em altmetria.

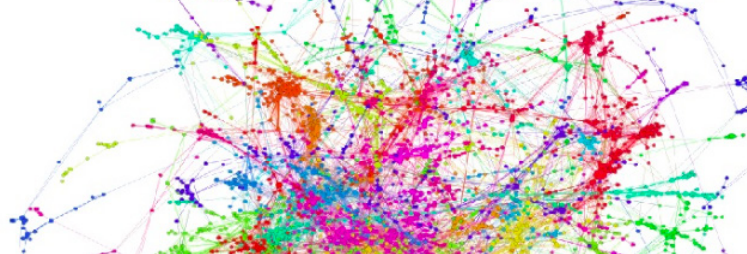


A altmetria tem sido considerada o campo de estudos que se dedica ao uso de dados cibernômicos para análises cientométricas (GOUVEIA, 2017) e nasce com a possibilidade de mensuração rápida de indicadores da atenção *on-line* que produtos da ciência alcançam, logo após sua publicação. Assim, o valor de seu indicador parece expressar a atenção *on-line* por meio da disseminação e outros tipos de interação que tais produtos geram na web, atingindo não apenas os interessados dentro da comunidade científica, mas também o grande público – cuja apropriação indica impacto social.

Embora as mídias sociais sejam anunciadas como grandes democratizadoras do conhecimento científico, o nível dessa democratização está muito relacionado a aspectos da economia geopolítica de cada dispositivo na comunicação científica, e tal como pode variar sua adesão em diferentes países e perfis de usuários, também pode apresentar indicadores de atenção variado quanto às áreas de conhecimento que cobrem e particularidades que apresentam.

O presente trabalho procura analisar o desempenho altmétrico de publicações brasileiras que tenham como autores pesquisadores bolsistas de produtividade em pesquisa (PQ) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). As bolsas PQs são destinadas a pesquisadores que se destaquem entre seus pares, valorizando sua produção científica segundo critérios normativos (CNPq, 2017), sendo atribuída a pesquisadores de todas as áreas, baseado não só na qualidade do projeto submetido, mas também na “qualidade” do pesquisador (WAINER; VIEIRA, 2013).

Existem pesquisas que se dedicaram à avaliação da produção científica resultante de pesquisas financiadas com bolsas de produtividade como a de Wainer e Vieira (2013), mas o foco geralmente é voltado para análise de métricas tradicionais e medidas bibliométricas. Sendo necessários estudos que atentem sobre a mesma produção, mas com um olhar não só do impacto acadêmico que recebem, a partir das citações, mas da atenção *on-line* que tais pesquisas despertam dentro e fora da comunidade científica, para o qual ainda não há trabalhos correlatos.



2 MATERIAL E MÉTODO

O banco de dados dessa pesquisa foi composto por uma amostragem de pesquisadores brasileiros, identificados por uma aplicação *Web scraping*, com extração de informações junto ao site do CNPq de todos os Bolsistas de Produtividades em vigência no Brasil em outubro de 2017, resultando em um total de $n = 14.609$. Em seguida foram tabuladas as suas publicações registradas no Currículo Lattes, a partir do pacote *GetLattesData* (PERLIN, 2017). Uma vez que publicações recentes costumam apresentar maior desempenho altmétrico, foram filtrados artigos de periódicos acadêmicos publicados entre 2016 e outubro de 2017 que possuem o *Digital Object Identifier* (DOI) informado no Lattes, $n = 99064$. Por fim, foram excluídos DOIs duplicados, referentes a artigos em coautoria, resultando em uma amostra final de $n = 69419$. Para artigos em coautoria com mais de um bolsista PQ só foi considerada a área de conhecimento do primeiro, levando em consideração a posição da ordem de autoria no trabalho.

A extração de dados da atenção *on-line* dessa produção foi realizada pelo pacote *rAltmetric* (rOpenSci, 2017) que consulta a *Application Programming Interface* (API) pública do *Altmetric* a partir do DOI das publicações e retorna uma planilha com os dados produzidos pela plataforma. Em seguida, os metadados fornecidos pelo CNPq (filiação, nível da bolsa PQ, Grande Área) e pelo Lattes (Idioma) foram cruzados, para produzir uma tabela única com as informações de interesse do trabalho.

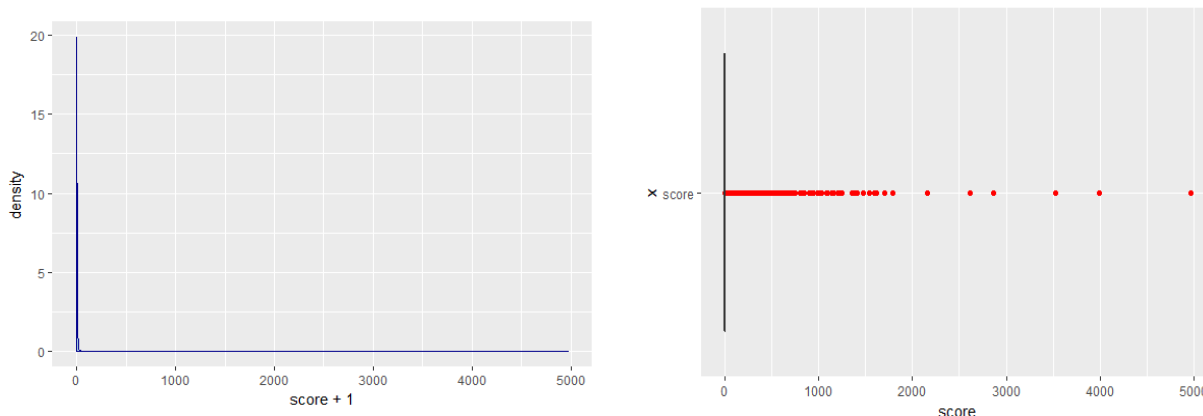
Vale lembrar que o *Attention Score Altmetric* é uma medida geral da atenção que um artigo ou conjunto de dados tem recebido *on-line*, e deve refletir a visibilidade (alcance) e a influência e engajamento (qualidade da atenção, local postado e reputação de quem posta) que pode variar de acordo com a fonte ou tipo de mídia considerada (ALTMETRIC, 2017). Os dados de atenção *on-line* são analisados segundo sua distribuição por densidade e variação; idioma da publicação e área de conhecimento dos pesquisadores; e por desempenho da média por tipo de fonte que forneceu seus valores altmétricos.



3 ANÁLISE E DISCUSSÃO DOS RESULTADOS

As soluções disponíveis para monitoramento e análise da atenção *on-line* e dos dados altmétricos da produção científica exige dessa produção padrões de identificação, como o DOI, que garantam entre outras vantagens sua distinção. Foi verificado que 73, 41% das publicações analisadas possui o identificador, o que parece revelar que os bolsistas PQ estão atentos a publicar seus resultados de pesquisa em periódicos com este padrão. A distribuição da variável *score* (atenção *on-line*) segundo sua densidade e variação pode ser vista na Figura 1. No primeiro caso, percebe-se que a densidade evidencia o comportamento de cauda longa da variável, isto é, a maior parte dos artigos não possui atenção *on-line*, ou seja, *score altmetrics* = 0, enquanto que bem poucos artigos possuem um índice elevado. Uma escala logarítmica em *x* foi considerada, mas não mudou muito a distribuição por conta do número muito elevado de zeros.

FIGURA 1 - DISTRIBUIÇÃO DA ATENÇÃO ON-LINE POR GRAUS DE DENSIDADE EVARIAÇÃO



Fonte: Dados da pesquisa (2017).

O segundo caso, reforça o comportamento do primeiro, os dados mostram que todas as medidas de posição estão próximas de 0 e a variável *score* possui um número muito grande de *outliers*, variando de 0 a 4966.648. A distribuição da média da atenção *on-line* por idioma (Gráfico 1) e por área de conhecimento (Gráfico 2) também foram observadas, tendo em vista que caracterizam o desempenho da produção analisada.

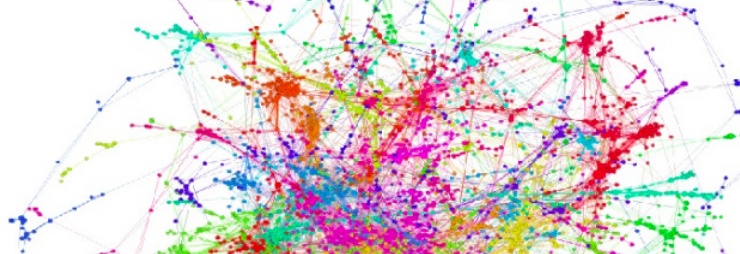
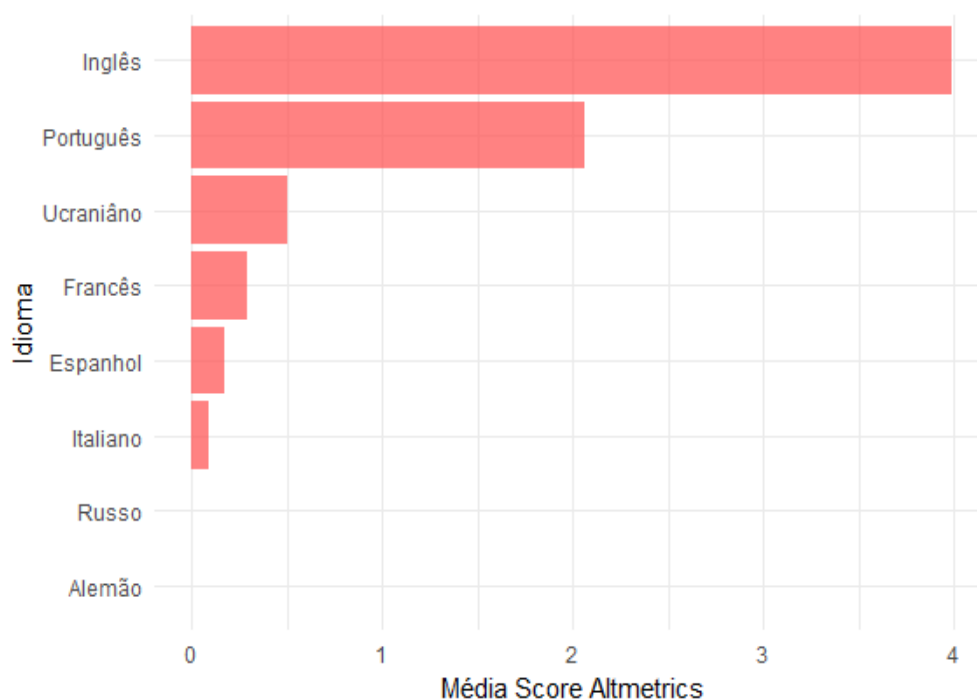
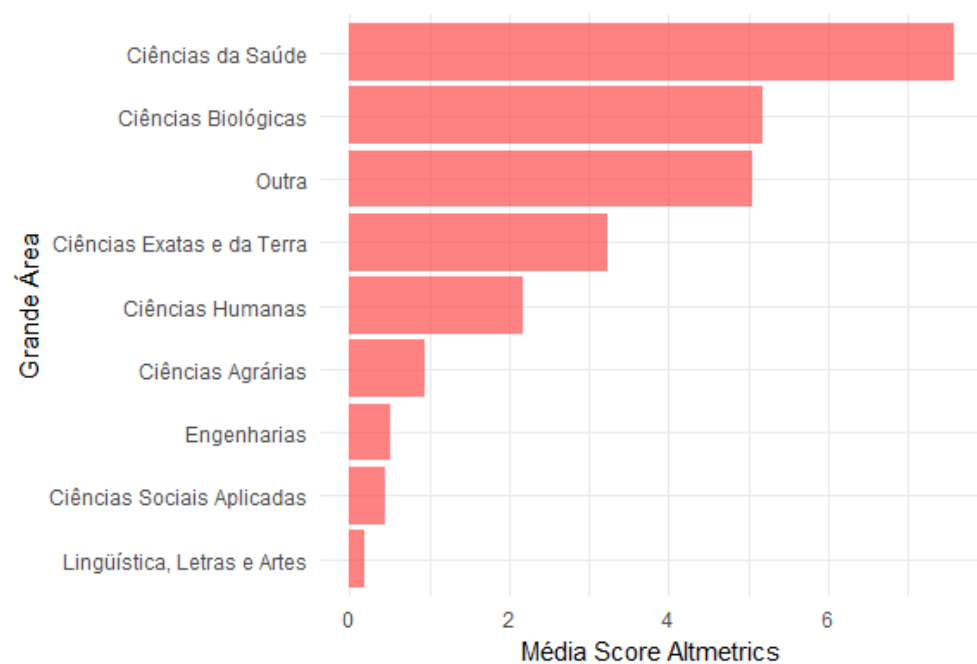


GRÁFICO 1 - MÉDIA DA ATENÇÃO *ON-LINE* POR IDIOMA



Fonte: Dados da pesquisa (2017).

GRÁFICO 2 - MÉDIA DA ATENÇÃO *ON-LINE* POR ÁREA DE CONHECIMENTO



Fonte: Dados da pesquisa (2017).



Na contramão da noção de democratização do conhecimento científico atribuído às mídias sociais, algumas pesquisas mostram que muitas vezes a análise de seu uso reproduz as mesmas disparidades de visibilidade e das métricas tradicionais, como o fato das publicações de países emergentes tenderem a ter menor visibilidade nas plataformas de redes sociais (ALPERIN, 2015) e do inglês ser a língua dominante (FAUSTO et al., 2012). Neste último caso, o resultado obtido com sua representação da média da atenção *on-line* por idioma, o Gráfico 1 apenas confirma tal comportamento. Quanto ao Gráfico 2, a média da atenção *on-line* indica melhor desempenho das Ciências da Saúde, com 7.56 pontos; e das Ciências Biológicas, com 5.17. O resultado é similar ao do estudo de Costas, Zahedi e Wouters (2014) no qual as Ciências da Saúde e Biológicas também se apresentam como áreas de conhecimento que obtiveram maior atenção *on-line*, com valores altmétricos superiores às demais. Mas no geral, os valores obtidos pela produção dos pesquisadores brasileiros são muito baixos.

O *Attention Score Altmetric* é composto por mais de 10 fontes distintas como portais de notícias (imprensa), blogs, mídias sociais, gerenciadores de referência, dentre outros. A Tabela 1 apresenta a distribuição da média da atenção *on-line* da produção analisada pelas fontes mais expressivas (*Mendeley*, *Twitter*, *Facebook*, *Imprensa* e *Blogs*). Fontes como *Wikipedia*, *YouTube*, *Reddit* tiveram incidência com valores muito baixos e foram desconsideradas.

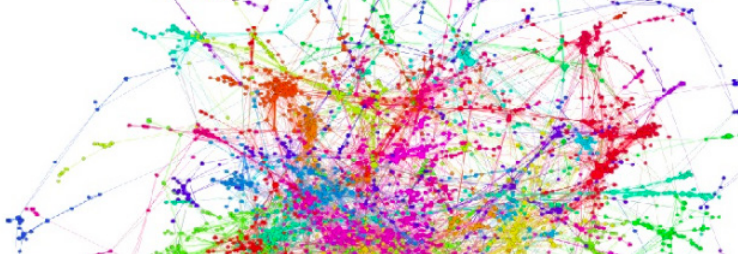
**TABELA 1 - MÉDIAS DE MENÇÕES DAS ÁREAS DE CONHECIMENTO POR TIPO DE FONTE**

Grande Área	Mendeley	Twitter	Facebook	Imprensa	Blogs
Ciências Agrárias	10.45	3.85	1.55	3.50	1.37
Ciências Biológicas	15.43	7.09	2.25	6.51	1.82
Ciências da Saúde	14.38	12.61	2.80	11.37	2.02
Ciências Exatas e da Terra	11.33	6.56	1.57	7.53	2.27
Ciências Humanas	8.10	5.97	1.85	9.40	1.92
Ciências Sociais Aplicadas	10.14	3.79	1.26	1.33	1.05
Engenharias	11.80	3.51	1.22	3.57	2.05
Linguística, Letras e Artes	3.70	3.29	1.29	0.00	1.00
Outra	14.53	10.19	2.05	8.98	1.62

Fonte: Dados da pesquisa (2017).

O primeiro ponto a se destacar é o bom desempenho do *Mendeley*, seguido do *Twitter* e a baixa cobertura do *Facebook*, que muito embora seja a mídia social mais utilizada no Brasil, alcançou média de menções bem inferior às demais não perdendo apenas para a cobertura em blogs. O resultado é muito similar ao de estudos que analisaram a diferença de atenção *on-line* entre as fontes dos dados. Tanto na pesquisa de Priem, Piwowar e Hemminger (2012) que analisou dados altmétricos da coleção *Public Library of Science* (PLOS), quanto na de Alperin (2015) que verificou a atenção *on-line* da coleção *Scientific Electronic Library Online* (SciELO) de países da América Latina, foi identificada maior cobertura de dados vindos do *Mendeley* com média de leituras bem superior às menções no *Twitter*, por sua vez, com médias superiores às postagens no *Facebook*.

É interessante notar as diferenças de desempenho das áreas de conhecimento por tipo de fonte. O *Mendeley*, por exemplo, registra valores maiores nas Ciências Biológicas (15.43) – resultado similar ao estudo de Alperin (2015) – enquanto as Ciências da Saúde possuem maior cobertura do *Twitter* (12.61) e do *Facebook* (2.80). Outro destaque é o desempenho de mídias da imprensa com médias que as conferem uma cobertura maior para as Ciências da Saúde (11.37), seguido das Ciências Humanas (9.40) e nenhuma cobertura para as áreas de Linguística, Letras e Artes. As áreas com melhor desempenho em postagens em blogs são as de Ciências Exatas e da Terra (2.27) e Engenharias (2.05).



Para Alperin (2015), essas flutuações entre as áreas quanto aos níveis de cobertura são indicativas de diferentes usos de cada tipo de fonte, potencialmente indicando que públicos variados estão sendo alcançados por cada disciplina ou práticas diversas em relação às mídias sociais e aos gestores de referências.

4 CONSIDERAÇÕES FINAIS

A produção brasileira de determinada área do conhecimento de autoria de pesquisadores bolsistas de produtividade do CNPq reflete a agenda de pesquisa e constitui corpo de conhecimento privilegiado sobre o pensamento dessa área no país, expressando compromissos com correntes teóricas, metodológicas e de pensamentos acerca de sua historicidade, atuação e tendências.

O estudo procurou avaliar essa produção sob a perspectiva de seu impacto social medido, não pelas citações, mas pela atenção *on-line* que tais pesquisas despertam dentro e fora da comunidade científica ao serem mencionadas em mídias sociais. Estudos qualitativos são necessários para averiguar conteúdos e contextos dessas menções e aprimorar o quadro altmétrico da produção analisada.

O baixo resultado encontrado nos índices e nas médias de menções, seja de maneira geral, por área de conhecimento ou tipo de fonte, nos fazem questionar se isso ocorre devido a uma não cultura do compartilhamento de informações científicas nas mídias sociais ou se indica, como apontado por Alperin (2015), uma questão de limitação e cobertura da ferramenta *Altmetric*, que precisa sempre ser considerada para análises do impacto altmétrico da produção científica nacional.

REFERÊNCIAS

ALPERIN, J. P. Geographic variation in social media metrics: an analysis of Latin American journal articles. **Aslib Journal of Information Management**, v. 67, n. 3, p. 289-304, 2015.



ALTMETRIC. **How is the Altmetric score calculated?** 2017. Disponível em: <<https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-attention-score-calculated->>. Acesso em: 11 dez. 2017

CNPq. Conselho Nacional de Desenvolvimento Científico e Tecnológico. **Bolsas no país. Produtividade em Pesquisa- PQ.** 2017. Disponível em: <<http://cnpq.br/apresentacao13/>>. Acesso em 11 dez. 2017.

COSTAS, R.; ZAHEDI, Z.; WOUTERS, P. Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. **JASIST**, v. 66, n.10, jul. 2014. <http://dx.doi.org/10.1002/asi.23309>

DAVENPORT, H., BECK, J. **The Attention Economy:** Understanding the New Currency of Business. Boston, Massachusetts: Harvard Business School Press, 2001.

GOUVEIA, F. C. A altmetria e a interface entre a ciência e a sociedade. **Trab. educ. saúde**, Rio de Janeiro, v. 14, n. 3, p. 643-645, dez. 2016 . <http://dx.doi.org/10.1590/1981-7746-sip00126>

PERLIN, M. S. **Package GetLattesData:** downloading and reading bibliometric data from Lattes. 2017. Disponível em: <<https://msperlin.github.io/2017-09-04-Package-GetLattesData/>>. Acesso em: 03 nov. 2017.

PRIEM, J.; PIWOWAR, H. A.; HEMMINGER, B. M. Altmetrics in the wild: Using social media to explore scholarly impact. **ArXiv.org**, 2012. Disponível em: <<http://arxiv.org/abs/1203.4745>>. Acesso em: 03 dez. 2017.

PRIEM, J.; TARABORELLI, D.; GROTH, C.; NEYLON, C. **Altmetrics:** a manifesto. 2010. Disponível em: <<http://altmetrics.org/manifesto>>. Acesso em: 16 ago. 2016.

WAINER, J.; VIEIRA, P. Avaliação de bolsas de produtividade em pesquisa do CNPq e medidas bibliométricas: correlações para todas as grandes áreas. **Perspect. ciênc. inf.**, Belo Horizonte, v. 18, n. 2, p. 60-78, jun. 2013. <http://dx.doi.org/10.1590/S1413-99362013000200005>