

Artigos de periódicos em acesso aberto: citações distribuídas em repositórios institucionais

Caterina Groposo Pavão

Universidade Federal do Rio Grande do Sul

caterina@cpd.ufrgs.br

Janise Silva Borges da Costa

Universidade Federal do Rio Grande do Sul

janise@cpd.ufrgs.br

Manuela Klanovicz Ferreira

Universidade Federal do Rio Grande do Sul

manuelakf@cpd.ufrgs.br

André Rolim Behr

Universidade Federal do Rio Grande do Sul

andre.behr@cpd.ufrgs.br

Resumo

Este trabalho propõe uma metodologia que possibilite coletar dados distribuídos, que comprovem o aumento das citações de artigos publicados em periódicos científicos e disponíveis em repositórios institucionais de acesso aberto, por meio de uma curva de citação que mostre, ao longo do tempo, o número de novas citações distribuídas em repositórios institucionais (RI), para os artigos disponibilizados em uma determinada janela temporal. Hoje, é inevitável e iminente demonstrar concretamente, aos pesquisadores, por meio da quantificação do aumento do impacto das publicações, as vantagens do depósito em repositórios institucionais de acesso aberto. Agregar serviços de valor aos repositórios requer estudos especializados, que vão além da contagem de acessos e *downloads*. Indicadores bibliométricos podem ser utilizados para provocar novos olhares em pesquisas futuras, que forneçam dados do uso real da produção científica disponível em acesso aberto e fornecer informações relevantes sobre um determinado tipo de documento ou um autor.

Palavras-chave: Repositórios institucionais, Acesso aberto, Artigos de periódicos, Análise de citações

Journal articles in open access: citations distributed in institutional repositories

Abstract

This work proposes a methodology that allows the collection of distributed data, which proves the citation increasing of published articles in scientific journals and available in open access institutional repositories, through a citation curve that shows, over time, the number of new citations distributed in institutional repositories (IR), for articles made available in a given time window. Nowadays, it is inevitable and imminent to concretely demonstrate to researchers the advantages of filing in open access institutional repositories by quantifying the increased impact of publications. Adding value services to repositories requires specialized studies, going beyond counting accesses and downloads. Bibliometric indicators can be used to elicit new insights into future research that provide data on the actual use of scientific production available in open access and relevant information about a particular type of document or author.

Key-words: Institutional repositories, Open access, Journal articles, Citations analysis

Introdução

As instituições de ensino superior são responsáveis por grande parte da produção científica do país. Os documentos gerados como produto de pesquisa, em sua maioria, encontram-se dispersos dentro das instituições ou armazenados em bibliotecas, o que pode dificultar o acesso devido às barreiras de tempo e espaço ou, ainda, acessíveis por meio de publicações ou bases de dados cujas assinaturas têm custos bastante elevados.

Os repositórios institucionais (RI) pretendem contribuir para a melhoria do sistema de comunicação científica e influenciar positivamente no avanço científico daquilo que se traduz em benefícios para a ciência e para a sociedade. Apesar de terem surgido no Brasil há mais de dez anos, não há registro de estudos específicos sobre as suas potencialidades para a gestão e controle da produção científica de uma instituição de ensino superior ou de pesquisa.

Estudos evidenciam a premência de agregar, aos RI, serviços que demonstrem não apenas quanto, quando e como os documentos são acessados e baixados, mas também, quanto, quando e como os documentos são citados. Para tanto, é necessário realizar medições sistemáticas do percurso dos documentos nos repositórios.

A medição do impacto dos documentos pressupõe a contagem de citações, mas quando essa contagem deve ser feita nos documentos de um repositório são encontradas dificuldades relativas aos mecanismos a serem utilizados. Os dados de uso de um RI são

gerados automaticamente no servidor, porém, os dados de citações exigem a atuação dos gestores do repositório. Segundo Barrueco (2008, tradução nossa), esta é uma operação dispendiosa, pois obter níveis aceitáveis de precisão exigirá a dedicação de consideráveis recursos humanos e técnicos. Devido, precisamente, aos custos, o autor argumenta que seria aconselhável fazer a extração de citações em nível de provedor de serviços, sendo que este, em âmbito nacional, seria o responsável pela colheita de documentos e seu processamento para suas próprias análises e por divulgá-las para domínio público. Portanto, teria um serviço que, em termos de protocolo OAI-PMH, atuaria como coletor e como provedor de dados, por sua vez, os dados poderiam subsidiar a tomada de decisão dos gestores de instituições de ensino e pesquisa e agências de fomento.

Fundamentação teórica

A visibilidade vem constituindo-se uma questão crucial, não só para os pesquisadores, mas também para suas instituições. Em consequência das vantagens do acesso aberto, dentre as quais destaca-se o aumento significativo da visibilidade das pesquisas, os pesquisadores vêm exercendo pressão sobre as editoras para que removam as barreiras de acesso a suas publicações. Isso pelo entendimento de que os recursos despendidos com pesquisa só se justificam se seus resultados estiverem amplamente acessíveis. Esta nova visão sobre a publicação científica desafia modelos já estabelecidos. Costa (2006), em seu artigo, chamava a atenção para a necessidade de encarar o desafio e mudar o modelo do processo de comunicação científica, tanto no que concerne aos papéis tradicionalmente desempenhados por seus atores, quanto no que diz respeito aos formatos da publicação científica.

Sob essa ótica, os RI podem constituir-se em instrumentos de reordenação quanto aos critérios de visibilidade e impacto da literatura científica, do pesquisador e das instituições, visto que agora ele passa a produzir para um universo que extrapola as dimensões dos mecanismos de disseminação até agora utilizados.

Para pesquisadores e instituições, é essencial alcançar visibilidade e impacto, que podem ser constatados, entre outros indicadores, pelas citações dos trabalhos publicados. Portanto, é muito importante que os repositórios assegurem a interoperabilidade e a preservação digital, a fim de garantir a divulgação, visibilidade, acessibilidade e uso de longo prazo a seus conteúdos digitais.

Os resultados do estudo realizado por Eysenbach (2006) comprovam que o acesso aberto à literatura científica tem potencial para acelerar o reconhecimento e a disseminação dos resultados de pesquisa, porém seus efeitos reais são controversos. O autor realizou uma análise bibliométrica longitudinal de um recorte de artigos em acesso aberto e de acesso restrito e evidenciou, entre outras coisas, que o número médio de citações, bem como a proporção de artigos citados pelo menos uma vez foi significativamente maior no grupo de periódicos em acesso aberto.

Na medida em que são criados seguindo princípios e padrões internacionais e com metas de apoiar e dar visibilidade à pesquisa e ao ensino, os repositórios devem ser estudados com vistas ao valor potencial inexplorado, em um sentido mais amplo e que pode garantir um novo papel para as bibliotecas universitárias, estreitando as relações com grupos de pesquisas, programas de pós-graduação, pesquisadores e gestores, com novos serviços que venham apoiar a missão institucional.

Portanto, agregar serviços de valor aos repositórios institucionais requer estudos especializados, que vão além da contagem de acessos e *downloads*. Indicadores bibliométricos podem ser utilizados para provocar novos olhares em pesquisas futuras, que forneçam dados do uso real da produção científica disponível em acesso aberto e fornecer informações relevantes sobre um determinado tipo de documento ou autor.

Metodologia

Os estudos empíricos serão realizados em repositórios institucionais brasileiros, tanto no que se refere às publicações citadas quanto às citantes. Poderão ser desenvolvidos estudos com diferentes abordagens, a saber: quantitativos, por exemplo, para contagem de citações; qualitativos, quando for necessário coletar dados por meio de entrevistas com pesquisadores ou gestores de repositórios ou, ainda, quali-quantitativa.

Com relação à forma como os dados serão coletados, analisados e os resultados interpretados, os procedimentos do estudo serão de caráter teórico, buscando-se os relacionamentos entre conceitos, ideias e características do objeto estudado e empírico, adotando outros tipos de procedimentos metodológicos. Não se trata, predominantemente, de realizar estudo de caso, porém, visa retratar de forma aprofundada e exhaustiva o fenômeno das citações de artigos após sua disponibilização em um RI de acesso aberto, permitindo a comparação com outros casos, com particularidades semelhantes, ou ainda com artigos que não foram disponibilizados em acesso aberto, na mesma área e/ou no mesmo período.

A vantagem está em explorar a realidade de um determinado RI, de um periódico específico, uma área do conhecimento, um determinado programa de pós-graduação ou, ainda, um grupo pré-determinado de autores de uma instituição ou grupo de pesquisa. Estes procedimentos permitirão realizar análises não apenas estatísticas, mas os dados poderão constituir um conjunto de elementos que permitam identificar causas, efeitos e eventuais soluções para o problema da contabilização de citações distribuídas em repositórios institucionais.

Inicialmente, baseados na metodologia apresentada nos estudos de Barrueco (2001 e 2008) e Eysenbach (2006), são propostas três etapas para a execução do estudo, a saber:

- 1) Coleta de dados – inclui definir o ambiente de trabalho e restringir o número de repositórios que serão analisados para extrair os documentos citados, neste caso, os artigos de periódicos e os documentos que servirão como base dos documentos

citantes. Nesta etapa será selecionado o RI que fornecerá os dados dos artigos de periódicos para os quais se pretende verificar o impacto por meio das citações distribuídas e os repositórios de onde serão extraídos os dados dos documentos citantes. Será realizada uma análise dos índices e esquema de metadados dos repositórios selecionados, a fim de identificar possível incompatibilidade na descrição dos tipos de documentos. A seguir, será realizado o *download* do texto completo dos potenciais documentos citantes e a conversão do formato original para um formato que permita ser analisado, por exemplo: PDF (*Adobe Portable Document Format*) para ASCII (*American Standard Code for Information Interchange*) ou de HTML (*HyperText Markup Language*) para XML (*Extensible Markup Language*).

- 2) Análise – abrange procedimentos para encontrar, na seleção dos tipos de documentos citantes, a lista de referências e, posteriormente, identificar os elementos que a compõem. Neste contexto, pode-se identificar e isolar a frase ou frases com as quais o autor se referiu ao artigo citado.
- 3) *Link* – nas referências coletadas, nos documentos citantes, procede-se à identificação do artigo citado e se o mesmo foi referenciado na sua versão em formato eletrônico, dentro do ambiente de dados definido para análise. Em caso afirmativo, será estabelecido um *link* entre os identificadores do documento citante e o artigo citado.

Os dados deste estudo, por estarem concentrados em citações, limitam-se a estimar o impacto dos artigos de periódicos em outras pesquisas e não no usuário do conhecimento, mas pode-se formular hipóteses para testá-los em estudos futuros como, por exemplo: que existe uma vantagem para os pesquisadores que utilizam resultados de pesquisas disponibilizados em RI, em acesso aberto; que permite transformar conhecimento mais rapidamente; aumentar produtividade; diminuir gastos e acelerar o consumo da informação produzida em âmbito nacional e internacional.

A aplicação desta metodologia pretende preencher uma carência de estudos relativos ao impacto dos artigos de periódicos publicados por pesquisadores de instituições de ensino superior do Brasil e disponibilizados em acesso aberto, tanto pela Via Verde como pela Via Dourada. A primeira é a via do autoarquivamento, que trata do arquivamento realizado pelos próprios autores de artigos científicos já publicados ou aceitos para publicação, a partir da autorização dos editores. A segunda abrange os periódicos científicos eletrônicos cujo acesso aberto a seus conteúdos é garantido pelos próprios editores (ALVES, 2008).

Este tipo de estudo apresenta possibilidades que podem vir a fazer parte da engrenagem da gestão do conhecimento institucional. Pode sinalizar o rumo a seguir em estudos futuros, com a intenção de contribuir com o conhecimento sobre repositórios e a gestão da produção científica institucional e, por conseguinte, evidenciar como a evolução tecnológica dos sistemas de informação e as iniciativas em prol do acesso aberto contribuem

para ampliar e dinamizar o processo de comunicação científica, assim como, maximizar o impacto alcançado pelos artigos de periódicos e, como consequência, a maior probabilidade de citação.

Resultados

Os resultados a serem atingidos a partir da realização deste estudo são muitos. Destacam-se, no entanto, em uma primeira etapa, os seguintes:

- a) verificar se os documentos citantes referenciam os artigos em formato eletrônico com o *link* do repositório da instituição onde foi produzido;
- b) buscar evidências do aumento de citações dos artigos, a partir da sua disponibilização em RI de acesso aberto;
- c) propor a divulgação institucional dos resultados, de modo a incentivar sua utilização para a gestão da pesquisa e o estímulo ao depósito em RI e em acesso aberto;
- d) ampliar e adaptar a metodologia para estudos de citações de outros tipos de documentos depositados em RI, com citação distribuída;
- e) propor novos estudos, da mesma natureza, por áreas do conhecimento, por publicação periódica, por autores de uma determinada instituição, entre outros, e
- f) divulgar os resultados, por meio da publicação de artigos e trabalhos em eventos da área.

Conclusões

Entende-se que os RI não devem conter apenas ricos recursos digitais, eles também devem fornecer um ambiente que integre a prestação de serviços de informação para apoiar as atividades acadêmicas e a aprendizagem de maneira eficaz, a tomada de decisão para a gestão eficiente dos recursos aplicados nas pesquisas e maximizar a visibilidade institucional e, por conseguinte, dos seus pesquisadores. Assim, ao conceber um RI, é importante que este alie qualidade da informação nele depositada, descrição consistente dos metadados, recursos de busca apropriados à comunidade a que se destina, desenho da arquitetura de informações que influencie positivamente na construção do conhecimento e na melhoria do desempenho da comunidade acadêmica e que forneça serviços de valor agregado que provoquem mudanças no perfil de utilização dos repositórios por parte dos pesquisadores e gestores institucionais.

Realizar uma análise de citações no âmbito dos RI, a fim de verificar o impacto dos artigos de periódicos produzidos numa determinada instituição, por meio das citações recebidas após o depósito em repositório de acesso aberto, não é tarefa simples, entretanto, será útil na medida em que novos serviços sejam criados, com o intuito de melhorar e/ou diversificar os resultados das iniciativas de acesso aberto e interoperabilidade entre repositórios e entre provedores de dados e de serviços. Outra opção de melhoria também é proporcionada pela integração com outros dados disponíveis nos repositórios como, por

exemplo, os já conhecidos dados de acessos e de *downloads* e a possibilidade de criar serviços personalizados por meio da identificação dos usuários.

Referências

ALVES, V. B. A. (2008) – Open archives: via verde ou via dourada?. *Ponto de Acesso* [Em linha]. Vol. 2, Nº 2. [Consult. 08 fev. 2018]. Disponível na Internet: <URL: <http://www.portalseer.ufba.br/index.php/revistaici/article/view/1780/2172>>.

BARRUECO, J. M. (2008) – Medición del uso e impacto de documentos distribuídos através de repositorios institucionales. *BiD: Textos universitaris de biblioteconomia i documentació* [Em linha]. Nº 20. [Consult. 28 mar. 2018]. Disponível na Internet: <URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=2747404>>.

BARRUECO, J. M. (2002) – Reference Linking: un nuevo concepto para facilitar el acceso a la literatura científica. *El Profesional de la Información* [Em linha]. Vol. 11, Nº 4. [Consult. 28 mar. 2018]. Disponível na Internet: <URL: <http://eprints.rclis.org/4011/>>.

COSTA, S. M. S. (2006) – Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. *Ciência da Informação* [Em linha]. Vol. 35, Nº 2. [Consult. 26 mar. 2018]. Disponível na Internet: <URL: <http://www.scielo.br/pdf/ci/v35n2/a05v35n2.pdf>>.

EYSENBACH, G. (2006) – Citation advantage of open access articles. *PLoS Biology*. [Em linha]. Vol. 4, Nº 5 [Consult. 20 mar. 2018]. Disponível na Internet: <URL: <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0040157>>.