



## Importância do Acesso Aberto a Dados de Pesquisas na Formação de Cientistas

### Importance of Open Access to Research Data in the Training of Scientists

Andréa Silva\*, Rossana Moura\*\*

\*University of Brasília, Brasil, \*\*University of Aveiro, Portugal

#### Resumo

Apresenta estudo sobre o mapeamento inédito de pesquisadores e especialistas que trabalham com o tema do Acesso Aberto a Dados de Pesquisa (AADP) no Brasil e sua importância no fomento a pesquisas científicas e à iniciação de pesquisadores. A e-Science inova na forma do fazer ciência apresentando intensa demanda de uso das TICs e gestão dos dados de pesquisas. Por conseguinte, o ambiente informacional tem impacto direto nos padrões de comunicação científica, particularmente quanto às pesquisas cooperativas, cada vez mais corriqueiras, ao compartilhamento e reúso de recursos informacionais, bem como quanto à disseminação dos resultados de pesquisas. As metodologias utilizadas foram a pesquisa bibliográfica, entrevistas, consultas a sites de instituições de pesquisa e a bases de dados científicos, mapeamento de dados disponibilizados em sites de organizações de fomento à pesquisa no Brasil e no exterior, sites de fóruns de pesquisadores, relatórios governamentais internacionais e consultas a diretórios de repositórios. O estudo revelou que políticas voltadas ao AADP promovem, dentre outros benefícios, o estímulo à pesquisa colaborativa; a eficiência do gasto público a partir da prevenção da duplicação de esforços numa mesma área de investigação; aumentos na qualidade dos resultados de pesquisa; maior participação dos cidadãos no processo científico e aceleração do processo de inovação a partir do acesso empresarial aos dados de pesquisa, fatores relevantes na capacitação dos cientistas atuais e formação de futuros pesquisadores.

*Palavras-chave:* acesso aberto, dados de pesquisas, e-Science, formação de cientistas.

#### Abstract

It presents a study about the unpublished mapping of researchers and specialists working on Open Access to Research Data (OARD) in Brazil and its importance in the promotion of scientific research and the initiation of researchers. E-Science innovates in the form of doing science presenting intense demand for the use of ICTs and management of research data. Therefore, the informational environment has a direct impact on the patterns of scientific communication, particularly in relation to increasingly common cooperative research, the sharing and reuse of information resources, and the dissemination of research results. The methodologies used were bibliographic research, interviews, queries to sites of research institutions and scientific databases, mapping of data made available on websites of research promotion organizations in Brazil and abroad, forums of researchers' forums, government reports Repositories directories. The study revealed that policies aimed at OARD promote,

among other benefits, the stimulation of collaborative research; The efficiency of public expenditure from the prevention of duplication of effort in the same area of research; Increases in the quality of research results; Greater participation of citizens in the scientific process and acceleration of the innovating process from business access to research data, relevant factors in the training of current scientists and the training of future researchers.

*Keywords:* Open access, research data, e-Science, training of scientists.

Este trabalho tem origem no estudo acerca das novas práticas de produção colaborativa de conhecimento científico, que utilizam as Tecnologias da Informação e Comunicação (TICs) como principal ferramenta de apoio. No campo da ciência, o emprego das tecnologias computacionais vem colaborando com a simulação de fenômenos complexos em que são explorados grandes volumes e uso intensivo de dados (Albagli, Maciel & Abdo, 2015). As recentes mudanças tecnológicas, político-institucionais e as inovações sociais demandaram alterações também no fazer ciência, agora em formato menos solitário e cada vez mais colaborativo, em que pesquisadores trabalham em rede baseada na internet, conhecida como e-Science.

Associada ao significado do termo ciência aberta, a e-Science propõe transcender o campo científico, como defendido por Bourdieu (2004), promovendo maior envolvimento e interlocução da ciência com outros segmentos sociais e tipos de saberes, provocando impacto direto nos padrões de comunicação científica. Nesse contexto, destaca-se o uso intensivo de dados de pesquisa (Organisation for Economic Co-Operation and Development, [OECD], 2007). Organizações internacionais propuseram o desenvolvimento de políticas que contemplem o Acesso Aberto a Dados de Pesquisa (AADP) como forma de promover o estímulo à pesquisa colaborativa e a eficiência do gasto público, uma vez que evita-se a duplicação de esforços na mesma área de investigação (European Commission, [EC], 2012, 2016, United States Department of Agriculture [USDA], 2014, Australian Research Council [ARC], 2015.).

De acordo com a European Commission (2016), o acesso aberto às publicações científicas trata do acesso on-line, gratuito para qualquer usuário. Além dos aspectos básicos – como o direito de ler, baixar e imprimir – inclui o direito de copiar, distribuir, buscar,

associar, rastrear e selecionar dados de fontes diversas. Apesar do interesse crescente da comunidade científica, a questão dos dados abertos na ciência demanda a superação de barreiras técnicas, legais e culturais (Machado, 2015). No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) desenvolve diversas ações para a promoção do acesso aberto a publicações científicas há mais de uma década. Em 2016, o Ibiict iniciou pesquisa com o objetivo de identificar práticas, posicionamentos e estratégias de disponibilização de dados em plataformas de acesso aberto, que possam vir a subsidiar um Programa Brasileiro de Acesso Aberto a Dados de Pesquisa (PBAADP), que promova a importância e a prática do compartilhamento de dados primários de pesquisa científica em todo o território nacional, como já ocorre na comunidade científica internacional (Research Information Network [RIN], 2008, [RIN] & National Endowment for Science, Technology and The Arts [NESTA], 2010).

Este trabalho relata o estudo do Ibiict com vistas a obter mapeamento de instituições internacionais e atores brasileiros que lidam com tema, direta ou indiretamente, e como o AADP promove a qualidade dos resultados de pesquisa, a maior participação dos cidadãos no processo científico e a aceleração do processo de inovação, fatores relevantes na capacitação dos cientistas atuais e formação de futuros pesquisadores.

### **Método**

Diante do desafio de identificar demandas para o desenvolvimento de um PBAADP, a pesquisa buscou responder às seguintes questões norteadoras: Quem são os atores brasileiros que, de alguma forma, estão associados ao tema do AADP no Brasil? Quais são os temas relevantes ao debate, que auxiliarão o processo de estabelecimento de um PBAADP? Desta forma, sistematizar a informação sobre o estado da arte da pesquisa e da prática de acesso aberto a dados de pesquisa no Brasil, visando subsidiar proposta de PBAADP passou a ser o objetivo geral do trabalho.

A metodologia de coleta de dados baseou-se em pesquisa bibliográfica, entrevistas com profissionais da área, consultas a sites de instituições de pesquisa no Brasil e no exterior, coleta de dados disponibilizados em sites de organizações de fomento à pesquisa nacional e internacional, sites de fóruns de pesquisadores, relatórios governamentais internacionais e consultas aos diretórios de repositórios. Para entender a complexidade dos fenômenos e objetos estudados foi utilizada combinação de métodos, principalmente na fase de coleta e análise de dados. Abordagens quantitativas e qualitativas permitiram analisar as representações e opiniões, amparadas por dados estatísticos que possibilitaram complementação necessária para avaliar a complexidade dos fenômenos e fatos específicos das amostras (Goldenberg, 2007).

### **Participantes**

O processo do mapeamento de áreas e grupos de pesquisa brasileiros de AADP foi realizado a partir de consulta ao Diretório de Grupos de Pesquisas do

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Esse diretório possui cadastro com mais de 35 mil profissionais, vinculados a grupos de pesquisa de instituições brasileiras cadastradas na Plataforma Lattes, atuantes em diversas áreas do conhecimento.

Já para a realização do mapeamento das instituições internacionais que lidam com o AADP, foram consideradas organizações governamentais e não governamentais, agências de fomento à pesquisa, universidades, fundações, entidades de classe dentre outros, relacionadas ao tema de AADP e que têm representatividade na comunidade científica internacional.

### **Procedimento**

O processo de pesquisa teve início em setembro de 2016, com a realização de estudos e debates sobre temas pertinentes ao AADP, em bibliografia previamente selecionada pelos orientadores do projeto, em que se buscou entender os conceitos e identificar experiências internacionais de AADP passíveis de aproveitamento no desenvolvimento de um programa nacional dessa temática. Os debates foram profícuos também no processo de definição de fontes de informação e metodologia para o mapeamento de instituições e atores que atuam no país, agrupados por segmentos de atuação definidos a partir de suas linhas de pesquisa, agora categorizadas como subáreas relacionadas ao AADP.

A segmentação dos atores foi feita a partir de consulta ao Diretório de Grupos de Pesquisas do CNPq, em que foram usados os parâmetros de pesquisa "Acesso Aberto", "Dados de Pesquisa" e "Dados Abertos". De forma complementar, foram realizadas buscas no Portal de Periódicos Capes e Google Acadêmico, no intuito de identificar pesquisadores brasileiros que publicaram trabalhos indexados com as mesmas palavras-chave e que, portanto, estariam associados ao tema do AADP. Como critério de filtro, foram definidas buscas em publicações recentes, dos últimos cinco anos, dentro dos temas parametrizados. Após análise detalhada em cada um dos grupos quanto às linhas de pesquisas e atividades realizadas, concluiu-se que em nenhum deles havia menção direta ao tema AADP ou a dados de pesquisa, mais especificamente.

Ampliando as buscas, foram realizadas entrevistas com alguns desses profissionais, selecionados pelo critério de produção científica, contatos profissionais prévios e proximidade física, que representaram os segmentos mapeados na primeira fase: Políticas Públicas para Acesso à Informação, Agropecuária, Saúde e Gestão de Dados. O propósito das entrevistas foi identificar especificidades sobre as melhores práticas em AADP das diversas áreas do conhecimento que fossem aplicáveis à realidade nacional, revelando potenciais idiosincrasias, ampliando a possibilidade de adesão dos principais públicos-alvo – pesquisadores, políticos e gestores de instituições de fomento à pesquisa em âmbito nacional.

No mapeamento das instituições internacionais que lidam com o AADP foram feitas consultas a sites de diversas instituições relacionadas à temática. Parte das consultas apontavam para outros sites, que indicavam

repositórios e/ou outras instituições. Para validar as instituições, foi consultada a seção do menu *about us*, em que procurou-se associação efetiva ao tema e eventuais categorias que se encaixavam dentro da temática do AADP. Nessa fase, os sites identificados foram agrupados em quatro grandes áreas: União Europeia, Reino Unido, Estados Unidos e Multicontinentes, que englobava países de continentes diferentes. Para efeito deste estudo, foram consideradas organizações governamentais e não governamentais, agências de fomento à pesquisa, universidades, fundações, entidades de classe dentre outros, relacionadas ao tema de AADP.

**Resultados**

Na primeira etapa do estudo foram identificadas as principais instituições internacionais e mapeados alguns atores do cenário nacional que lidam com o tema AADP. Identificados os atores de projeção nacional que lidam com o tema AADP, foi realizada sistematização das informações coletadas, o que gerou uma base de dados que apresenta lista de atores e categorias de atuação relacionadas ao tema. Nessa fase inicial, foram identificados 67 atores provenientes de organizações distintas, predominantemente associados a instituições públicas de pesquisa, como demonstrado no Gráfico 1.

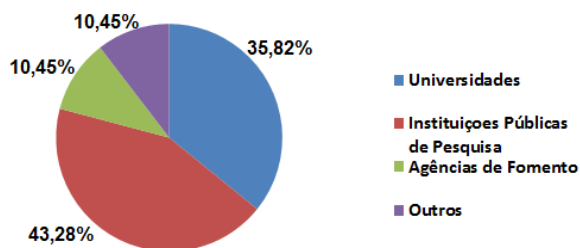


Gráfico 1 - Mapeamento de Atores: Instituições Participantes

O Gráfico 2 mostra a distribuição das instituições internacionais, mapeadas entre as quatro localizações mencionadas. Nota-se que a maioria das instituições estão localizadas no Reino Unido, situação que pode sofrer algumas alterações a partir da saída do Reino Unido da União Europeia (*Brexit*).

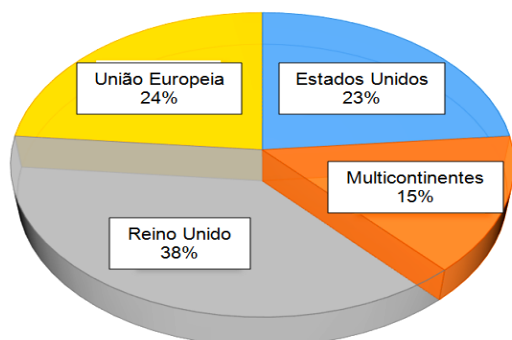


Gráfico 2 – Localização das Instituições

Em relação às áreas do conhecimento que as instituições da amostra se dedicam, foram identificadas

outras, para além das que trabalham com todas as áreas acadêmicas – são as instituições especializadas em apenas uma área de conhecimento e, em alguns casos, em poucas áreas correlacionadas. Foram mapeadas 10 áreas entre as instituições especializadas, a saber: Arqueologia, Astrofísica/Astronomia, Ciências Ambiental/Natural, Ciência da Informação, Ciência e Tecnologia, Ciências Sociais e Humanidades, Ciências Sociais e Políticas, Dados Numéricos e Geoespaciais, Graciologia e História.

O mérito principal da pesquisa foi a produção de mapeamento inédito de atores – pesquisadores, especialistas e gestores públicos – que trabalham com o tema do AADP no Brasil, incluídos na tabela de Mapeamento de Atores Nacionais. Da mesma forma, as instituições internacionais relacionadas ao tema AADP, em qualquer que sejam as áreas do conhecimento, também foram registradas na tabela Instituições Internacionais Relacionadas ao Tema AADP (Moura & Silva, 2017). O esforço de pesquisa visava ainda à identificação das principais áreas de interesse dentro no tema que foram devidamente categorizadas em subáreas, baseadas nas linhas de pesquisas dos atores identificados. A proposta é que, a partir destas categorias, o IBICT possa dar sequência aos estudos criando grupos de interesse (GIs) para discussão e sistematização de grupos de trabalho (GTs). Posteriormente, os GTs devem conduzir novas pesquisas, ampliando o debate e produção científica em torno do AADP.

**Discussão**

Dentre os diversos fatores que impulsionam a e-Science, pode-se identificar que o desenvolvimento das TICs vem amparando a exploração de grandes volumes de dados oriundos da multi e interdisciplinaridade dessa nova forma de fazer ciência. Cada vez mais, explora-se o diálogo entre teorias, experimentos e simulações, o que exige uma diversidade de tecnologias e colaboração, particularmente quanto à formulação de modelagens ou metodologias para análise dos dados de pesquisas (Appel, Maciel & Albagli, 2016).

As comunidades de práticas científicas, os padrões internacionais, estratégias e características de sistemas de informação para o compartilhamento de dados têm apontado a importância de incluir profissionais que lidam com o AADP na definição de políticas nacionais para a área (OECD, 2007, RIN, 2008, RIN & NESTA, 2010), haja vista iniciativas não governamentais que têm gerado contribuições proeminentes para evolução do contexto do compartilhamento de dados de pesquisas, como o caso do *Panton Principles* (PP, 2010) – site criado por pesquisadores com o propósito de apoiar e angariar contribuições na disseminação dos princípios fundamentais da causa do AADP.

O estudo demonstrou ainda que, as instituições internacionais, guiadas por um entendimento comum da Europa – unidos por fatores históricos, sociais, culturais e políticos, bem como por razões científicas e econômicas – vêm promovendo o intercâmbio de informações e experiências entre academias, além de incentivarem a excelência e a ética na condução da

pesquisa, bem como abordagens inter, trans e multidisciplinares em todos os empreendimentos científicos. Quanto ao mapeamento preliminar de atores nacionais relacionados direta ou indiretamente com AADP, a pesquisa possibilitou identificar que cultura de uso do AADP no Brasil ainda é pouco disseminada, embora existam inúmeras iniciativas voltadas ao acesso aberto.

Nesse contexto, foi identificada a relevância do papel das TICs na formação de pesquisadores, de forma a promover a integração entre diversos profissionais e setores da sociedade que interferem na prática científica-tecnológica em atividades diversas, desde o planejamento até a avaliação dos resultados finais (Trigueiro, 2001). Vale salientar ainda que, para lidar com as TICs e os desafios da interdisciplinaridade, os pesquisadores são desafiados rotineiramente pela enorme quantidade de informações. Para tanto, necessitam ser letrados digital e informacionalmente. O letramento digital viabiliza a capacidade de lidar com o universo computacional. Esse processo é construído a partir da associação existente entre o letramento digital e o letramento alfabético, como no caso da utilização de programas que realizam processamento de texto em computadores – desde a escrita e edição na tela, construção de frases e parágrafos, dentre outros. O letramento informacional vai além, possibilitando o desenvolvimento de competências na busca e uso da informação, de forma crítica, disponível em vários suportes e canais, impressos ou eletrônicos. (Azevedo & Gasque, 2017).

O estudo realizado demonstrou que, quanto maior a integração promovida entre integrantes da comunidade científica, maior é a expectativa de rapidez na produção e difusão da informação e do conhecimento. Tal fato mostra-se relevante para a formação de novos pesquisadores, pois o pesquisador que compartilha estudos, trabalhando a reflexão, planejamento e desenvolvimento das questões, com mesmo afinco que produz solitariamente, está caminhando na direção de um saber interdisciplinar, onde há a coexistência da parceria e da solidão, que exige comprometimento do pesquisador com os parceiros do grupo de pesquisa (Fazenda, José & dos Santos, 2015).

### Referências

- Appel, A.L. ; Maciel, M.L. & Albagli, S. (2016). A e-Science e as novas práticas de produção colaborativa de conhecimento científico. *Revista Internacional de Ciencia Y Sociedad*, v. 3, pp. 41-52. Obtido em 20 junho, 2017. Obtido em 14 de junho, 2017. Recuperado de <http://journals.epistemopolis.org/index.php/cienciayso/article/view/470/78>.
- Australian Government. Australian Research Council, [ARC], (2015). Open Access Policy. Versão 2015.1. Obtido em 24 novembro, 2016. Recuperado de [http://www.arc.gov.au/sites/default/files/filedepot/Public/Policy%20&%20Strategy/ARC%20Open%20Access%20Policy/ARC\\_Open\\_Access\\_Policy\\_V2015.1\\_17Aug15.pdf](http://www.arc.gov.au/sites/default/files/filedepot/Public/Policy%20&%20Strategy/ARC%20Open%20Access%20Policy/ARC_Open_Access_Policy_V2015.1_17Aug15.pdf).
- Azevedo, I.C.M. & Gasque, K.C.G.D. (2017). Contribuições dos letramentos digital e informacional na sociedade contemporânea. *Transinformação*, 29(2), pp. 163-173. <https://dx.doi.org/10.1590/2318-08892017000200004>.
- Bourdieu, P.( 2004). Usos sociais da ciência: por uma sociologia clínica do campo científico. São Paulo: UNESP.
- European Commission, [EC], (s.d.). Open Access Pilot in FP7. Obtido em 20 agosto, 2016. Recuperado de [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/open-access-pilot\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-pilot_en.pdf).
- \_\_\_\_\_. (2012). On access to and preservation of scientific information. Obtido em 20 novembro, 2016. Recuperado de [https://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/recommendation-access-and-preservation-scientific-information\\_en.pdf](https://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf)
- \_\_\_\_\_. (2016). Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. Version 3.0. Obtido em 20 de agosto, 2016. Recuperado de [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf).
- Fazenda, I.C.A., José, M.A.M., & dos Santos, C.A.M. (2016). Formar Pesquisadores Interdisciplinares. *Revista Ciências Humanas*. v. 9, n 1, edição 16, pp. 62 – 69. Obtido em 12 de junho de 2017. Recuperado de <http://www.rchunitau.com.br/index.php/rch/article/view/276>.
- Gasque, K. C. G. D. (2017). Metacognição no processo de letramento informacional. *Revista Brasileira de Biblioteconomia e Documentação*, v. 13, pp. 177-195. Obtido em 14 de fevereiro de 2017. Recuperado de <http://repositorio.unb.br/handle/10482/23426>.
- Goldenberg, M. (2007). *A arte de pesquisar: como fazer pesquisa qualitativa em ciências sociais*. 3. ed. Rio de Janeiro: Record.
- Instituto Brasileiro de Informação em Ciência E Tecnologia, [IBICT], (2016). Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã. Obtido em 19 set. 2016. Recuperado de <http://livroaberto.ibict.br/Manifesto.pdf>.
- Machado, J. Dados abertos e ciência aberta. (2015). Em Albagli, S.; Maciel, M.L. & ABdo, A.H. (Orgs.). *Ciência aberta, questões abertas* [versão eletrônica pdf]. pp. 201-229. Obtido em 01 de setembro de 2016. Recuperado de <http://livroaberto.ibict.br/handle/1/1060>.
- Moura, R.O.C.. & Silva, A.C.M. (2017). Desenvolvimento, modernização e criação de novas estratégias para organização, preservação e acesso à informação científica, tecnológica e social no Brasil: compartilhamento em Acesso Aberto dos Dados de Pesquisas. Relatório final de atividades do projeto – CNPq. IBICT. [Manuscrito inédito].
- Organisation for Economic Co-Operation and Development, [OECD], (2007). Principles and Guidelines for Access to Research Data from Public Funding. Obtido em 17 de agosto de 2016. Recuperado de <http://www.oecd.org/sti/sci-tech/38500813.pdf>.

- Panton Principles [PP], (2010). Principles for Open Data in Science. Recuperado de <https://pantonprinciples.org/>.
- Research Information Network, [RIN], (2008). To Share or not to Share: Publication and Quality Assurance of Research Data Outputs. Obtido em 01 de setembro de 2016. Recuperado de <http://www.rin.ac.uk/system/files/attachments/To-share-data-outputs-report.pdf>.
- \_\_\_\_\_; National Endowment for Science, Technology and The Arts, [NESTA], (2010). Open to All? Case studies of openness in research. Obtido em 17 de agosto de 2016. Recuperado de [http://www.rin.ac.uk/system/files/attachments/NESTA-RIN\\_Open\\_Science\\_V01\\_0.pdf](http://www.rin.ac.uk/system/files/attachments/NESTA-RIN_Open_Science_V01_0.pdf).
- Research Data Alliance, [RDA]. [s.d]. Creating and Managing RDA Groups. Recuperado de <https://www.rd-alliance.org/creating-and-managing-rda-groups.html>.
- Trigueiro, M.G.S. (2001). A formação de Cientistas: Necessidades e Soluções. Em Baumgarten, M.. (Org.). A Era do Conhecimento: Matrix ou Ágora?. pp. 61-70. Porto Alegre / Brasília: Editora da Universidade UFRGS / Editora UnB, v. 1.
- United States Department of Agriculture, [USDA], (2014). Implementation Plan to Increase Public Access to Results of USDA-Funded Scientific Research. Obtido em 14 de janeiro de 2017. Recuperado de <https://www.usda.gov/documents/USDA-Public-Access-Implementation-Plan.pdf>.

### **Agradecimentos**

À diretora do Ibict, Cecília Leite, pelo apoio institucional à pesquisa. Às orientações de Cristiane Rauen e Leonardo Lazarte, respectivamente Tecnologista e Coordenador-Geral de Tecnologia da Informação no Ibict na época do estudo.

Esta pesquisa foi realizada no âmbito do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), órgão do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) do Brasil e financiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).