

A PLATAFORMA DE BLOGS CIENTÍFICOS RESEARCHBLOGGING.ORG PARA AVALIAÇÃO ALTERNATIVA DO INTERESSE, USO E ALCANCE DE PESQUISAS

Eixo temático: Informetria e Webometria

Modalidade: Apresentação oral

1 INTRODUÇÃO

O surgimento e a expansão de novas tecnologias e ferramentas baseadas na Internet estão proporcionando meios para incrementar a comunicação científica, melhorando o acesso aos resultados de pesquisa e seu compartilhamento para um público mais amplo, potencializando a avaliação de outras dimensões do impacto dos resultados das pesquisas. Este contexto requer novos sistemas de gestão para tornar o conhecimento científico mais acessível e utilizável, incluindo metrias alternativas para avaliar o alcance e qualidade da informação científica publicada. Entre as novidades tecnológicas, as mídias sociais surgidas com o advento da chamada web 2.0 oferecem um forte potencial para avaliações complementares dos resultados da investigação científica. Ferramentas e serviços online como blogs, Twitter, Facebook e gerenciadores de bibliografia a exemplo do CiteULike, do Connotea, do Mendeley e do Zotero apontam para alternativas de análise sobre o real interesse, uso e alcance das publicações científicas.

Entre essas novas tecnologias destacam-se os blogs de ciência como importantes filtros para a recuperação de informação científica publicada. A blogosfera científica cresceu significativamente nos últimos anos, devido à importância atribuída à divulgação científica, engajando cientistas na escrita sobre seus temas de pesquisa e interesse, para seus pares e para o público em geral, numa linguagem acessível através de postagens em blogs especializados (BONETTA, 2007), que ampliam o alcance e funcionam como filtros para a literatura publicada. Verifica-se uma tendência positiva para a agregação entre os blogs de ciência, principalmente através de condomínios e plataformas online. Este estudo descreve a plataforma Research Blogging (<http://www.researchblogging.org>), um agregador de blogs

científicos que postam sobre literatura publicada revisada por pares, verificando seu potencial como alternativa para a avaliação no atual contexto de mudanças da comunicação científica.

2 RESEARCH BLOGGING: BREVE HISTÓRICO E CARACTERIZAÇÃO

A plataforma Research Blogging (RB) foi criada em 2007 por um blogueiro de ciências e desenvolvida em colaboração com o Seed Media Group, através da criação de um ícone especial para distinguir postagens sobre pesquisa publicada revisada por pares, indexando-as e reunindo-as numa coleção online através de um agregador de blogs, sendo uma fonte útil para leitores interessados em pesquisa de ponta comentada por cientistas e especialistas em suas respectivas áreas e funcionando com um filtro eficaz para recuperar e identificar pesquisas acadêmicas sérias e evitar a propagação de conteúdos duvidosos. De fato, os blogs cadastrados na plataforma RB seguem diretrizes elaboradas pela própria comunidade de blogueiros de ciência para garantir a qualidade dos conteúdos disponibilizados. Atualmente a plataforma RB dispõe de mais de 1.230 blogs ativos, com mais de 26.960 postagens sobre a pesquisa publicada revisada por pares em assuntos que vão da antropologia à zoologia, organizadas em categorias e em sete diferentes idiomas: inglês, alemão, chinês, espanhol, italiano, polonês e português.

Este trabalho investigou quantitativamente o número de blogs e postagens da plataforma RB, sua distribuição através dos idiomas e temáticas adotados pelo agregador de blogs, as revistas científicas de Acesso Aberto (AA) e restritas e as visualizações das postagens, estabelecendo correlações entre metrias tradicionais.

3 METODOLOGIA

3.1 COLETA E TRATAMENTO DOS DADOS

Foi realizado um estudo exploratório e de abordagem quantitativa para a coleta de dados nas postagens do site Blogging Research. A pesquisa foi realizada em janeiro de 2012 e incluiu todo o período disponível desde a criação da plataforma, considerando postagens publicadas entre 01 de novembro de 2007 e 31 de dezembro de 2011 e apenas aquelas discutindo artigos publicados em revistas científicas revisadas por pares. Excluíram-se

postagens com referências a livros, anais de eventos, manuais e outras fontes online ou offline. Também foram excluídos blogs sem um endereço online ativo. Os dados obtidos foram tratados, listados em planilhas e categorizados segundo o número total de blogs, de postagens, de tópicos e idiomas adotados pela plataforma, o número de artigos citados e de títulos de revistas científicas, restritas e de acesso aberto, e o número total de visualizações das postagens.

3.2 ANÁLISE ESTATÍSTICA E COMPARATIVA ENTRE MÉTRICAS

As citações de artigos científicos e as visualizações das postagens foram contadas para cada revista científica citada nos blogs da plataforma RB. Verificou-se se ambas variáveis poderiam ser descritas por uma lei de potência da forma $\Pr(x) \sim (x/x_{\min})^{-\alpha}$, onde x é o valor da variável de soma, $x_{\min}$ é o valor mínimo de x que segue um comportamento da lei de potência e α é o parâmetro de escalonamento. $x_{\min}$ e α foram estimados utilizando uma otimização por máxima-verossimilhança e a adequação da distribuição à lei de potência foi avaliada através do teste de Kolmogorov-Smirnoff (CLAUSET; SHALIZI; NEWMAN, 2009; MILLER, 1956). Em seguida, usou-se o $x_{\min}$ para definir o limite inferior da amostra para a análise estatística, incluindo apenas as revistas com citações nos blogs e visualizações superiores ao $x_{\min}$ estimado para ambas as variáveis. Comparou-se os resultados dessas revistas citadas com 7 diferentes métricas da base *Journal Citation Reports* (JCR), da Thomson Reuters, a saber: 1. Fator de Impacto (FI, *Impact Factor*); 2. Total de artigos; 3. Total de citações; 4. Meia-vida de citações (*Cited Half-Life*); 5. Índice de imediatez (*Immediacy index*); 6. *Eigenfactor*; e 7. Influência do Artigo (*Article Influence Score*). As correlações entre as variáveis de contagem do RB e as variáveis do JCR foram avaliadas através do teste estatístico ρ de Spearman. Por fim, para verificar se os artigos em acesso aberto foram mais citados do que o esperado, comparou-se a proporção de citações em AA nos blogs com a proporção de artigos AA na amostra, através de um teste binomial. Todas as análises foram realizadas no programa estatístico R, v 2.14.2 (R DEVELOPMENT CORE TEAM, 2012).

4 RESULTADOS E DISCUSSÃO

Durante o período analisado (2007-2011) a plataforma Research Blogging coletou, registrou, indexou e compartilhou 26.969 postagens de 1.236 blogs (Fig. 1), distribuídas nas

seguintes categorias: Biologia, 9.787 postagens (36% do total); Ciências da Saúde (compreendendo Saúde e Medicina), 4.177 postagens (15% do total); Psicologia, 3.401 postagens (13%); Neurociência, 2.495 (9%); Ciências Sociais Aplicadas, 1.108 (4%); Antropologia 1.058 (4%), Química 879 (3%); Física, 835 (3%), Geociências, 518 (2%), Educação e Pesquisa, 438 (2%); Astronomia, 407 (2%), Ciências da Computação/ Engenharias, 239 (1%), Ecologia, 221 (1%); Filosofia, 152 (1%); e Matemática com 77 postagens. A categoria Outros obteve 1.177 postagens (4%) (Fig. 2). O idioma dominante foi o inglês, com 1.008 blogs e 22.660 postagens publicadas, seguido pelo português, com 65 blogs e 1.013 postagens. O espanhol apresentou 52 blogs, com 1.456 postagens; o alemão 36 blogs e 742 postagens; o italiano 32 blogs com 449 postagens; o polonês 24 blogs e 512 postagens; e o chinês 19 blogs com 137 postagens publicadas (Fig. 3). No período analisado 19.000 postagens citaram e linkaram 26.171 artigos científicos publicados em 3.350 diferentes revistas (Fig. 4). A grande área mais coberta por títulos de periódicos foi a de Ciências da Saúde, com 1.071 títulos, seguida de Ciências Sociais Aplicadas, com 796 títulos; Ciências Biológicas com 599 títulos; Exatas e Ciências da Terra, 530 títulos; enquanto que a área Multidisciplinar apresentou 308 títulos de periódicos citados e as Humanidades, 46 títulos. Em relação às citações, a tendência encontrada não segue uma estreita relação com o Fator de Impacto (FI) das revistas, e verificaram-se grupos de periódicos mais citados com revistas de alto FI e outros com FI de menos de 10. Os testes de Kolmogorov-Smirnoff indicaram que as citações e as visualizações seguem a lei de potência de distribuições ($p < 0,001$). O $x_{\min}$ de 3 valores de citações indicou a inclusão de 1.334 revistas a partir das quais 820 tinham informações cientométricas disponível no JCR. As correlações entre as métricas JCR e as contagens do RB foram geralmente baixas (Tabela 1). Todas as correlações foram inferiores a 0,40, exceto entre o Total de Citações e o Eigenfactor, e ambas as contagens RB (citações e visualizações) apresentaram uma correlação particularmente baixa com a métrica Vida Média de Citações (*Cited Half-Life*). As correlações médias entre as citações e as visualizações (RB) com as métricas do JCR foram de 0,33 e 0,29, respectivamente. Sobre as revistas de acesso aberto, os resultados mostraram que 11,7% das citações (3.055 de 26.171) vieram de revistas AA, um valor quatro vezes maior do que a observada em citações verificadas na Wikipedia (2,8%), conforme achados em pesquisa de Evans e Krauthammer (2011). Estas citações vieram de 7,2% das revistas em AA presentes na amostra (242 de 3.350). As diferenças entre a proporção de citações a AA e artigos em AA disponíveis foram

significativas sob um teste binomial ($p = 1.822e-48$) (Figura 5). Ao contrário de citações da Wikipédia (EVANS; KRAUTHAMMER, 2011), seis das revistas mais citadas foram AA: PLoS ONE (com mais de 1.000 citações), Psychological Science e PLoS Biology (entre 201 e 350 citações); e PLoS Medicine, Pediatrics e PLoS Pathogens (101 a 200 citações).

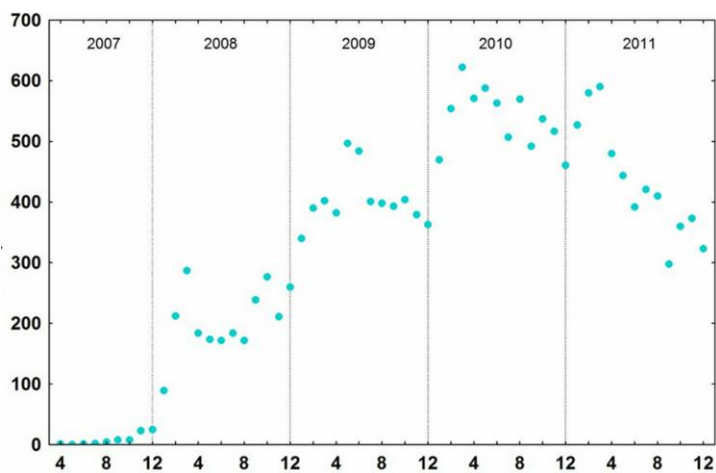


Figura 1: Total de postagens na Plataforma Research Blogging

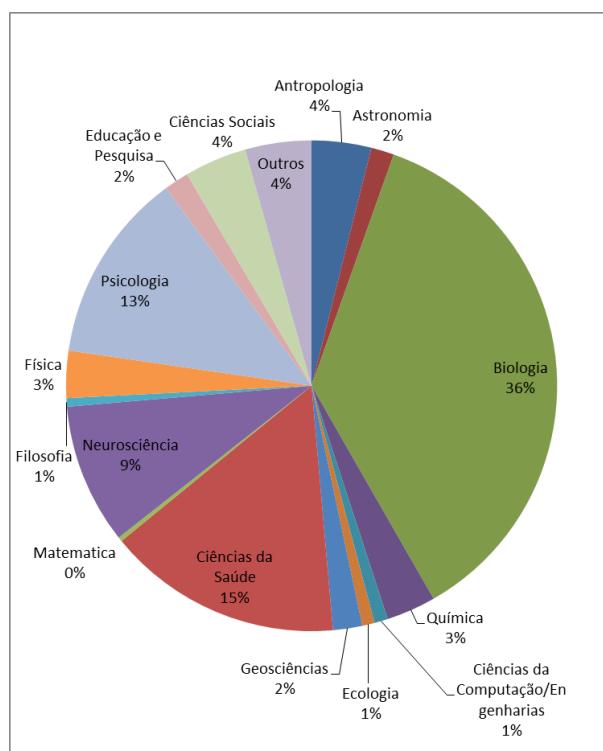


Figura 2: Distribuição de postagens por categorias da plataforma Research Blogging

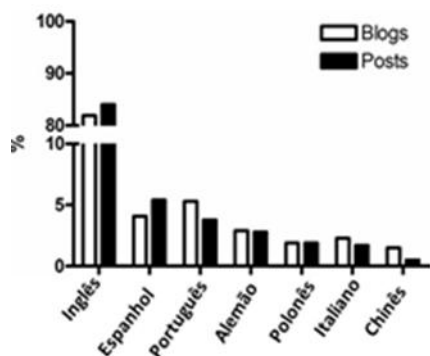


Figura 3: Distribuição das postagens da Plataforma Research Blogging por idioma

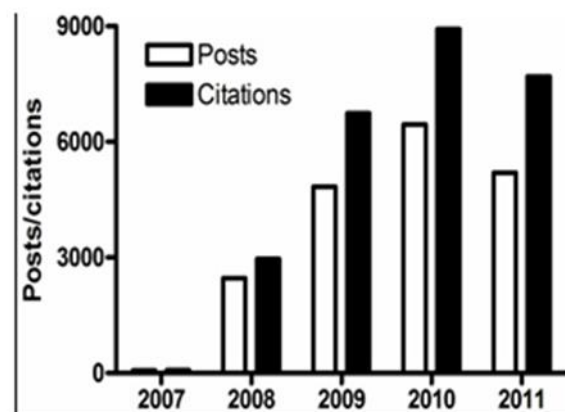


Figura 4: Número Total de Postagens e Citações

	Citações (blog)	Visualizações
IF	0.3757	0.3188
Total de Artigos	0.2937	0.2590
Total de Citações	0.4087	0.3649
Índice de Imediatez	0.3469	0.3003
Meia-Vida	0.0491	0.0617
Eigenfactor	0.4339	0.3806
Influência do Artigo	0.3841	0.3215

Tabela 1: Correlação estatística entre as contagens RB e as metrias JCR

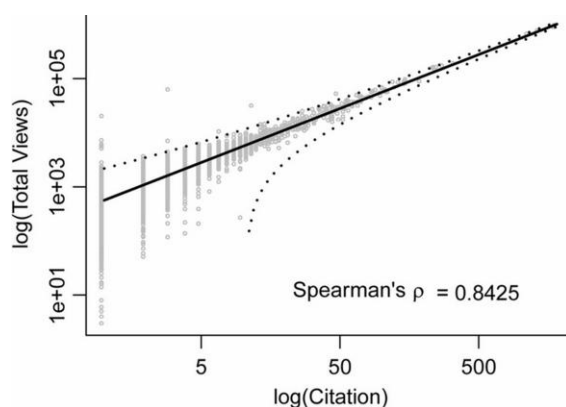


Figura 5: Correlação entre o número de visualizações e as citações

5. CONSIDERAÇÕES FINAIS

As análises sugerem que as citações e as visualizações nas postagens da plataforma Research Blogging avaliam aspectos da produção científica diversos das análises cientométricas tradicionais. De fato, as correlações encontradas aqui são menores do que as correlações médias entre todas as métricas de citação indicadas por Bollen et al (2009), como FI, Índice de Imediatez, Citações por documento, Meia-Vida de Citações, etc, indicando que métricas baseadas em blogs científicos como os do RB avaliam uma dimensionalidade diferente da qualidade e do alcance das revistas científicas, sugerindo a possibilidade de avaliações de aspectos complementares da qualidade da pesquisa. Neste sentido, as correlações das métricas encontradas no RB são consistentes com a chamada *Altmetrics* (*Altmetrics*) (PRIEM; PIWOWAR; HEMMINGER, 2012), indicando que os blogs de ciência organizados e sistematizados como o RB podem ser considerados para o estudo dessas metrias alternativas. Tais metrias alternativas mostram-se promissoras para complementar a avaliação de revistas científicas, dizendo algo sobre seu impacto e alcance social através de citações e visualizações em mídias sociais como os blogs de ciência.

REFERÊNCIAS

- BOLLEN, J.; VAN DE SOMPEL, H.; HAGBERG, A.; CHUTE, R. A principal component analysis of 39 scientific impact measures. *PLoS ONE*, v. 4, n. 6, e6022, 2009. doi:10.1371/journal.pone.0006022.
- BONETTA, L. Scientists Enter the Blogosphere. *Cell*, v. 129, n. 3, p. 443-445, 2007.
- CLAUSET, A.; SHALIZI, C.R.; NEWMAN, M.E.J. Power-Law Distributions in Empirical Data. *Siam. Review*, v. 51, p. 661-703, 2009.
- EVANS, P.; KRAUTHAMMER, M. Exploring the use of social media to measure journal article impact. In: *AMIA Annual Symposium Proceedings*, p. 374-378, 2011.
- MILLER, L.H. Table of Percentage Points of Kolmogorov Statistics. *J Am Stat Assoc.*, v. 51, p. 111-121, 1956.
- PRIEM, J.; PIWOWAR, H.A.; HEMMINGER, B.M. Altmetrics in the wild: Using social media to explore scholarly impact. 2012 [preprint]. Disponível em: <http://arxiv.org/html/1203.4745v1>. Acesso em: 22 Abril 2012.
- R DEVELOPMENT CORE TEAM. *R: A Language and Environment for Statistical Computing*. Viena: R Foundation for Statistical Computing, 2012.