

# cadernos

2016  
2

# BAD

Atas da  
7.<sup>a</sup> Conferência Luso-Brasileira  
sobre Acesso Aberto

ISSN 0007-9421



9 770007 942009 >



associação portuguesa de  
bibliotecários, arquivistas e documentalistas

---

## Ficha Técnica

---

**Cadernos BAD, 2016, N 2, jul-dez**  
**Atas da 7ª Conferência Luso-Brasileira sobre Acesso Aberto**

### **Comissão Organizadora**

Ângelo Fonseca, Instituto Politécnico de Viseu, Portugal  
Bianca Amaro, Instituto Brasileiro de Informação em Ciência e Tecnologia – Ibict, Brasil  
Eloy Rodrigues, Universidade do Minho, Portugal  
João Mendes Moreira, FCCN, Portugal  
José Carvalho, Universidade do Minho, Portugal  
José Santos Costa, Instituto Politécnico de Viseu, Portugal  
Tainá Batista, Instituto Brasileiro de Informação em Ciência e Tecnologia – Ibict, Brasil  
Vasco Vaz, FCT, Portugal  
Washington Luís Ribeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia – Ibict, Brasil

### **Comissão Científica**

Anabela Serrano, Escola Superior de Estudos Industriais e de Gestão do Instituto Politécnico do Porto, Portugal  
Angélica Conceição Dias Miranda, Universidade Federal do Rio Grande – FURG, Rio Grande/RS, Brasil  
Bianca Amaro, Instituto Brasileiro de Informação em Ciência e Tecnologia – Ibict, Brasil  
Carlos Fiolhais, Universidade de Coimbra, Portugal  
Caterina Groposo Pavão, Universidade Federal do Rio Grande do Sul, Brasil  
Cristina Ribeiro, DEI– Faculdade de Engenharia d Universidade do Porto/ INESC TEC, Portugal  
Daniel Silva, Instituto Politécnico de Viseu, Portugal  
Eloísa da Conceição Príncipe de Oliveira, Ministério da Ciência, Tecnologia e Inovação – MCTI / Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, Brasil  
Fernando César Lima Leite, Faculdade de Ciência da Informação, Universidade de Brasília, Brasil  
Flavia Goulart Garcia Rosa, Universidade Federal da Bahia, Brasil  
Joaquim Sousa Pinto, Universidade de Aveiro, Portugal  
José dos Santos Costa, Instituto Politécnico de Viseu, Portugal  
Lídia Brandão Toutain, Universidade Federal da Bahia, Brasil

---

Lilian Alvares, Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, Brasil  
Lúcia Pato, Instituto Politécnico de Viseu, Portugal  
Luís Fernando Sayão, CNEN, Brasil  
Manuela Barreto Nunes, Universidade Portucalense (Biblioteca Geral), Portugal  
Maria João Amante, ISCTE-IUL – Instituto Universitário de Lisboa, Portugal  
Maria Manuel Borges, Universidade de Coimbra, Portugal  
Miguel Ferreira, KEEP Solutions, Portugal  
Sandra Maria Gouveia Antunes, Instituto Politécnico de Viseu, Portugal  
Sely Costa, Universidade de Brasília, Brasil  
Simone R Weitzel, UNIRIO/Escola de Biblioteconomia, Brasil  
Sueli Mara Ferreira, Universidade de São Paulo, Brasil  
Suzana Amaro, Instituto Politécnico de Viseu, Portugal  
Maria Teresa Antas de Barros, Instituto Politécnico de Viseu, Portugal  
Teresa Costa, FCT/FCCN, Portugal

---

## Sumário

### Editorial

7ª ConFOA: Do Acesso Aberto à Ciência Aberta	1
<i>A Comissão Organizadora da 7ª Conferência Luso-Brasileira Acesso Aberto</i>	

### Comunicações

A influência das revistas científicas de acesso aberto para o depósito e publicação dos dados de pesquisa	3
<i>Teila de Oliveira Carvalho</i>	
Ritmos da informação/comunicação de ciência dos centros de investigação em Portugal	16
<i>Fábio Ribeiro</i>	
Marketing digital em portais de periódicos científicos de acesso aberto	31
<i>Juliana Aparecida Gulka, Elaine Rosangela de Oliveira Lucas, Ronaldo Ferreira de Araújo</i>	
Peças de teatro manuscritas (séculos XVIII e XIX) em acesso aberto da Escola Superior de Teatro e Cinema: projeto apoiado pela Fundação Calouste Gulbenkian	44
<i>Luísa Maria Lousã Marques</i>	
Repositórios de dados de pesquisa para as ciências da saúde	57
<i>Lucas Paganine, Michelli Costa</i>	
Os repositórios das instituições de ensino superior portuguesas : Estudo comparativo	71
<i>Maria Eduarda Pereira Rodrigues, Maria João Amante, Clarisse Pais, Susana Lopes, Teresa Segurado, António Moitinho Rodrigues</i>	
Repositórios de dados de pesquisa no mundo	80
<i>Michelli Costa, Tiago Braga</i>	
Panorama do autoarquivamento nos repositórios institucionais portugueses	96
<i>Viviane Santos de Oliveira Veiga, Luis Guilherme Gomes Macena, Cícera Henrique da Silva, Maria Manuel Borges</i>	
Acesso aberto, direitos autorais e a inclusão sociocultural das pessoas portadoras de deficiência	106
<i>Allan Rocha Souza, Alexandre de Serpa Pinto Fairbanks, Wemerton Monteiro Souza</i>	
O retrato situacional das revistas científicas brasileiras	116
<i>Danyelle Mayara Silva, Teila de Oliveira Carvalho, Bianca Amaro, Tainá Batista de Assis</i>	
Integrar uma rede de repositórios no ecossistema de gestão de ciência nacional	125



*José Carvalho, João Mendes Moreira, Cátia Laranjeira, Fernando Ribeiro, Paulo Graça, Paulo Lopes*

## **Pecha Kucha**

Perfil dos usuários do repositório institucional da Universidade Federal da Bahia <i>Flávia Rosa, Susane Barros, Rodrigo França Meirelles</i>	134
Repositórios temáticos e memória: a constituição da educação em saúde no Brasil por meio da trajetória de Virgínia Schall <i>Denise Nacif Pimenta, Leandro da Conceição Borges, Kizi Mendonça de Araújo, Maria Cristina Soares Guimarães, Cícera Henrique da Silva</i>	145
Avaliação de repositórios institucionais brasileiros: Uma proposta de método de avaliação <i>Ana Kelly Alves Lameira</i>	153
Implementação de métricas alternativas no portal oasisbr <i>Danielly dos Santos Ribeiro, Luiza Moreira Camargo</i>	168
Gestão de dados: Sobreposições ou convergências entre infraestruturas? <i>Pedro Moura Ferreira, Bárbara Rodrigues</i>	179
Acesso aberto à informação científica: diretrizes, políticas e modelos de repositórios científicos para Moçambique <i>Horácio Francisco Zimba, Ranito Zambo Waete, Aidate Mussagy</i>	187
Social Dendro: Aplicação de técnicas das redes sociais à gestão de dados de investigação <i>João Rocha da Silva, Nelson Pereira</i>	202
O cenário dos livros eletrônicos científicos em acesso aberto <i>Tainá Batista de Assis</i>	212
Integração de plataformas institucionais: o caso EUDAT <i>Ricardo Carvalho Amorim, Fábio Silva</i>	220
Promover o acesso aberto com o Open Journal System : duas instituições, duas realidades <i>Paula Seguro-de-Carvalho, Rui Teófilo</i>	232
Vocabulários controlados na descrição de dados de investigação no Dendro <i>Yulia Karimova, João Aguiar Castro</i>	241
Projeto TAIL—Gestão de dados de investigação da produção ao depósito e à partilha (resultados preliminares) <i>Cristina Ribeiro, João Rocha da Silva, João Aguiar Castro, Ricardo Carvalho Amorim, João Correia Lopes</i>	256
Bibliotecas, repositórios institucionais e direitos autorais <i>Bianca Amaro</i>	265
O RI no contexto da gestão da produção científica: Proposta de um questionário para aferição de níveis de satisfação e novas necessidades <i>Maria João Pinto, Sofia Fernandes</i>	275
Construção de um repositório de dados oceanográficos <i>Ricardo Amorim, João Castro, Inês Garganta, Artur Rocha, Gabriel David</i>	281

## 7ª ConfOA: Do Acesso Aberto à Ciência Aberta

Este número especial dos Cadernos BAD é inteiramente dedicado à publicação da maioria dos trabalhos apresentados nos formatos de Comunicação e Pecha Kucha na [7ª Conferência Luso-Brasileira de Acesso Aberto](#) (ConfOA), que se realizou no Instituto Politécnico de Viseu, nos dias 2, 3 e 4 de novembro de 2016.

Dando continuidade ao assinalável êxito das edições anteriores, a realização do evento resultou da parceria estabelecida entre o próprio Instituto Politécnico de Viseu, a Fundação para a Ciência e a Tecnologia (FCT), a Universidade do Minho – Serviços de Documentação (SDUM) e o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), envolvendo perto de 200 participantes de 6 países (Brasil, Espanha, França, Moçambique, Portugal e Reino Unido).

A ConfOA, com realização alternada entre Portugal e Brasil, tem como objetivo reunir as comunidades portuguesa e brasileira, que desenvolvem atividades de investigação, desenvolvimento, gestão de serviços e definição de políticas relacionadas com o Acesso Aberto ao conhecimento e com a ciência aberta, com o propósito de promover a partilha, discussão e divulgação de conhecimentos, práticas e investigação sobre estas temáticas, em todas as suas dimensões e perspetivas.

A escolha do Instituto Politécnico de Viseu para acolher a 7ª ConfOA, não só assegurou a diversidade geográfica da realização deste evento, como proporcionou excelentes condições e um acolhimento exemplar a todos os participantes. A realização da ConfOA pela primeira vez no interior de Portugal registou um número bastante elevado de submissões de Comunicações, *Pechas Kuchas* e Pósteres. Foram submetidas 83 propostas: 55 do Brasil, 27 de Portugal e uma de Moçambique, tendo sido avaliadas 75, ou seja 90% das submissões. As restantes 8 não se enquadravam nos temas e critérios do evento.

Para além das sessões de apresentação dos trabalhos selecionados, o programa da ConfOA foi organizado em torno de três sessões: sessão inaugural sobre Ciência Aberta, com Stephen Curry – Orador convidado da *Imperial College London*; uma mesa sobre o Acesso Aberto e Ciência Aberta na Lusofonia, com Maria Fernanda Rollo – Secretária de Estado da Ciência, Tecnologia e Ensino Superior, Bianca Amaro – em representação do Instituto Brasileiro de Informação em Ciência e Tecnologia, João Nuno Ferreira – em

---

representação da Fundação para a Ciência e Tecnologia e Horácio Francisco Zimba docente na Universidade Eduardo Mondlane e, por último, uma sessão dedicada à Publicação em Acesso Aberto com a participação de Martin Eve da *Open Library of Humanities*, Flávia Rosa da Editora da Universidade Federal da Bahia e Delfim Leão da Imprensa da Universidade de Coimbra.

Os trabalhos que aqui se publicam resultam de 11 das 12 Comunicações e 15 das 17 *Pechas Kuchas* apresentadas na 7ª ConfOA, abrangendo um largo espectro temático, que vai desde a gestão de dados científicos, até à publicação de revistas científicas ou as políticas de Acesso Aberto, passando pelas questões dos direitos de autor, da bibliometria, da interoperabilidade e dos sistemas de gestão de informação científica.

Gostaríamos de agradecer a disponibilidade dos Cadernos BAD em acolher, uma vez mais, a publicação de todos os trabalhos anteriormente referenciados, contribuindo para a partilha e atualização dos últimos desenvolvimentos do Acesso Aberto, particularmente, no mundo lusófono.

Por fim, convidamos todos os leitores desta publicação e demais interessados a participarem na 8ª Conferência Luso-Brasileira sobre Acesso Aberto que se irá realizar no Brasil, na [Fundação Oswaldo Cruz](#) (Fiocruz), nos dias 4, 5 e 6 de outubro de 2017.

Contamos com todos e com cada um em particular!

*A Comissão Organizadora da 7ª Conferência Luso-Brasileira sobre Acesso Aberto*

---

# A influência das revistas científicas de acesso aberto para o depósito e publicação dos dados de pesquisa

---

Teila de Oliveira Carvalho

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

[teillacarvalho@ibict.br](mailto:teillacarvalho@ibict.br)

## Resumo

Toda pesquisa científica é pautada no levantamento, elaboração e análise de dados que geram conteúdos científicos. Com a finalidade de dar visibilidade a esta questão, a disponibilização dos dados de pesquisa tem ganhado cada vez mais notoriedade dentro da comunidade científica. E como agente da comunicação científica, algumas revistas já têm apoiado e incentivado o compartilhamento de dados como requisito para publicação. Diante desse cenário, este estudo buscou analisar a proporção e a forma como as revistas em acesso aberto das áreas de Ciências e Medicina do Brasil e de Portugal indexadas no DOAJ têm tratado a questão do depósito e publicação dos dados. Como resultado, percebeu-se que, mesmo que de forma ainda discreta, as revistas brasileiras e portuguesas de acesso aberto têm se mobilizado para que os dados estejam à disposição da comunidade de forma completa e acessível. Conclui-se que a maioria das publicações que exigem ou sugerem que os dados sejam disponibilizados reconhecem a importância do acesso ao conteúdo completo das pesquisas.

**Palavras-chave:** Dados abertos, Dados de pesquisa, Revistas Científicas, Banco de dados, Acesso aberto

## The influence of open access scientific journals for the deposit and publication of research data

### Abstract

All scientific research is based in the survey, preparation and analysis of data that generate scientific content. In order to give visibility to this issue, the availability of research data has increasingly gain notoriety within the scientific community. And as agents of scientific communication, some journals already have supported and

encouraged the sharing of data as a requirement for publication. In this scenario, this study investigated the proportion and how the Brazil and Portugal 's open access journals on areas of Science and Medicine indexed in DOAJ have addressed the issue of deposit and publication of data. As a result, it could be noticed that, even if still discreet, the open access journals of Brazil and Portugal have mobilized themselves so that the data is available completely and accessible to the community . It is concluded that most publications that require or suggest that the data are available recognize the importance of access to the full content of the research.

**Key-words:** Open data, Research data, Scientific journals, Database, Open access

## A ciência e os dados de pesquisa

A ciência é formada por diversos fatores e processos que trabalham para a geração de um mesmo produto, o conhecimento. Todas as etapas para a transformação de uma pesquisa em algo acessível à comunidade científica estão em constante adaptação às necessidades do contexto vivido pela sociedade.

Dizer que a ciência está evoluindo, o que envolve, entre diversos outros fatores, o desenvolvimento da tecnologia, significa dizer que seus diferentes ramos estão em um processo de transformação. Nesse cenário, Molloy (2011) acredita que o surgimento da *web* permitiu, além de novas possibilidades de compartilhamento, colaboração e análise, que a ciência explorasse novas formas de investigação. Cada etapa de uma pesquisa agora se destaca de uma forma individual para que não só o produto final seja aproveitado, mas todos os elementos sejam vistos como algo que pode gerar novos conhecimentos.

Além disso, a busca pelo acesso aberto às pesquisas científicas também tem gerado uma alteração em todas as fases da ciência. Esse movimento não mais solicita apenas o acesso aos resultados finais, mas tem influenciado a sociedade a requerer também que os dados gerados e obtidos durante todo o processo da pesquisa estejam disponíveis livremente aos demais pesquisadores e, por isso, segundo Piwowar, Day e Fridsma (2007), é um fator que pode aumentar a pressão para o compartilhamento de dados.

Sayão e Sales (2014) acreditam que os limites da busca pelo acesso aberto se expandiram com o que a “Declaração de Berlim sobre o Acesso Aberto ao Conhecimento em Ciências e Humanidades”, que foi publicada em 2003, definiu como contribuições de acesso livre. Nessa declaração os “resultados de pesquisas científicas originais, dados brutos e metadados, fontes originais, representações digitais de materiais pictóricos e gráficos e material acadêmico multimídia” são consideradas contribuições que devem está livre de barreiras de acesso.

« A expansão do conceito de acesso livre – um pilar de importância crítica para a prática de uma ciência mais aberta – não esta circunscrita somente as publicações acadêmicas tradicionais, como são os artigos de periódicos; suas demandas avançam para outros conteúdos que incluem, de forma privilegiada, a disponibilização aberta e de forma inteligível de dados de pesquisa.» (Sayão e Sales, 2014, p. 77)

Dessa forma, é perceptível que a questão sobre o acesso aos dados de pesquisa está ganhando cada vez mais destaque e passa a ser um dos grandes centros de discussão dentro da comunidade científica. Hoje já não basta publicar o produto final das pesquisas, disponibilizar os dados utilizados já está se tornando uma exigência da sociedade para que todo estudo esteja propício a reutilização. Heidorn (2008) acredita que o princípio da replicabilidade é a base para a ciência. E Lehnert e Hsu (2015) e Molloy (2011) afirmam que o processo científico se torna mais eficiente e benéfico para a sociedade quando os dados estão disponíveis abertamente e de forma útil, tornando a investigação científica transparente e reproduzível.

### **O que são dados de pesquisa?**

Toda produção científica inicia-se tendo como base o levantamento e análise de dados. A literatura afirma que os dados constituem a base de toda a teoria científica (Heidorn, 2008; Butlin, 2011; Whitlock, 2011 e Callaghan *et al*, 2013). Dessa forma, a pesquisa é pautada na elaboração de dados que geram conteúdos científicos. Molloy (2011) acredita que o levantamento, a análise, a publicação, a reanálise, a crítica e a reutilização dos dados é o que constrói a ciência.

Percebe-se ainda que a pesquisa científica produz e coleta dados que são muito variados e heterogêneos, que possuem diversos formatos e passam por diferentes processos dependendo de cada disciplina em que está inserido e dos objetivos da própria pesquisa. Portanto, o conceito de dados é muito amplo e complexo por envolver uma grande gama de contextos.

Marques (2014) apresenta um conceito generalizado para dados de pesquisas quando afirma que eles são a massa de informações primárias que serve como base para as conclusões do artigo quando são analisados e interpretados. Já o relatório da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) de 2007 conceitua dados de pesquisa como registros factuais (escores numéricos, documentos textuais, imagens e sons) utilizados como materiais originais para a investigação e que são habitualmente aceitos como recursos necessários para validar os resultados da pesquisa científica pela comunidade. Borgman (2010) cita o conceito de dados de um relatório do Conselho Nacional de Pesquisa para mostrar a sua complexidade, pois nesse relatório “dados são fatos, números, letras e símbolos que descrevem um objeto, ideia, condição, situação, ou outros fatores”. Além desses tipos de dados, Sturges *et al* (2014) e Sayão e Sales (2015) acrescentam também vídeos, áudio, software, algoritmos, equações, animações, modelos, simulações, dados geográficos, mapas geológicos, ontologias, o conteúdo da web, modelos de dados e muito mais.

Além da diversidade de conceitos para dados, entende-se que eles possuem uma variedade de funções. Sayão e Sales (2015) apontam ainda que a visão dos dados como

apenas subprodutos dos processos de pesquisa está sendo alterada devido ao reconhecimento da importância informacional que eles possuem dentro da ciência. Os dados não são mais vistos apenas como um levantamento necessário para fundamentar uma pesquisa, eles são uma rica fonte de descobertas e de geração de conhecimento. E por isso Uhler e Schröder (2007) acreditam que o uso dos dados já não está apenas vinculado ao seu propósito original e seus produtores, mas já podem ser considerados recursos autônomos.

Sabendo-se que a função da ciência e de sua divulgação é a geração de novos conhecimentos, tem-se os dados como parte fundamental desse processo que, quando acessíveis, geram uma série de vantagens para a comunidade científica. Callaghan *et al* (2013) acreditam que a internet tem auxiliado para que conjuntos de dados estejam disponíveis para qualquer pessoa usá-los mais rapidamente e facilmente. Mas, para que isso ocorra, é necessária uma movimentação de todos os agentes da comunidade científica em prol do compartilhamento de dados de forma a incentivar os pesquisadores a depositarem e publicarem seus dados em meios confiáveis e que garantam a durabilidade e acessibilidade deles. Para isso, muitas barreiras precisam ser quebradas principalmente entre os próprios pesquisadores. É necessário mostrar para a sociedade científica os benefícios que o compartilhamento de dados pode gerar para a ciência.

### **O compartilhamento de dados de pesquisa**

A disponibilização dos dados de pesquisa tem ganhado cada vez mais notoriedade dentro da comunidade científica. Antes a discussão que era voltada apenas para a visibilidade dos resultados, agora tem garantido aos dados um papel de destaque como um fator importante de propagação científica.

Por entender que o desenvolvimento da ciência vai além da publicação dos resultados finais de uma pesquisa, a própria sociedade espera que os pesquisadores disponibilizem seus dados. Atualmente busca-se um aprofundamento das pesquisas através da análise de todo o conjunto de dados produzidos durante o processo de criação. Outro fator determinante para essa movimentação é o fato exposto por Costello (2009), Chavan e Penev (2011) onde a grande maioria dos dados são frutos de pesquisas financiadas com recursos públicos e por isso deveriam ser publicados, citados, utilizados e reutilizados por quem tivesse interesse em fazê-lo. Parr (2007) afirma ainda que apenas o fato de os dados terem um possível valor para a comunidade científica já deveria ser uma razão para exigir a sua publicação.

Hoje a busca pelo compartilhamento de dados vai além da ideia de preservar e comprovar aquilo que foi produzido. Assim como a publicação dos resultados por meio de artigos ou livros, os dados, quando acessíveis ao público, são uma ferramenta de geração de conhecimento de forma mais rápida e com um custo menor. Ter esse entendimento de que um dado recolhido para uma pesquisa específica pode ser alvo de muitos estudos e que isso

gera uma economia de tempo e gastos, em alguns casos inclusive viabilizando pesquisas que de outra forma seriam impossíveis de se realizar, também desperta na comunidade científica a necessidade de apoiar o compartilhamento de dados. Sayão e Sales (2015, p. 21) acreditam que “o acesso e compartilhamento de dados de pesquisa contribui de forma significativa para que a ciência avance e maximize os investimentos aplicados em programas de pesquisa”.

Além disso, defender a importância que o compartilhamento de dados gera para a ciência é mostrar um universo de vantagens tanto para o próprio cientista que faz a coleta dos dados originais como para a futura geração de pesquisadores. Para Piwowar, Day e Fridsma (2007), entre os benefícios que a partilha dos dados oferece para a comunidade científica estão o incentivo a múltiplas perspectivas sobre determinado assunto, a ajuda na identificação de erros, o fato de desencorajar fraudes, o aumento do uso eficiente de recursos financeiros, além de evitar a coleta de dados duplicados. Whitlock (2011) acredita que os dados, como parte importante da carreira de um pesquisador, podem gerar novas descobertas científicas quando arquivados, o que também aumenta as chances dos cientistas que recolheram esses dados receberem os créditos.

Apesar da evidente necessidade de compartilhar os dados, algumas preocupações ainda cercam os cientistas, o que pode desmotivá-los a disponibilizar seus dados em acesso aberto. Piwowar e Chapman (2008) citam o tempo que pode levar para partilhar os dados por ser necessário formatar, documentar e enviar; a possibilidade de as conclusões originais passar por uma reanálise; o receio dos investigadores por excessivos pedidos de assistência; a diminuição da vantagem competitiva para futuras publicações; e a complicação para depositar os dados que podem se tornar desorganizados ou se perder se não forem bem administrados. Bruna (2010) expõe outras quatro preocupações: o receio de conceder os dados que ainda estão sendo usados e analisados para outros estudos; a apreensão dos pesquisadores de não serem citados; casos em que os dados podem trazer problemas políticos, culturais ou biológicos quando acessíveis ao público; e o arquivamento de dados que exige uma estrutura que garanta a integridade e acesso aos dados por um longo período, além da necessidade de possuir uma entrada flexível para os diversos tipos de dados.

Diante desse cenário, não só os pesquisadores se posicionaram, mas de uma forma geral todos os agentes da comunicação científica tiveram que se manifestar e tomar partido em relação a esse assunto, reconhecendo os dados como fontes de informação. Lehnert e Hsu (2015) garantem que os governos, financiadores, instituições acadêmicas, organizações profissionais e editores começaram a emitir políticas e declarações que apoiam o acesso aberto aos dados. Esse posicionamento e geração de políticas são importantes para nortear os pesquisadores e garantir a segurança e o destino que os dados devem tomar para serem compartilhados. É por essa razão que muitas agências financiadoras e revistas científicas



começaram a exigir a partilha e depósito dos dados primários como requisito de financiamento e publicação das pesquisas (Piwowar e Chapman, 2008 ; Sayão e Sales, 2015).

Entende-se, porém, que apenas a exigência de ter que partilhar os dados não garante que estes permanecerão úteis para a ciência. É necessário que os dados estejam acessíveis e possam ser utilizados independentemente do formato em que foram produzidos. Por isso que o arquivamento de dados e o local para que isso ocorra precisam ser tratados de forma tão cuidadosa dentro das políticas. Klump *et al* (2006) e Costello (2009) acreditam que os centros de dados modernos podem ajudar os dados de pesquisas a atender os critérios de persistência e qualidade para serem considerados para publicações, garantindo metadados adequados para descrição dos dados, agrupando dados de diferentes fontes e ainda fornecendo ferramentas para explorar, visualizar e baixar os dados em formatos apropriados.

Esses centros podem ser tanto repositórios como as próprias revistas científicas que auxiliam na publicação formal dos dados. Eles colaboram para que os dados não se percam sendo depositados em locais que não oferecem os subsídios necessários para a preservação e manutenção da sua qualidade, com por exemplo nos sites pessoais.

Sayão e Sales (2014) e Lehnert e Hsu (2015) garantem que os repositórios de dados se incorporam rapidamente a essa estrutura de informação, permitindo que o acervo de dados seja adequadamente preservado e permaneça acessível para uso, reutilização e compartilhamento, agregando ainda mais valor aos dados a medida que eles vão sendo descobertos, explorados e citados. Além disso, as revistas também estão se apresentando como uma interligação entre os dados e seus resultados finais. Bruna (2010) acredita inclusive que as revistas, ao apoiar e facilitar o depósito de dados, ajudam os pesquisadores a cumprir com os requisitos para comunicação exigidos pelas agências de financiamento, além das obrigações éticas com os cidadãos que indiretamente financiam as suas pesquisas. Portanto, as revistas científicas aumentam a sua relevância quando se posicionam em favor da disponibilização dos dados para garantir a difusão completa das pesquisas científicas.

### **O papel das revistas no compartilhamento de dados**

As revistas carregam a função de serem os principais meios de divulgação da ciência e, por isso, sofrem diretamente as consequências das mudanças informacionais. A cada nova demanda, as revistas precisam se adaptar para corresponder às expectativas da comunidade científica.

Quando surgiu a questão do compartilhamento de dados por parte dos pesquisadores, os editores perceberam a necessidade de incluir de alguma forma a discussão dentro das revistas, seja por meio das próprias publicações ou orientando os autores a depositarem seus dados em locais apropriados. Sturges *et al* (2015) e Whitlock (2011) consideram que os editores e as políticas das revistas têm um papel fundamental na

adoção do compartilhamento de dados no sentido de apoiar o arquivamento e dar o devido crédito aos cientistas quando ocorre o reuso dos dados, preocupando-se também com a vinculação dos dados depositados nos repositórios aos documentos originais. Uma alternativa para fazer essa interligação entre dados e artigos publicados é a solicitação de um número de registro nos repositórios e/ou o emprego de um identificador persistente para os dados.

Parr (2007) sugere ainda que cada revista apresente uma política com informações sobre onde, como e quando os dados devem ser depositados e associados as publicações. Esse posicionamento auxilia a quebra das barreiras impostas pelos pesquisadores para barrar a partilha dos dados coletados e garante o desenvolvimento da ciência por vias mais amplas envolvendo todos os produtos gerados nas pesquisas.

Existem ainda as revistas que não atuam apenas exigindo o compartilhamento de dados em repositórios, mas exercem também a função de preservação e curadoria. Isso ocorre quando elas permitem que dados sejam publicados como arquivos suplementares ou quando são consideradas revistas de dados.

Porém Whitlock (2011) considera que optar por arquivar os dados como materiais suplementares de artigos ainda não é a opção ideal, apesar de ser uma alternativa para permitir o acesso a eles e que é melhor do que mantê-los guardados. O autor acredita que dessa forma os dados estão mais propensos a se deteriorar com o tempo por não receberem todos os cuidados necessários para manutenção, o que se espera que ocorra quando são depositados em repositórios preparados para isso.

Já as chamadas revistas de dados assumem um papel ainda maior como publicação. Segundo Force *et al* (2016), esse tipo de revista surge como uma ponte entre as publicações tradicionais e a submissão de dados. Elas não trabalham apenas com a citação e resultados tirados dos dados, mas buscam apresentar de forma completa toda a estrutura e circunstância que envolve a coleta, análise e produção deles. Callaghan *et al* (2013) afirmam que uma revista de dados viabiliza ao leitor conhecer e entender quando, como e por que os dados foram coletados através de uma descrição detalhada do conjunto. Mas o autor não descarta a necessidade dessas revistas manterem vínculo com repositórios que armazenam o conjunto de dados para conservar e proteger os registros científicos. Chavan e Penev (2011) apontam também que um dos objetivos das revistas de dados é despertar a atenção da comunidade científica quanto a existência dos dados.

Ciente desse importante papel que as revistas assumem quando apoiam o compartilhamento de dados, este estudo visa analisar a medida e a forma como as revistas em acesso aberto do Brasil e de Portugal têm tratado a questão do depósito e publicação dos dados.

## Metodologia

Para alcançar o objetivo proposto, foi adotada uma abordagem mista, ou seja, tanto quantitativa quanto qualitativa, buscando-se estudar o caso das revistas brasileiras e portuguesas das áreas de Ciências e Medicina quanto à exigência de partilha de dados. A escolha dessas áreas foi feita pelo histórico de produção de dados e a grande presença delas na literatura sobre compartilhamento de dados científicos.

A amostra da pesquisa foi formada pelas revistas das áreas selecionadas que estão indexadas no *Directory of Open Access Journals* (DOAJ), principal diretório de revistas de acesso aberto da atualidade. Esse diretório tem por objetivo aumentar a visibilidade e facilitar o uso de revistas científicas e acadêmicas de acesso aberto. Atualmente o DOAJ possui mais de 9 mil revistas indexadas e o Brasil é o país com o maior número de revistas no diretório com 874 publicações cadastradas. Portugal tem ao todo 71 revistas indexadas nesse sistema.

A pesquisa no DOAJ para seleção dos títulos foi realizada utilizando-se os filtros de tipos de documentos recuperados [*Journals vs Articles*], país de publicação [*country of publisher*] (Brasil e Portugal) e assunto [*Subject*] (Ciências e Medicina). Nos casos das publicações recuperadas nas duas áreas de corte, ciências e medicina, efetuou-se uma análise do seu conteúdo interno para identificar a temática mais abordada por elas e, para efeito de descrição de dados, foram encaixadas em apenas um grupo.

E com o intuito de verificar a exigência de depósito em bases de dados ou a publicação na própria revista dos dados figurados nos artigos a elas submetidos, foi feita a análise documental das informações dispostas nas páginas das revistas recuperadas na busca e, mais detalhadamente, das suas diretrizes para autores.

## Resultados

A busca realizada em abril do ano de 2016 no DOAJ com os filtros propostos recuperou um total de 231 revistas brasileiras e portuguesas a serem analisadas no perfil definido para a pesquisa. Desse total, o sistema retomou 216 revistas do Brasil, sendo 77 da área de ciências e 139 da medicina, e 15 revistas de Portugal, que engloba cinco de ciências e 10 de medicina. Com a análise detalhada das páginas das revistas e de suas políticas de submissão, foi identificado que 90 dessas revistas exigem algum tipo de depósito de dados dos autores que submetem suas pesquisas, sendo das brasileiras 15 de ciências e 71 de medicina, e das portuguesas uma de ciências e três de medicina (Figura 1).

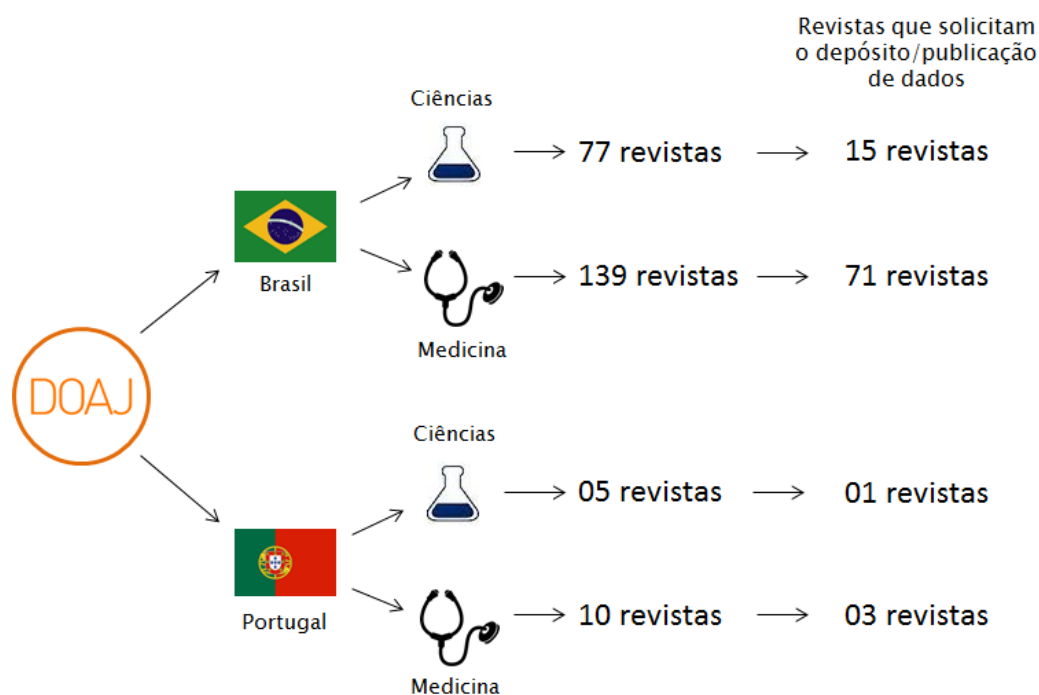


Figura 1: Número de revistas que solicitam o depósito/publicação de dados

O tratamento das revistas em relação ao compartilhamento de dados envolve a especificação dos dados a serem compartilhados. Os tipos de dados que as publicações exigem dependem da área específica de cada uma e, por isso, é possível identificar uma diversidade de tipos de dados nas políticas das revistas. Esses dados podem ser dados brutos, sequência de genomas, dados de cristalografia, dados taxonômicos, tipos de espécies, sequências de nucleotídeos, compostos sintetizados, registro de genes, ensaios randomizados e ensaios clínicos. Sendo que é predominante nas revistas da área médica a exigência do registro dos ensaios clínicos, o que ocorreu em 64 das publicações do Brasil e em três de Portugal, mas uma brasileira da área de ciências também apresenta essa solicitação.

O depósito dos ensaios clínicos é motivado em 55 revistas em apoio às políticas para registro desses materiais da Organização Mundial da Saúde (OMS) e do Comitê Internacional de Editores de Revistas Médicas (ICMJE), reconhecendo a importância do registro e divulgação internacional das informações dos estudos clínicos em acesso aberto. Nestes casos, as revistas orientam que os ensaios clínicos sejam depositados em Registros de Ensaio Clínicos validados pelos critérios estabelecidos pela OMS e ICMJE. No site do ICMJE (<http://www.icmje.org/>) está disponível uma lista com os registros de ensaios aceitos pelo comitê e 23 das revistas brasileiras de medicina optaram por indicar essa lista para os autores.

Observa-se uma variedade tipológica também para tratar dos locais onde os dados devem ser depositados. É possível encontrar nas políticas das revistas além de banco de

dados, centro de dados, repositórios, museus públicos, instituições legais, bancos públicos e registros públicos. Além disso cerca de sete revistas ainda aceitam o material como arquivo suplementar ao texto principal.

Como forma de auxílio aos pesquisadores, as publicações indicam alguns bancos de dados como sugestão para o depósito. Os bancos mais comuns nas políticas das revistas são:

- GenBank do National Center for Biotechnology Information – NCBI (<https://www.ncbi.nlm.nih.gov/genbank/>)
- ClinicalTrials.gov (<https://clinicaltrials.gov/>)
- Registro Brasileiro de Ensaio Clínicos – ReBEC (<http://www.ensaiosclinicos.gov.br/>)
- European Molecular Biology Laboratory – EMBL (<http://www.embl.org/>)

E quanto às publicações que atuam diretamente com os dados e se posicionam como revistas de dados, foi localizada nesse grupo analisado apenas uma revista nessa situação. A revista é brasileira da área de ciências e pode ser considerada uma revista especificamente de dados, ou seja, o conteúdo publicado são os próprios dados científicos.

## Conclusão

A discussão em entorno do depósito dos dados tem atingido de forma significativa a comunidade científica. Busca-se atualmente um detalhamento maior do conteúdo disponível nas pesquisas já publicadas através da análise dos dados coletados durante todo o processo científico.

Porém percebe-se que muitos são os obstáculos impostos para dificultar a partilha de dados, mas está cada vez mais claro que o compartilhamento dos dados é hoje uma questão primordial no universo da ciência. E como uma das principais interessadas nesse assunto, as revistas começaram a se posicionar como um agente motivador desse compartilhamento. O incentivo por parte das revistas de uma publicação formal dos dados em repositórios ou até mesmo nelas é uma maneira de fornecer aos dados e aos seus pesquisadores uma garantia de acesso e preservação.

Por meio do levantamento feito na base do DOAJ nas áreas de ciências e medicina, foi possível verificar que, mesmo que de forma ainda discreta, as revistas brasileiras e portuguesas de acesso aberto têm se mobilizado para que os dados estejam à disposição da comunidade de forma completa e acessível. Os editores, principalmente os da área de medicina, têm se apoiado em decisões de organizações internacionais para poderem exigir de seus pesquisadores o depósito dos dados utilizados.

A opção de depositar os dados em repositórios próprios para esse fim ainda é a melhor opção e a mais recomendada nas políticas das revistas. Alguns bancos de dados têm se destacado como opções para receber os dados de pesquisa e são na sua maioria

repositórios internacionais. Quando sugeridos pelas revistas, geralmente são bancos de dados específicos das áreas cobertas por essas publicações.

Com isso, é possível concluir que a maioria das publicações que exigem ou sugerem que os dados sejam disponibilizados reconhecem a importância do acesso ao conteúdo completo das pesquisas. Além de compreender que os dados estão ganhando uma visibilidade que antes era apenas das tabelas e gráficos presentes nos artigos e trabalhados para o simples entendimento do resultado final de uma pesquisa. Isso ocorre devido a essa busca pelo compartilhamento dos dados de pesquisa e o movimento pelo acesso aberto a produção científica que tem expandido a sua atuação.

### Referências bibliográficas

BERLIN DECLARATION (2003) – Declaração de Berlim sobre acesso livre ao conhecimento nas ciências e humanidades. [Em linha]. Berlim. [Consult. 22 Set. 2016]. Disponível na Internet: <URL: [https://openaccess.mpg.de/67693/BerlinDeclaration\\_pt.pdf](https://openaccess.mpg.de/67693/BerlinDeclaration_pt.pdf)>.

BORGMAN, Christine L. (2010) – Research data: who will share what, with whom, when, and why? In *China-North America Library Conference*. Pequim, China: 2010. [Consult. 19 Set. 2016]. Disponível na Internet: <URL: <http://works.bepress.com/borgman/238/>>.

BRUNA, Emilio M. (2010) – Scientific journals can advance tropical biology and conservation by requiring data archiving. *Biotropica* [Em linha]. Vol. 42, Nº 4. [Consult. 15 Abr. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.1111/j.1744-7429.2010.00652.x>>. ISSN 1744-7429.

BUTLIN, R. (2011) – Data archiving. *Heredity* [Em linha]. Vol. 106, Nº 709. [Consult. 15 Set. 2016]. Disponível na Internet: <URL: <http://www.nature.com/hdy/journal/v106/n5/full/hdy201043a.html>>. ISSN 1365-2540.

CALLAGHAN, Sarah [et al.] (2013) – Processes and procedures for data publication: a case study in the geosciences. *International Journal of Digital Curation* [Em linha]. Vol. 8, Nº 1. [Consult. 15 Set. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.2218/ijdc.v8i1.253>>. ISSN 1746-8256.

CHAVAN, Vishwas; PENEV, Lyubomir (2011) – The data paper: a mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics* [Em linha]. Vol. 12 (Suppl 15), Nº S2. [Consult. 20 Set. 2016]. Disponível na Internet: <URL: <http://www.biomedcentral.com/1471-2105/12/S15/S2>>. ISSN 1471-2105.

COSTELLO, Mark J. (2009) – Motivating online publication of data. *BioScience* [Em linha]. Vol. 59, Nº 5. [Consult. 18 Set. 2016]. Disponível na Internet: <URL: <http://bioscience.oxfordjournals.org/content/59/5/418.full>>. ISSN 1525-3244.

FORCE, Megan [et al.] (2016) – Research Data in Journals and Repositories in the Web of Science: Developments and Recommendations. *Bulletin of IEEE Technical Committee on Digital Libraries* [Em linha]. Vol. 12, N° 1. [Consult. 15 Set. 2016]. Disponível na Internet: <URL: [http://www.ieee-tcdl.org/Bulletin/current/papers/IEEE-TCDL-DC-2016\\_paper\\_3.pdf](http://www.ieee-tcdl.org/Bulletin/current/papers/IEEE-TCDL-DC-2016_paper_3.pdf)>. ISSN 1937-7266.

HEIDORN, P. Bryan (2008) – Shedding light on the dark data in the long tail of science. *Library Trends* [Em linha]. Vol. 57, N° 2. [Consult. 16 Abr. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.1353/lib.0.0036>>. ISSN 1559-0682.

KLUMP, Jens [et al.] (2006) – Data publication in the open access initiative. *Data Science Journal* [Em linha]. Vol. 5, N° 15. [Consult. 17 Set. 2016]. Disponível na internet: <URL: <http://doi.org/10.2481/dsj.5.79>>. ISSN 1683-1470.

LEHNERT, Kerstin; HSU, Leslie (2015) – The new paradigm of data publication. *Elements* [Em linha]. Vol. 11, N° 5. [Consult. 25 Set. 2016]. Disponível na Internet: <URL: [http://elementsmagazine.org/archives/e11\\_5/e11\\_5\\_dep\\_toolkit.pdf](http://elementsmagazine.org/archives/e11_5/e11_5_dep_toolkit.pdf)>. ISSN 1811-5209.

MARQUES, Fabrício (2014) – Ciência transparente. *Pesquisa FAPESP* [Em linha]. N° 218. [Consult. 18 Set. 2016]. Disponível na Internet: <URL: <http://revistapesquisa.fapesp.br/2014/04/24/ciencia-transparente/>>.

MOLLOY, Jennifer C. (2011) – The open knowledge foundation: open data means better science. *PLoS Biology* [Em linha]. Vol. 9, N° 12. [Consult. 25 Set. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.1371/journal.pbio.1001195>>.

OECD (2007) – *OECD Principles and Guidelines for Access to Research Data from Public Funding*. [Em linha] Paris: Organisation for Economic Co-Operation and Development. [Consult. 12 Set. 2016]. Disponível na Internet: <URL: <https://www.oecd.org/sti/sci-tech/38500813.pdf>>.

PARR, Cynthia S. (2007) – Open Sourcing Ecological Data. *BioScience* [Em linha]. Vol. 57, N° 4. [Consult. 15 Set. 2016]. Disponível na Internet: <URL: <http://bioscience.oxfordjournals.org/content/57/4/309.full>>. ISSN 1525-3244.

PIWOWAR, Heather A.; CHAPMAN, Wendy W. (2008) – A review of journal policies for sharing research data. In *ELPUB 2008 Conference on Electronic Publishing*. Toronto, Canada: 2008. [Consult. 19 Set. 2016]. Disponível na Internet: <URL: [http://elpub.scix.net/data/works/att/001\\_elpub2008.content.pdf](http://elpub.scix.net/data/works/att/001_elpub2008.content.pdf)>.

PIWOWAR, Heather A.; DAY, Roger S.; FRIDSMA, Douglas B. (2007) – Sharing detailed research data is associated with increased citation rate. *PLoS ONE* [Em linha]. Vol. 2, N° 3. [Consult. 15 Abr. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.1371/journal.pone.0000308>>. ISSN 1932-6203.

SAYÃO, Luís F.; SALES, Luana F. (2014) – Dados abertos de pesquisa: ampliando o conceito de acesso livre. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde* [Em linha]. Vol. 8, Nº 2. [Consult. 15 Set. 2016]. Disponível na Internet: <URL: <http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/611>>. ISSN 1981-6278.

SAYÃO, Luís F.; SALES, Luana F. (2015) – Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores [Em linha]. Rio de Janeiro : CNEN/IEN. [Consult. 20 Ago. 2016]. Disponível na Internet: <URL: [http://www.cnen.gov.br/images/CIN/PDFs/GUIA\\_DE\\_DADOS\\_DE\\_PESQUISA.pdf](http://www.cnen.gov.br/images/CIN/PDFs/GUIA_DE_DADOS_DE_PESQUISA.pdf)>. ISBN 978-85-61905-03-3.

STURGES, Paul [et al] (2014) – Access to research data: addressing the problem through journal data sharing policies. In *IATUL Conference Proceedings*. [Consult. 16 Set. 2016]. Disponível na Internet: <URL: <http://docs.lib.purdue.edu/iatul/2014/openaccess/3/>>.

STURGES, Paul [et al] (2015) – Research data sharing: developing a stakeholder-driven model for journal policies. *Journal of the Association for Information Science and Technology* [Em linha]. Vol. 66, Nº 12. [Consult. 12 Set. 2016]. Disponível na Internet: <URL: <http://onlinelibrary.wiley.com/doi/10.1002/asi.23336/abstract>>. ISSN 2330-1643.

UHLIR, Paul F.; SCHRÖDER, Peter (2007) – Open data for global Science. *Data Science Journal* [Em linha]. Vol. 6 [Consult. 16 Set. 2016]. Disponível na Internet: <URL: <http://doi.org/10.2481/dsj.6.OD36>>. ISSN 1683-1470.

WHITLOCK, Michael C. (2011) – Data archiving in ecology and evolution: best practices. *Trends in Ecology and Evolution* [Em linha]. Vol. 26, Nº 2. [Consult. 18 Set. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.1016/j.tree.2010.11.006>>. ISSN 0169-5347.



---

## Ritmos da informação/comunicação de ciência dos centros de investigação em Portugal

---

Fábio Ribeiro

Centro de Estudos de Comunicação e Sociedade – Universidade do Minho

[fabio.ribeiro@ics.uminho.pt](mailto:fabio.ribeiro@ics.uminho.pt)

### Resumo

A ciência envolve hoje um conjunto de procedimentos que não se esgotam no manuseamento de métodos e técnicas, materiais e enquadramentos teóricos. A preocupação de muitos cientistas reside na visibilidade científica do seu trabalho, pela valorização pessoal ou progressão na carreira, entre outros motivos. De braço dado com uma “novilíngua” muito própria (Martins, 2012), a ciência passou a estar na encruzilhada de *rankings*, visualizações, *downloads*, citações, bases de indexação de revistas, projetos e bolsas, ou da internacionalização.

Na expectativa de contribuir para uma reflexão sobre a disseminação do trabalho científico, este texto tenta responder a dois desafios: 1) recuperar algum do debate conceptual sobre o que significa informar/comunicar sobre ciência (Wolton, 2009; Carvalho & Cabecinhas, 2004), propondo a complementaridade entre os dois conceitos, em que o primeiro trata a publicação de conteúdos científicos e o segundo sugere uma vertente relacional entre investigadores e públicos (não) especializados; 2) analisar a informação e as ações de comunicação de ciência de centros de investigação em Portugal. Pretende-se caracterizar o posicionamento destes centros em ambientes online (*site*, blogues, redes sociais, etc.), avaliando questões críticas como modos de atuação nestes contextos, conteúdos habitualmente publicados, ações de formação com grupos específicos, relação com os média, interação com repositórios institucionais.

**Palavras-chave:** comunicação; ciência; centros de investigação; FCT; Portugal.

Science Information/Communication Patterns in research unities in Portugal

### **Abstract**

Handling with science nowadays is no longer limited to methods, techniques, materials and theoretical frameworks. Scientists seem to be pretty aware that their scientific outcomes must be available for the general public and entities, not only for personal motivations, but also to progress in professional terms. Hence, rankings, articles' views and downloads, citations and indexations, grants and funded projects are brought into a "new shaped scientific language" (Martins, 2012).

Following the premise to contribute for the debate about the visibility of science, this research seeks to answer two questions: 1) what is exactly "inform" and "communicate" about science, from a theoretical point of view? In this sense, it is proposed a new paradigm of "information science", dealing only with content publication, and "communication science", as an interactive strategy to enhance the proximity among researchers and the general public (Wolton, 2009; Carvalho & Cabecinhas, 2004); 2) observing these two concepts, what kind of content is published by research centres in Portugal, in websites and social media, in terms of scientific outputs and events, clipping from the media and interaction with institutional repositories.

**Keywords:** communication; science; research centres; FCT; Portugal.

### **Visibilidade do trabalho científico – uma breve discussão conceptual**

O contacto com a informação científica decorre, sobretudo, em ambiente escolar, numa primeira fase do percurso educativo. Consoante a opção estratégica assumida individualmente, pode dar-se o caso de um afastamento circunstancial entre o cidadão e a informação de cariz científico, que se esgota nos bancos da escola ou na universidade. Nestas circunstâncias, os média têm a capacidade de preencher este eventual desencontro, promovendo um papel ativo na publicação de conteúdos científicos. Neste quadro inerente à visibilidade do trabalho do cientista, e a sua longevidade no tempo, existe um debate francamente amplo sobre as condições pelas quais se estabelece a comunicação da ciência, tendo em conta a construção de um edifício comunicativo de uma pós-modernidade técnica e tecnológica que atribui novas configurações ao significado social de "ser cientista".

Ainda que o trabalho do cientista seja valorizado positivamente em termos sociais, existem porventura fatores endógenos que obstam igualmente a uma maior expressão comunicativa por parte dos atores científicos, dada a pressão para publicar, a preocupação permanente com rankings científicos, a progressão na carreira ou mesmo a estabilidade do próprio mercado de trabalho nesta área, refém de diferentes perceções e investimentos díspares entre governos (Martins, 2012). O reconhecimento difuso do público, tido como a níveis residuais pelas matérias científicas, concretiza-se num dos principais dados revelados pelo Pew Research Center, dos EUA, publicado em fevereiro de 2015. Segundo o documento, 84% de cientistas considerava que o desinteresse dos cidadãos pelas circunstâncias do

trabalho científico se assume numa importante desmotivação neste contexto. Deste modo, podemos especular que este facto deverá condicionar também a predisposição dos cientistas em estabelecer ações de contacto com públicos não especializados, por exemplo.

Não obstante fatores internos, que são discutíveis pela transversalidade que evocam, a emancipação da comunicação de ciência sofreu com as várias representações sociais desenvolvidas sobre a ciência, ao longo das décadas. Na verdade, é com uma imagem algo debilitada que a ciência – tida aqui de uma ponta de vista genérico – sai da II Guerra Mundial, devido a ligação estratégica que alguns regimes totalitários manifestaram junto de determinados cientistas, com os conhecidos episódios de ligação da engenharia ao fabrico das bombas atómicas ou mesmo ao estabelecimento de campos de concertação nazi (Felt et al., 1995). Este cinzentismo a que a ciência havia sido vergada provocou-lhe uma herança histórica difícil de suplantar. Assim, numa tentativa de contrariar este ambiente pouco pessimista, determinados setores da comunidade científica estabeleceram laços oportunamente construídos para reforçar e solidificar a credibilidade social da ciência, demonstrando os efeitos práticos e benéficos do conhecimento científico, em ambiente escolar, através de autênticas manobras de relações públicas junto dos média (Greenberg, 2001). O avanço tecnológico, sobretudo na área das telecomunicações, com implicações diretas e de fácil acesso ao cidadão comum, foram transfiguradas progressivamente até finais dos anos 90, dotando a ciência de uma imagem já mais positiva, comparativamente àquela que manifestamente exibia em meados do século XX (Kurath & Gisler, 2009). Esta dificuldade na emancipação da comunicação da ciência pode até ler-se na criação relativamente tardia da revista científica *Public Understanding of Science*, da SAGE, que apenas surgiu em 1992, ao contrário de diversos títulos desta editora, que surgiu em 1965.

O que pode estar em causa neste quadro de emergência dos desafios da credibilização e visibilidade da ciência, nas exigências de entidades financiadores e cobranças a nível público e mediático, será provavelmente a passagem de um “modelo de défice cognitivo” para um “modelo interativo” (Carvalho & Cabecinhas, 2004), em que já se assume as limitações de um conhecimento parco de públicos não especializados sobre ciência para a adoção de comportamentos mais interativos com este grupo de pessoas, que se concretizam em ações de formação ou demonstração de ciência.

A questão fundamental neste contexto reside, portanto, no impacto da ciência na sociedade em geral, a sua relação com os públicos, e a própria motivação dos cientistas para comunicar sobre o seu trabalho no espaço público. Neste sentido, e devido ao seu percurso como “comunicador de ciência” e jornalista nos EUA, Hillier Krieghbaum acompanhou com particular interesse a forma como os leitores de jornais consumiam e se relacionavam com as notícias de teor científico. Como recorda o autor (1970), a National Association of Science Writers (NASW) levou a cabo um estudo, em 1967, para identificar a pertinência que os leitores atribuíam aos conteúdos científicos. Em quase duas mil pessoas inquiridas, um

grande número – 83% – reconheceu a necessidade de se conhecer melhor a investigação científica para que a sociedade encontre “um mundo melhor”. Analisando do ponto de vista mediático, existem até poucos “comunicadores de ciência” de excelência, com carisma, visibilidade e notoriedade. À exceção de poucos casos de investigadores – que se tornaram emblemáticos e mediáticos, como Carl Sagan e Stephen Hawking – existe porventura uma sensação de afastamento deliberado dos cientistas com públicos não especializados.

Ainda que esta percepção pareça nítida, não será tão clara a definição conceptual sobre “comunicar ciência”. Particularmente difusa no debate académico, a vulgarização destes dois termos, associados a universos relativamente conhecidos do espaço público e da quotidianidade, prestam-se fundamentalmente à formulação de diferentes entendimentos sobre aquilo que evoca. A incerteza sobre as dimensões integrantes do conceito consolida-se em determinados entendimentos:

“Frequentemente associada à divulgação pública do conhecimento científico, a comunicação da ciência tem várias outras facetas, da comunicação entre cientistas à ficção científica no cinema. O próprio conceito de ciência, na expressão ‘comunicação da ciência’, pode ser problematizado: falamos de investigação científica, apenas, ou também das aplicações da pesquisa, alargando então o conceito à tecnologia?” (Carvalho & Cabecinhas, 2004: 5).

O conceito de “comunicação de ciência”, provavelmente o mais fácil para tentar ilustrar esta realidade, aparece na obra de Burns et al. (2003) sob a sigla “Scicom”. Os autores estabelecem assim uma “analogia da vogal” para definir este termo, na medida em que cada dimensão apresentada – iniciada por uma vogal – pretende demonstrar a essência do conceito, nos seguintes termos: *Awareness* (consciencialização), incluindo representações partilhadas com familiares e conhecidos sobre novos aspetos da ciência; *Enjoyment* (divertimento), como uma resposta afetiva fruto de uma valorização positiva da ciência como entretenimento ou arte; *Interest* (interesse), através do contacto voluntário com a ciência ou a forma como ela se dá a conhecer; *Opinions* (opinião), com a formação e reconstrução de entendimentos ligados ao conhecimento científico; *Understanding* (compreensão), que lida com a assimilação e explicação autónoma de conteúdos, processos e fatores inerentes à ciência. Embora a metáfora exprima alguma vertente ambiciosa neste contexto, outros autores identificam a especialização de contacto com os públicos como a força motriz de um renovado conceito de “comunicação estratégica de ciência”, tendo em conta a necessidade de delinear uma “estratégia que se baseie na descodificação das descobertas científicas para o público que não pertença a este meio” (Magalhães, 2015: 54).

Deste modo, e procurando contribuir para um debate que carece de uma explanação teórica ainda mais densa por parte dos estudiosos nesta área, esta investigação inspira-se na decisiva distinção entre “informação” e “comunicação”, proposta pelo sociólogo francês Dominique Wolton, em 2009, como conceitos antagónicos, mas complementares, nomeadamente com a indicação de que o primeiro se refere apenas ao conteúdo subjacente

ao ato comunicativo e o segundo à relação que se estabelece entre atores sociais, científicos. Deste modo, “informação científica” relaciona-se exclusivamente com todas as ações de publicação de conteúdos científicos, desprovida de uma interação entre investigadores e público/audiências específicas, como por exemplo, artigos em revistas científicas, livros, capítulos de livros, relatórios, etc. Por seu turno, a “comunicação científica” sugere uma vertente relacional, interativa, entre investigadores e a comunidade académica, civil, subjacente a uma troca de conhecimentos científicos, como são os casos dos congressos, seminários, das aulas abertas, demonstrações científicas (Noite Europeia dos Investigadores) ou uma reconfigurada interação via *site* e redes sociais.

Esta tipologia de conceitos presta-se essencialmente a uma conceção de disseminação da ciência tendo por base a sua natureza e essência, de conteúdos, e a possibilidade efetiva de ações concretas com públicos considerados numa perspetiva geral, embora percecionados como não especializados.

Tal categorização pretende sustentar o trabalho empírico que de seguida se apresenta, tendo como ambição contribuir de algum modo para uma radiografia de como os centros de investigação em Portugal promover a dinamização da visibilidade pública da ciência, a partir de uma determinada amostra.

### **Do debate conceptual às práticas no terreno: a visibilidade da ciência**

Sob o objetivo genérico de traduzir numa reflexão comparada o esforço e a atenção que determinadas dinâmicas se estabelecem junto da visibilidade da ciência e do trabalho realizado pelos investigadores, esta amostra procurou centrar-se em centros de investigação financiados pela Fundação para a Ciência e a Tecnologia (FCT), na última avaliação de unidades de I&D, em 2013.

Como resultado deste processo, 80% dos centros considerados obtiveram uma classificação entre “Bom” e “Excepcional”, como identifica a seguinte Figura 1.

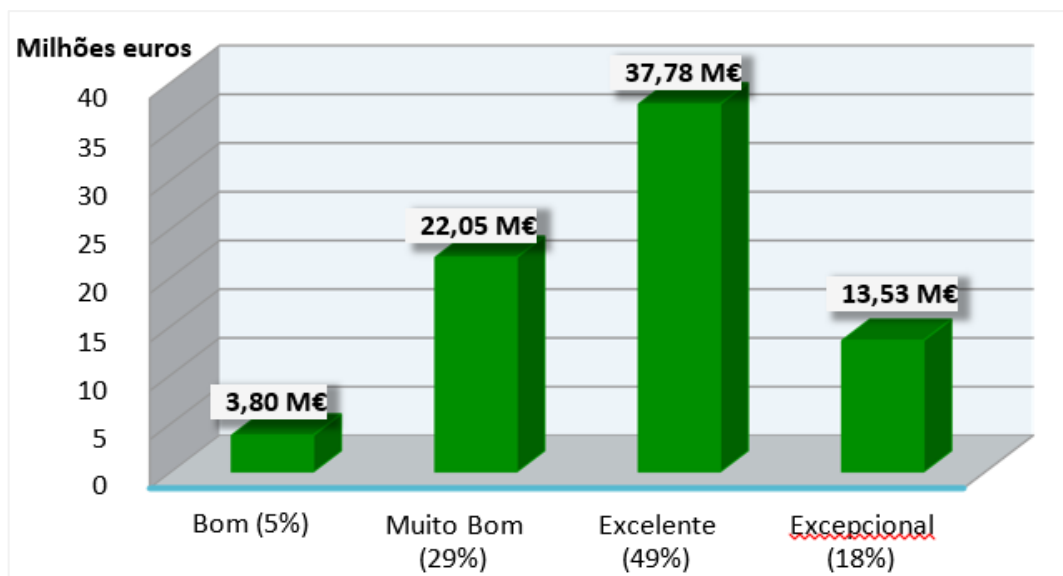


Figura 1: Avaliação de Unidades I&D 2013 – Resultados Globais. Documento de 25.05.2015. Fonte: FCT.

Estes centros agregavam 90% dos doutorados presentes a concurso, num montante global de investimento que superava os 77 milhões de euros, num horizonte temporal até 2020. A figura revela que foi na categoria “Excelente” que se concentrou a maior parte do financiamento das unidades de investigação.

Tendo em conta estes factos, seleccionou-se uma amostra de unidades de investigação com classificação “Excelente”, de dois painéis de avaliação (Ciências Sociais e Humanidades), num total de 25 centros de investigação<sup>i</sup> dispersos por todo o país, num período de observação datado de 20 de outubro de 2016. A Tabela 1 evidencia, portanto, os 24 centros de investigação analisados, orientadas a partir da perspetiva dos conceitos de “informação de ciência” e de “comunicação científica”, a partir de dois objetos de estudo: os *sítes* e as redes sociais.

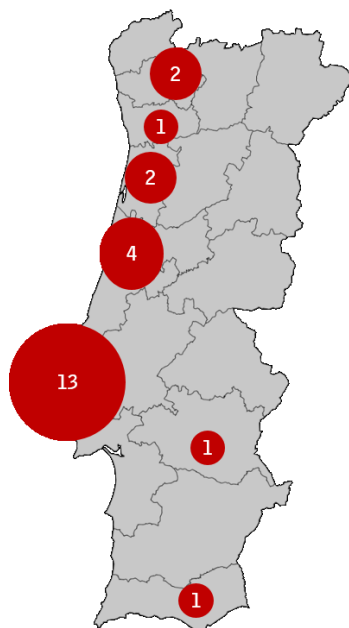
Nome	Acrónimo	Universidade	Número de doutorados
Centro de Estudos de Linguística Geral e Aplicada	CELGA	Universidade de Coimbra	35
Centro de Estudos de Sociologia e Estética Musical	CESEM	Universidade Nova de Lisboa	54
Centro de Estudos Geográficos	CEG	Universidade de Lisboa	67
Centro de História d'Aquém e d'Além-Mar	CHAM	Universidade Nova de Lisboa	113
Centro de Investigação de Políticas do Ensino Superior	CIPES	Universidade de Aveiro	20
Centro de Investigação do Núcleo de Estudos e Intervenção Cognitivo-Comportamental	CINEICC	Universidade de Coimbra	22
Centro de Investigação em Antropologia e Saúde	CIAS	Universidade de Coimbra	17
Centro de Investigação em Ciência Política	CICP	Universidade do Minho	37
Centro de Investigação em Psicologia	CIPsi g	Universidade do Minho (UM)	30
Centro de Literatura Portuguesa	CLP	Universidade de Coimbra	35
Centro de Psicologia da Universidade	CPUP	Universidade do Porto	43

do Porto			
Centro Interdisciplinar de Estudos de Género	CIEG	ISCSP - Universidade de Lisboa)	11
Centro Interdisciplinar de História, Culturas e Sociedades	CIDEHUS	Universidade de Évora	44
Instituto de Ciências Sociais	ICS-ULisboa	Universidade de Lisboa	87
Instituto de Etnomusicologia - Centro de Estudos em Música e Dança	INET-md	Universidade Nova de Lisboa	63
Instituto de Filosofia da Nova	IFILNOVA	Universidade Nova de Lisboa	43
Instituto de História Contemporânea	IHC	Universidade Nova de Lisboa	84
Interdisciplinary Center for Archaeology and Evolution of Human Behaviour	ICArEHB	Universidade do Algarve	12
Nova School of Business and Economics	Nova SBE	Universidade Nova de Lisboa	54
Research in Social Sciences and Management	CSG	Universidade Técnica de Lisboa	79
Research Unit in Business and Economics	CUBE	Universidade Católica Portuguesa	48
Unidade de investigação em Governança, Competitividade e Políticas Públicas	GOVCOPP	Universidade de Aveiro	44
Vidro e Cerâmica para Artes	VICARTE	Universidade Nova de Lisboa	12
William James Center for Research	WJCR	Instituto Universitário de Ciências Psicológicas, Sociais e da Vida	11

**Tabela 1:** Amostra dos centros de investigação estudados.

A análise que se procurará desenvolver não pretende discutir com particular insistência as dinâmicas próprias de cada centro, pelo que a lógica de observação generalista prevalece, eventualmente apenas contrariada para pontuar um ou outro caso que, pela sua natureza, se apresenta como inovador ou como símbolo de uma prática pouco conseguida.

Em relação ao número de doutorados, estes painéis concentram 1065 membros, num investimento de mais de 6,5 milhões de euros (6,631,276 €), sensivelmente 10% do investimento total em ciência no país. Formalmente, os 24 centros de investigação dispersam-se geograficamente no litoral (Mapa 1), com o distrito de Lisboa a reunir a maioria das entidades observadas (13), seguido de Coimbra (quatro), Aveiro e Braga (dois), com Évora e Faro num plano residual, com um centro de investigação “Excelente” cada. Note-se ainda a ausência das Regiões Autónomas neste contexto.



Mapa 1: Distribuição geográfica dos centros de investigação.

Após esta caracterização formal, estudaram-se as dinâmicas associadas aos *sites* e rede sociais destes centros. No que diz respeito à primeira dimensão (*site*), as seguintes categorias da observação procuram descrever a atividade das unidades de investigação:

– **Conteúdos comuns:** nesta categoria localizámos os conteúdos que surgiram com maior frequência: “quem somos” (identificação da estrutura do centro, direção e gestão); “investigação” (inclui indicação dos investigadores e linhas/grupos de pesquisa); “formação” (ações, cursos universitários ou seminários associados ao centro); “projetos de investigação” (em curso ou descontinuados, identificando a fonte de financiamento, se aplicável); “eventos” (série de acontecimentos que marcam a vida de um centro, entre congressos, anúncios para a chamada de comunicações, destaques dos investigadores nos média, etc.). Este conjunto de informações configura uma ideia de “comunicação institucional” das próprias unidades;

– **Enquadramento:** a maioria dos centros (17 em 24) opta por integrar o *site* numa instituição ou escola universitária a que pertence. Poucos preferem criar um endereço próprio – apenas 6 – e um deles funciona a partir de um blogue;

– **Idiomas:** no que se refere ao idioma de trabalho, 17 centros publicam os seus conteúdos em português e inglês, numa lógica bilingue. Cinco utilizam apenas o português e dois o inglês;

– **Disponibilização de conteúdos:** naquela que será provavelmente a dimensão mais subjetiva desta análise, pretendeu-se avaliar genericamente a quantidade de informações que os centros colocam à disposição do público, em termos dos menus apresentados e dos conteúdos que deles constam. Assim, 11 apresentam *sites* “muito completos”, cinco num



nível “completo”, quatro “razoável”, três “insatisfatório” e apenas um centro com uma apresentação de conteúdos “mediocre”;

– **Identidade visual:** dos 24 centros de investigação, apenas três não têm qualquer referência visual, um logotipo próprio ou imagem distintiva daquela que é a sua instituição;

– **Aspetos pouco comuns:** seis centros de investigação produzem uma *newsletter* regular, com as principais indicações sobre os acontecimentos que marcaram a agenda e a atualidade das unidades e dos investigadores. A indicação de parceiras com a sociedade civil ou outras comunidades académicas aparece pouco explorada (apenas no Inet-Md), tal como na indicação dos Prémios e outras distinções (novamente o Inet-Md e o CESEM) ou as oportunidades de emprego (CELGA, IcarEHB). O IFIL Nova mantém aquilo que define como um “metablog”, uma oportunidade proporcionada por uma plataforma que convida à interação do público com os investigadores sobre várias temáticas do âmbito do centro. O IHC apresenta um espaço autónomo para o recorte de notícias nos média sobre a atividade dos investigadores, que se designa por “clipping”. O CEG inova ao disponibilizar uma Fototeca e Mapoteca, espaços virtuais em que, dada a natureza daquele centro, se proporciona uma perspetiva mais visual, como fotografias e mapas. O ICS e o Nova SBE centralizam a sua atividade numa lógica de grupo, em que a maior parte do conteúdo publicado versa essencialmente sobre a instituição-mãe a que pertencem.

Numa segundo ponto desta observação, entendeu-se considerar as redes sociais na tentativa de perceber possíveis articulações com os *sites* próprios das instituições. Com efeito, as unidades não relacionam os conteúdos do *site* a estas redes, optando por colocá-las numa dinâmica isolada. A Tabela 2 oferece, portante, uma visão global da presença dos centros neste âmbito:

Rede social	Número de centros de investigação
Sem qualquer página	13
Facebook	9
Facebook e Twitter	1
Facebook, Twitter e Youtube	1

**Tabela 2:** Utilização das redes sociais pelos centros de investigação.

A maioria dos centros de investigação não apresenta qualquer presença nas redes sociais (13 contra 11) e este acaba por ser o dado mais relevante nesta matéria. O Facebook, por seu turno, reúne o maior número de preferências nos centros que apostam na dinamização de conteúdos neste espaço (11). A opção por múltiplas presenças em outras

redes que não a criada por Mark Zuckerberg, em 2004, representa apenas a preocupação de dois centros, sendo que um utiliza adicionalmente o Twitter e outro opta por uma estratégia aparentemente mais ambiciosa com a integração em três redes sociais. Esta inclinação pelo Facebook, todavia, encontra-se em linha com as indicações de estudos recentes de entidades com a Marktest que, em sucessivos relatórios anuais intitulados “Os Portugueses e as Redes Sociais”, sublinham a utilização preferencial dos internautas nacionais por esta rede.

Outra forma de avaliar globalmente a prestação dos centros de investigação nas redes sociais prende-se com a atividade clássica que desenvolvem nestes espaços – a colocação de conteúdos (*posts*) – e ainda a manifestação de ‘gostos’/seguidores/subscritores gerados entre o público, que pode de algum modo ser indicador de envolvimento ou, pelo menos, de um acompanhamento mínimo daquilo que é publico nestes espaços.

No caso do Facebook, nos 11 centros que dinamizam uma página nesta plataforma, existem claramente ritmos diferentes<sup>i</sup>, entre os seguidores (*‘like’*) e o número de *posts* contabilizados. No entanto, parece seguro admitir que o número elevado de seguidores associados a uma página determina um volume importante de conteúdos publicados, pelo menos nos primeiros casos representados no Gráfico 1.

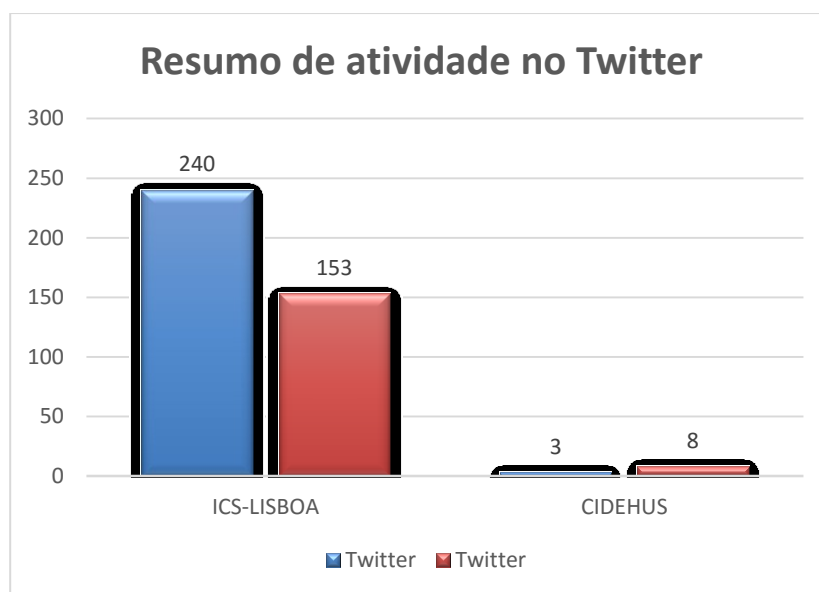


**Gráfico 1:** Número de seguidores e *posts* em 2016 no Facebook, nos centros de investigação com conta ativa nesta rede social.

Ainda assim, no caso do CIDEHUS que ocupa a terceira posição com mais *posts*, surge apenas em sexto lugar no número de seguidores. Tendo em conta que a média de seguidores oscila aos 1946 utilizadores, apenas quatro centros superam essa fasquia (IHC,

ICS, CHAM e CEG). A média não ultrapassa os 178, superado pela maioria das unidades analisadas.

O Gráfico 2 sistematiza a atividade de apenas dois centros de investigação no Twitter, com atuações díspares. O ICS-Lisboa apresenta um nível elevado de *posts* com um número de seguidores interessantes (240). O CIDEHUS mantém apenas uma presença praticamente residual.



**Gráfico 2:** Número de seguidores e *posts* em 2016 no Twitter, nos centros de investigação com conta ativa nesta rede social.

Como rede social que menos atividade desperta, o Youtube é utilizado apenas pelo ICS de Lisboa, com 13 subscritores e 20 vídeos, um número considerável tendo em conta as práticas generalizadas no quadro dos centros de investigação analisados (Gráfico 3).



**Gráfico 3:** Número de subscritores e *posts* em 2016 no Youtube, nos centros de investigação com conta ativa nesta rede social.

### Silêncios que preocupam: considerações finais

A partir da proposta conceptual aqui ensaiada neste texto, parece claro que a “informação de ciência” domina perante as ações de “comunicação de ciência”, na medida em que os centros de investigação apostam preferencialmente na divulgação e publicação de conteúdos meramente descritivos da atividade que produzem e onde o eco sobre iniciativas concretas de contacto com públicos específicos fica claramente fora da publicitação do que é visivelmente acessível nos *sites* e nas redes sociais. Neste sentido, a informação de ciência promovida pelos centros confunde-se, em larga medida, com uma ideia de “comunicação institucional”, dada a descrição manifestamente simplista de muitos conteúdos publicados, sobretudo pela repetição de opções em vários centros.

A simples publicação de conteúdos, em detrimento da socialização efetiva dos mesmos, recorda o entendimento de Ahmad Said Esber, poeta árabe, numa recente entrevista ao Expresso, quando refere que “houve uma mudança, a cultura é cada vez mais a cultura da informação. Já não há cultura, só informação”. Esta posição confirma provavelmente a secundarização da “comunicação de ciência”, a que evoca uma relação e um contacto direto no sentido de alterar o outro interlocutor.

Os *sites* das unidades de investigação reúnem informação competente e completa, na maioria dos casos. A informação é satisfatória para quem procura saber qual o âmbito de ação daquelas entidades. Contudo, as páginas poderiam ter outro investimento, mais apelativo do ponto de vista estético. Os visitantes não conseguem igualmente interagir em nenhum aspeto, o que revela um estaticismo das opções tomadas neste caso. Os investigadores são identificados de modo também simplista, na maior parte das situações,

sem uma referência de contacto ou descrição substantiva do percurso académico e/ou profissional, pelo que não surpreende a ausência de ligações a redes de partilha da produção científica, como repositórios ou redes sociais como Academia.edu ou Researchgate.

A publicação de conteúdos em formato bilingue – português e inglês – será eventualmente coerente com as contínuas pressões para a verticalização da língua em torno do ambiente anglo-saxónico, frequentemente solicitadas em concursos públicos a financiamento em Portugal, por parte de entidades com a FCT. Imagina-se que a consolidação de alguns desses procedimentos, aliada a possíveis estratégias e parcerias de internacionalização, possam conduzir a este cenário. A identidade visual, como marca do próprio centro, cumpre-se na amostra, apesar de três unidades não apresentarem qualquer referência a este respeito.

Os *sites* dos centros ligam-se pouco às redes sociais. Como segunda dimensão em análise, e ainda que pudéssemos ter ido mais longe na análise dos conteúdos publicados, no envolvimento que suscita (comentários, partilhas, *trending topics*, etc.), os eixos abordados revelam algumas perplexidades, que começam justamente na ausência de dinamização de qualquer rede em 13 centros com responsabilidades importantes na exigência de visibilidade que a ciência deve apresentar, e ainda mais no quadro de financiamentos consideráveis.

Este estudo deixa outras questões a avaliar, eventualmente, no que toca às redes sociais, nomeadamente com a identificação concreta das reações que os *posts* nestes espaços geram, sobretudo naquelas de âmbito mais substantivo, como comentários ou *retweets*. Existe neste quadro uma dimensão participativa passiva (Ribeiro, 2013), sobretudo com mecanismos de reação rápida, possivelmente desprovidas de reflexividade e espírito crítico, como são os casos da ação ‘gosto’ ou *retweetar*, daí que poderia ser interessante medir o quadro das interações que são promovidas neste quadro. O aparente silêncio que se verifica nestes ambientes – visível pela ausência de níveis elevados de comentários – pode deixar interpretações de que a presença dos centros de investigação presta-se apenas a um nível mínimo de “informação da ciência”, de mera identificação da instituição, da publicação ou visibilidade superficial do trabalho realizado pelos investigadores. Esta investigação destaca, por isso, um evidente meta-discurso da ciência, em que cientistas falam apenas para cientistas, num circuito fechado com potencialidades para se expandir, mas com notórias resistências e silêncios.

## Referências bibliográficas

BURNS, T.; O'CONNOR, D. & STOCKLMAYER, S. M. (2003) – Science Communication: a contemporary definition. Public Understanding of Science. [Em linha]. Vol. 12 [Consultado 08 setembro 2016] Disponível na Internet: <URL: <http://pus.sagepub.com/content/12/2/183.extract> >. Online ISSN 1361–6609

CARVALHO, A.; Cabecinhas, R. (2004) – Comunicação da ciência: perspectivas e desafios. Comunicação e Sociedade. ISSN 2183–3575. Vol. 6 (2004), p. 5–10

CONFIBERCOM / Escola de Comunicação e Artes de São Paulo / SOCICOM – Federação Brasileira de Associações Científicas e Acadêmicas – 1º Congresso Mundial de Comunicação Ibero-Americana. MARTINS, Moisés de Lemos – A política científica e tecnológica em Portugal e as ciências da comunicação : prioridades e indecisões – São Paulo. 2012

EXPRESSO. Adonis: “O monoteísmo colonizou o nosso cérebro”. [Em Linha] Lisboa; Grupo Impresa [Consultado 15 de outubro de 2016] Disponível na Internet: <URL: <http://expresso.sapo.pt/cultura/2016-11-05-Adonis-O-monoteismo-colonizou-o-nosso-cerebro>>

FELT, U., NOWOTNY, H.; TASCHWER, K. (1995) – Wissenschaftsforschung: Eine Einführung. Frankfurt: Campus

FUNDAÇÃO PARA A CIÊNCIA E A TECNOLOGIA. Resultados Globais da Avaliação de Unidades I&D 2013. [Em Linha] Lisboa; FCT [Consultado 15 de outubro de 2016] Disponível na Internet: <URL: [www.fct.pt](http://www.fct.pt)>

GREENBERG, D. (2001) – Science, Money and Politics: Political Triumph and Ethical Erosion. Chicago: University of Chicago Press

KRIEGHBAUM, H. (1970) – A ciência e os meios de comunicação de massa. Rio de Janeiro: Edições Correio da Manhã

KURATH, Monika; GISLER, Priska (2009) – Informing, involving or engaging? Science communication, in the ages of atom-, bio-, and nanotechnology. Public Understanding of Science. [Em linha]. Vol. 18. No. 5 [Consultado 03 setembro 2016] Disponível na Internet: <URL: <https://www.hks.harvard.edu/sdn/articles/files/KurathGisler-InformingEngaging.pdf>>. Online ISSN 1361–6609

MAGALHÃES, Ricardina (2015) – A Comunicação Estratégica aplicada à divulgação da Ciência. O caso do Centro de Estudos de Comunicação e Sociedade – Observatorio (OBS\*) Journal. [Em linha]. Vol. 9. No. 4. [Consultado 10 setembro 2016] Disponível na Internet: <URL: <http://obs.obercom.pt/index.php/obs/article/view/850> >. ISSN 1646–5954

MARKTEST. Os Portugueses e as Redes Sociais 2016. [Em Linha] Lisboa; Grupo Marktest [Consultado 03 de outubro de 2016] Disponível na Internet: <URL: <http://www.marktest.com/wap/a/grp/p~96.aspx>>

PEW RESEARCH CENTER, INTERNET, SCIENCE & TECH – How Scientists Engage the Public. [Em Linha] Washington DC: Pew Research Center, Internet, Science & Tech. [Consultado 21 de

outubro de 2016] Disponível na Internet: <URL: <http://www.pewinternet.org/2015/02/15/how-scientists-engage-public/>>

RIBEIRO, Fábio Fonseca (2013) - A participação dos cidadãos nos média portugueses : estímulos e constrangimentos. Braga: Instituto de Ciências Sociais da Universidade do Minho. 512 p. Tese de doutoramento

WOLTON, D. (2009) - *Informer n'est pas communiquer*. ISBN: 978-227-106-8200 (2009). Paris: CNRS Éd., coll. Débats

---

<sup>i</sup> Por questões éticas e pelo distanciamento que se pretende concretizar, nomeadamente com a pertença institucional do autor deste artigo ao Centro de Estudos de Comunicação e Sociedade (CECS), da Universidade do Minho, optou-se por não incluir esta unidade de investigação na amostra estudada. Recorde-se que este centro inclui o grupo de 25 avaliados como “Excelente”, nos painéis de Ciências Sociais e no de Humanidades. À data da avaliação de unidades I&D, em 2013, o CECS contava com 53 membros integrados.

<sup>ii</sup> Nota ainda para a curiosa estratégia do Centro de Estudos de Sociologia e Estética Musical (CESEM), que opta por criar um grupo de discussão privado e por isso não foi integrado nesta lógica. O VICARTE apresenta uma página que em detrimento de “gostos” apresenta “amigos”. Por não utilizarem a mesma nomenclatura, e para efeitos de uma comparação efetiva e coerente, ficaram fora desta consideração.

---

## Marketing digital em portais de periódicos científicos de acesso aberto

---

Juliana Aparecida Gulka

Biblioteca Universitária da Universidade Federal de Santa Catarina (BU/UFSC)

[juliana.gulka@ufsc.com](mailto:juliana.gulka@ufsc.com)

Elaine Rosangela de Oliveira Lucas

Universidade do Estado de Santa Catarina (UDESC)

[lanilucas@gmail.com](mailto:lanilucas@gmail.com)

Ronaldo Ferreira de Araújo

Universidade Federal de Alagoas (UFAL)

[ronaldfa@gmail.com](mailto:ronaldfa@gmail.com)

**Resumo:** Discorre brevemente sobre os periódicos científicos, o Acesso Aberto e como as tecnologias disponíveis aplicadas a este último promoveu iniciativas de implantação dos Portais de Periódicos de Acesso Aberto. Reflete sobre o *marketing* científico digital bem como sua aplicação em Portais de Periódicos Científicos no que se refere a práticas voltadas aos pesquisadores, editores e instituições de pesquisa que atuam na dinâmica da comunicação científica. Reforça a necessidade da criação de um plano de *marketing* científico digital alinhando o Portal com as estratégias necessárias à promoção de seus periódicos, atendendo critérios dos indexadores, realizando divulgação científica, analisando feedbacks e monitorando os ambientes digitais periodicamente com a finalidade de atingir melhores resultados. Dentre os aspectos conclusivos considera-se que a gestão da presença digital de um Portal de Periódicos por meio do *marketing* corrobora com os princípios do acesso aberto, à medida que favorece a disseminação das publicações e garante a visibilidade e acessibilidade a informação científica.

**Palavras-chave:** *Marketing* científico digital, Visibilidade, Portal de Periódicos, Comunicação científica, Acesso Aberto



## Digital marketing in portals for scientific open access Journals

**Abstract:** The article intends to briefly present scientific journals, open access, and how the available applied technologies for the latter promoted implementation initiatives of portals for open access journals. The study also deals with scientific digital marketing and its application in portals for scientific journals with regard to practices aimed at researchers, editors, and research institutions operating in the field of the dynamics of scientific communication. The article reinforces the need to create a scientific digital marketing plan aligning the portal with the necessary strategies to promote their journals, fulfilling criteria of the indexers, conducting scientific publishing, analyzing feedback and monitoring the digital environments periodically in order to achieve better results. Among the conclusive aspects it is considered that the management of the digital presence of a portal for journals through marketing corroborates the principles of open access, as it benefits the dissemination of publications and ensures the visibility and accessibility of scientific information.

**Keywords:** Scientific digital marketing, Visibility, Portal for journals, Scientific communication, Open access

## Introdução

A transição do modelo impresso para o eletrônico demandou novas formas de gerir as coleções de periódicos e promoveu o surgimento de bases de dados e Portais de Periódicos Científicos. Esses Portais desempenham um papel de agrupamento de periódicos institucionais e auxiliam na gestão de qualidade das publicações. Com o movimento de Acesso Aberto, o uso de periódicos eletrônicos foi intensificado, de modo que os Portais também ganharam mais espaço à medida que forneciam o apoio a essa forma inovadora de comunicação científica.

Em um contexto de grande conectividade, um Portal de Periódicos precisa acompanhar as tendências de presença digital, além de utilizar plataformas e estratégias de forma planejada. O *marketing*, por sua vez, pode auxiliar no aproveitamento das oportunidades que o mundo digital oferece e potencializar essas estratégias. O propósito da presente comunicação é apresentar como o *marketing* digital é pertinente às iniciativas de um Portal de Periódicos Científicos de Acesso Aberto com base em uma pesquisa bibliográfica. Leva-se em conta que o Brasil estima a presença de 151 Portais de Periódicos Científicos (Silveira, 2016), vinculados a instituições públicas e privadas, o que estimula o estudo de produtos e serviços na referida área.

A implantação de um Portal de Periódicos contribui para: «[...] o aumento da visibilidade e do valor público das instituições» (Rodrigues; Fachin, 2010, p. 38), indicando a qualidade da entidade à medida que esta toma para si a responsabilidade pela preservação e disseminação dos periódicos, contribuindo para a expansão da comunicação científica.

Paralelas a isso, novas modalidades de *marketing* surgem para atender segmentos específicos, nos quais produtos intangíveis também apresentam valor de mercado (Bizzocchi, 2001).

O *marketing* científico, sendo um tipo de *marketing* especializado, contribui para o crescimento e alavancagem do mercado de produtos científicos (Araujo, 2015), podendo oferecer informação de qualidade com linguagem simples e acessível. Para Bomfá et al. (2009), este é justamente o desafio do *marketing* científico: levar o discurso da ciência a um público amplo, promovendo, divulgando e reconhecendo as pesquisas, projetando autores e alcançando prestígio e visibilidade. O objetivo do *Marketing Científico* é potencializado quando utilizado nos periódicos Científicos, sobretudo nos disponibilizados em Acesso Aberto.

### **Periódicos científicos e Acesso Aberto: contextualização histórica**

A comunicação com finalidade de passar conhecimentos adiante é uma prática habitual. Em relação à ciência, «ninguém pode afirmar quando foi que se começou a fazer pesquisa científica e, por conseguinte, quando, pela primeira vez, houve comunicação científica» (Meadows, 1999, p. 3), mas sabe-se que ela influencia a sociedade há séculos.

A fala e a escrita são os artifícios utilizados para a ciência ser comunicada. Os discursos registrados, no entanto, se materializam e ficam preservados além do espaço e tempo do próprio autor. A produção de materiais científicos foi fortemente influenciada pela imprensa, que, por sua vez, facilitou o surgimento dos periódicos como substitutos das trocas de cartas entre estudiosos e pesquisadores.

De acordo com Gulka (2016, p. 40), «foi nesse contexto que os periódicos científicos foram criados, buscando atender uma expectativa de transição para um formato que alcançasse grupos maiores e fosse, ao contrário das cartas, mais fácil de reproduzir. »

Em janeiro de 1665, em Paris (França), foi criado o periódico *Journal de Sçavans* (Meadows, 1999). Em março do mesmo ano, em Londres (Inglaterra), surgiu o *Philosophical Transactions* (Mueller, 2000). Embora sejam do mesmo período histórico, é o segundo que é reconhecido como o precursor do periódico científico moderno, por publicar somente materiais de cunho científico (Meadows, 1999; Bomfá, 2003).

Quanto às suas características, um periódico possui periodicidade, com frequência regular de fascículos, dentro de uma área específica do conhecimento ou com amplitude geral (Fachin; Hillesheim, 2006), e trazem contribuições de diferentes autores em uma mesma edição, seguindo padrões de normalização.

De acordo com Gulka (2016, p. 46), «ao periódico cabe a função de registrar e fazer circular informações, priorizando as descobertas científicas. A comunicação científica feita por meio de periódicos também permite que os cientistas e/ou autores sejam reconhecidos e recebam prestígio por suas pesquisas à medida que o trabalho é apreciado pelos pares. »

Para Suaiden (2008, p. 9), os periódicos científicos «[...] desempenham um papel importante por suas funções de registro, legitimação de autoria, disseminação da produção científica e memória da ciência. »

Em meados da década de 70, iniciou-se a contestação do modelo de publicação em periódicos impressos em formato tradicional (Mueller, 2006). Isso se deu principalmente em virtude dos elevados custos de manutenção de coleções em bibliotecas e de produção dos materiais. Para Silveira (2016, p. 65), «reconhecida mundialmente como *serial crisis* (em português: crise dos periódicos), é um fenômeno de muita oferta de informação científica e restrição de acesso, isto é, ao mesmo tempo que se produzia muita pesquisa científica, os pesquisadores não tinham acesso ao conteúdo delas, pois as instituições de pesquisa não davam conta de pagar para acessar o conteúdo das publicações científicas. »

Aliada à crise, a tecnologia da informação e da comunicação predizia um novo momento, e dessa forma foram originadas as primeiras iniciativas de acesso livre ao conteúdo acadêmico (Mueller, 2006).

As alternativas objetivavam a modificação do formato do periódico e o seu aprimoramento, visando a qualidade científica e a disseminação da informação.

«A mudança para suportes eletrônicos foi, certamente, a mais significativa, tendo sido implementada de forma gradativa pelas publicações. A mudança de suporte modifica o modo como se acessa, utiliza e assimila a informação. Os periódicos científicos eletrônicos são mais versáteis, rápidos e ignoram barreiras geográficas. » (Gulka, 2016, p. 49).

A partir da década de 90, os periódicos passaram a ser publicados de forma integral na *web*, e os leitores puderam, então, conseguir acessar todo o conteúdo de interesse de seus próprios computadores (Stumpf, 1996).

Em 1999, com a finalidade de discutir o assunto acerca do acesso à informação, a Convenção de Santa Fé, no Novo México, trabalhou no modelo que evoluiu para o *Open Archives Initiative* (OAI) (Fausto, 2013). De acordo com Triska e Café (2001, p. 92), a convenção estabeleceu os «princípios básicos de uma nova filosofia para a publicação científica». Esforços internacionais também resultaram no *Budapest Open Access Initiative* (BOAI), em 2001, culminando em duas iniciativas: a Via Verde, para os repositórios e o autoarquivamento, quando o pesquisador deposita seus arquivos de modo aberto; e a Via Dourada, para os periódicos de Acesso Aberto. » (Budapest Open Access Initiative, 2002; Harnad et al., 2004). Do BOAI também resultou o primeiro protocolo de interoperabilidade entre fontes digitais, o *Open Access Initiative – Protocol for Metadata Harvesting* (OAI-PMH), e foram definidos os princípios do movimento: acessibilidade (interoperabilidade), fidedignidade (revisão por pares) e disseminação (autoarquivamento) (Triska; Café, 2001).

No Brasil, alguns projetos de iniciativa do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) buscam seguir os princípios do Acesso Aberto. São eles: Sistema Eletrônico de Administração de Conferências (SOAC), Sistema Eletrônico de Teses e

Dissertações (TEDE), Sistema para Construção de Repositórios Institucionais Digitais (DSpace) e *Open Journal System* (OJS), intitulado de Sistema Eletrônico de Editoração de Revistas (SEER) (Instituto Brasileiro de Informação em Ciência e Tecnologia, 2016a).

O movimento de Acesso Aberto em consonância com a tecnologia fez com que a circulação de periódicos aumentasse, muitas vezes ficando dispersos pela rede. Essas mudanças culminaram em alterações na forma como as academias lidam com as publicações científicas, resultando em um novo modo de pensar, ocasionando a criação de espaços e serviços diferenciados no que tange aos periódicos eletrônicos. Nesse contexto, foram criados os Portais de Periódicos.

### **Portais de Periódicos Científicos**

Os Portais de Periódicos Científicos são consequência do Acesso Aberto, que por conta da dinâmica para romper as barreiras de acesso em conjunto com as tecnologias disponíveis fez surgir um novo ambiente, responsável por reunir os periódicos científicos que se encontravam dispersos pela rede.

Marra e Weitzel (2015) comentam que foi o Acesso Aberto que facilitou as iniciativas de implantação dos Portais de Periódicos, sobretudo em universidades. A implantação de um Portal de Periódicos propicia questões estratégicas, ao contribuir para: «[...] o aumento da visibilidade e do valor público das instituições» (Rodrigues; Fachin, 2010, p. 38), indicando a qualidade da entidade à medida que esta toma para si a responsabilidade pela preservação e disseminação dos periódicos, contribuindo para a expansão da comunicação científica.

Para Oliveira e Cunha (2016) a implantação de um Portal organiza os periódicos de uma instituição e define a responsabilidade dos atores envolvidos. Além disso, a responsabilidade institucional fornece mais credibilidade, seja através da identidade visual, conselho editorial ou políticas editoriais, seja em relação às competências necessárias na preservação dos dados, padrões, suporte técnico e treinamentos, já que «o papel da instituição concentra-se nas questões que se referem ao coletivo de todos os periódicos.» (Garrido; Rodrigues, 2010, p. 61). Dessa forma, um Portal funciona como agregador de periódicos de uma mesma instituição, e esta, por sua vez, funciona como uma metaeditora (Rodrigues; Fachin, 2010).

Do ponto de vista administrativo, a instituição passa a oferecer um apoio no gerenciamento dos periódicos científicos, com estrutura no que tange aos recursos humanos, tecnológicos e financeiros (Carvalho, 2014).

«Enquanto os periódicos isoladamente são responsáveis por todo o processo editorial, quando reunidos em Portais, várias questões passam a ser de interesse comum e podem ser partilhadas, otimizando recursos, padronizando processos, viabilizando processos colaborativos e exigindo diretrizes e normas específicas. A organização das publicações científicas em meio digital em Portal requer, além de espaço seguro para armazenamento dos dados, estrutura para pesquisas, testes, cursos em fluxo contínuo, oferta

de bolsas para alunos da graduação e pós-graduação e estabelecimento de parcerias entre os órgãos da instituição envolvidos. » (Rodrigues; Fachin, 2008, p. 8)

Um Portal de Periódicos científicos funciona como centralizador e, em termos de gestão, precisa de uma estrutura planejada, com responsabilidade técnica, institucional e operacional. Para Gulka (2016, p. 59),

«um Portal de Periódicos é um ambiente em Acesso Aberto, padronizado, que busca reunir em um mesmo local diferentes periódicos científicos, e auxiliar no gerenciamento destas revistas por meio de serviços como hospedagem, configuração e customização, suporte a editores, preservação digital, segurança e interoperabilidade de metadados, indexação em bases de dados e sustentabilidade editorial.»

De acordo com Ferreira (2008a, p. 11), um Portal de Periódicos científicos pode garantir «a disseminação, a visibilidade e a acessibilidade da produção científica, buscando gerar mecanismos que garantam, preservem e melhorem sua visibilidade, uso e responsabilidade social. »

O Portal de Periódicos possui funções educativas, tecnológicas, sociais e políticas. Conforme afirma Silveira (2016, p. 88),

«a função educativa do Portal é disponibilizar produtos e serviços que viabilizem aos editores e suas equipes a desenvolver a competência informacional voltada para a editoração científica, os princípios do Acesso Aberto e as funções da comunicação científica. A função tecnológica é servir e prover por melhores recursos tendo em vista as mudanças sociais e tecnológicas vigentes. A função social e política é garantir o direito ao acesso às informações públicas, bem como ser um articulador político institucional e despertar a conscientização da comunidade científica dos benefícios das fontes em Acesso Aberto. »

Em relação aos serviços oferecidos por um Portal de Periódicos científicos de Acesso Aberto, Silveira (2016) baseia-se na literatura e lista os principais segmentos de atuação: assessoria e capacitação; controle, normalização, edição e indexação; segurança e preservação; avaliação, métricas; relatórios de gestão e, por fim, *marketing* científico digital.

### **Desbravando o *Marketing* Científico Digital**

Visto como a junção do '*marketing* científico' com o '*marketing* digital', o *marketing* científico digital é considerado uma «estratégia empregada em produtos da ciência, aliada à comunicação científica e comunicação digital, com o intuito de oferecer serviços alinhados às necessidades dos usuários, visando à promoção de periódicos, pesquisas e pesquisadores, com foco na visibilidade científica. » (Araújo, 2015, p. 72).

Entre as vantagens que o ambiente digital proporciona ao *marketing* científico, estão a dinâmica de propagação de mensagens, sua rapidez, e seus aspectos de interatividade e colaboração. De acordo com Araújo (2015, p. 72), «a escolha das estratégias de *marketing* científico digital precisa refletir o tipo de imagem que se quer transmitir, bem como atender ao perfil dos clientes/usuários que possui ou que se pretende alcançar. » Dessa forma, o *marketing* científico digital pode promover as publicações científicas, à medida que seus

recursos agregam valor e projetam o nome dos pesquisadores, inclusive internacionalmente (Bomfá et al., 2009).

Vale salientar, ainda, que «a presença digital verdadeira tem que ser orientada por objetivos e gerar oportunidades reais de negócios para a entidade que a coloca em prática» (Strutzel, 2015b). Ou seja, utilizar os canais digitais para atingir um objetivo bem definido da instituição é um fator crucial para o retorno dessa presença.

Sendo assim, é possível descrever, de forma simplificada, que a presença digital seria a «existência de algo no ambiente digital» (Gabriel, 2010, p. 249), ou ainda a representação da «[...] existência de uma entidade nas mídias digitais e sociais» (Strutzel, 2015a, p. 87), e pode ser, de acordo com os autores, categorizada em três grupos: própria, espontânea e paga.

Por fim, é possível refletir sobre o papel do *marketing* científico digital no aumento da visibilidade dos Portais de Periódicos Científicos. Para tanto, é preciso apresentar esse tipo de *marketing* como forma de promoção e divulgação dos Portais de Periódicos para os quais seus editores devem: construir e manter uma presença online; oferecer um conteúdo adequado aos ambientes que atuam; e estabelecer uma atuação responsiva.

### **Possibilidades de *marketing* científico digital para Portais de Periódicos Científicos**

As possibilidades de aplicação do *marketing* científico digital refletem práticas voltadas aos pesquisadores, editores e instituições de pesquisa que, por meio dos artigos, periódicos científicos e Portais de Periódicos, atuam na dinâmica da comunicação e divulgação científica.

De acordo com Gulka (2016, p. 185), «a Ciência da Informação possui aspectos interdisciplinares que contemplam o trabalho realizado no âmbito do *marketing* digital, que alia conceitos de tecnologia, internet, administração, etc.», sendo que os profissionais que atuam nessa área têm a necessidade de se ajustar às novas demandas de trabalho que surgem por causa dos ambientes digitais.

Ao assumir o contexto do *marketing* científico digital, os Portais de Periódicos passam também a cumprir critérios estabelecidos por indexadores que refletem diretamente na disseminação dos periódicos que agrupam. Um dos indexadores a exigir esse tipo de atuação é a SciELO, que trata de *marketing* e divulgação, exigindo que os periódicos indexados realizem as atividades de forma individual ou coletiva. O Portal de Periódicos, a medida que representa os interesses dos periódicos institucionais de forma coletiva, tem plena responsabilidade em auxiliar os editores no cumprimento dos critérios, bem como de atuar na vanguarda das tendências do mundo digital.

Um exemplo de divulgação científica realizada em território brasileiro, que segue os princípios do *marketing* científico digital, é o 'Projeto FAPESP 2013/2014 - os periódicos

científicos brasileiros: estratégias para expandir e melhorar a comunicação com a sociedade', desenvolvido pelo Portal de Revistas da Universidade de São Paulo (USP) em parceria com o Laboratório de Estudos Avançados em Jornalismo da Universidade de Campinas (UNICAMP). Tal iniciativa objetiva elaborar conteúdos de cunho científico, mas com linguagem acessível, visando a divulgação dos resultados de pesquisas científicas produzidas.

Em sua pesquisa em nível de mestrado, Silveira (2016) estabeleceu uma proposta de diretrizes com elementos estruturantes para a gestão, manutenção e sustentabilidade de Portais de Periódicos, nos quais cita, entre outras coisas, o serviço de marketing digital para a divulgação do Portal e dos periódicos. A justificativa é que «a divulgação da informação maximiza a visibilidade. Ampliando as formas da divulgação científica, aproxima-se a sociedade da ciência produzida pela universidade. Aumentando a disseminação, conseqüentemente aumenta a visibilidade da produção. »

A autora recomenda fortemente a criação de um planejamento de *marketing* científico digital que envolva o departamento de jornalismo e comunicação da instituição, e cita como ferramentas e estratégias a utilização de: Facebook, Twitter, email *marketing*, notícias, press releases e *blog*.

De acordo com Vicente (2015, p. 31), «termos como Facebook, Twitter, Google+, LinkedIn, Youtube, Instagram não frequentavam o cotidiano, bem como o meio científico. Provindas de uma nova sensibilidade global, essas novas nomenclaturas interferem – direta ou indiretamente – no cotidiano de milhões de pessoas. »

Há uma variedade grande de canais e ferramentas, além das apontadas por Silveira (2016) e Vicente (2015), que podem ser utilizadas com o intuito de fortalecer a presença digital de um Portal e seus periódicos, mas salienta-se que sua adoção e gestão dependem especificamente do planejamento da instituição.

As mídias sociais são aliadas na divulgação científica e na geração de debates e discussões no ambiente online. No uso do Facebook, especificamente, destaca-se a fanpage do Portal de Periódicos da Universidade Federal de Santa Catarina, que conta com mais de 6.500 curtidas e reúne interessados em acompanhar os rumos da ciência e tecnologia e as novidades da comunicação científica.

Já o Twitter, foi apontado por um estudo pioneiro em 2013, por Juan Pablo Alperin, da Universidade de Stanford, como o canal que mais registrou menções e recomendações aos artigos indexados pela SciELO (Marques, 2014).

O *blog*, por sua vez, é uma alternativa quando o *site* institucional é estático, viabilizado somente em uma estrutura pré-definida. *Blogs* são páginas digitais que possuem entradas de texto, também chamadas de *posts*, em ordem cronológica e por assunto.



Utilizam geralmente linguagem mais informal e possuem espaço para comentários (Strutzel, 2015a).

No âmbito do *marketing* científico digital, os *blogs* possibilitam uma maior interação entre autores, leitores e editores, ampliando a visibilidade e disseminação tanto para a comunidade específica quanto para a sociedade em geral (Príncipe, 2013).

Quanto aos Portais de Periódicos Científicos, um *blog* funcionaria não apenas como um local para curadoria de conteúdo em relação a seus periódicos, mas para criação de conteúdo autêntico referente ao mundo científico.

De acordo com Príncipe (2013), os *blogs* científicos são pouco utilizados no Brasil, mas muito difundidos na Europa e nos Estados Unidos. Para a autora, esses *blogs* podem funcionar além da divulgação científica, como uma rede de discussão de pesquisas científicas.

Exemplos brasileiros que podem ser citados com relação à adoção de *blogs* científicos, embora não sejam mantidos por Portais de Periódicos, são: SciELO em Perspectiva<sup>1</sup> e o *Blog* da Associação Brasileira de Editores Científicos<sup>2</sup> (ABEC).

As duas iniciativas contam com a colaboração de convidados que publicam textos referentes a temas pertinentes da ciência, promovendo a reflexão e o debate, contando com a divulgação científica das revistas indexadas, no caso do *blog* da SciELO, e das revistas associadas, no caso da ABEC.

Os *blogs* citados como exemplo também utilizam da estratégia de envio de *newsletter*, via email *marketing*, outro canal que apresenta potencial de uso pelos Portais de Periódicos.

O email *marketing* ou *newsletter* possibilitam a segmentação, personalização, mensuração e rastreamento (dependendo da plataforma utilizada), integração com outras mídias, assincronicidade, permanência, difusão digital, preço baixo, possibilidade de envio de mensagem multimídia, riqueza de conteúdo, capacidade de viralização e possibilidades de relacionamento (Gabriel, 2010).

O fato do email *marketing* permitir a integração com outras mídias amplifica sua utilização em conjunto com o *blog*. De acordo com Gabriel (2010), o email é um dos principais canais de relacionamento e compartilhamento de informações, sendo uma das fontes de informação essenciais para leitores de *blogs*, que o utilizam para receber as novas notícias.

---

<sup>1</sup> <http://blog.scielo.org/#>

<sup>2</sup> <http://www.abecbrasil.org.br/novo/blog-list-layout/>



De acordo com Gabriel (2010), novas tecnologias – como mídias sociais, por exemplo – não causam a obsolescência das antigas, como é o caso do email *marketing*. Suas estratégias visam estabelecer um canal de conexão com o público interessado nos conteúdos publicados.

Outro fator que deve ser considerado é a integração dos canais utilizados, de forma que uma plataforma se alimente e também estimule a outra, promovendo uma mescla e criando oportunidades de *marketing* (Gabriel, 2010). Um exemplo disso seria utilizar o Facebook para publicar chamadas dizendo que em outro canal tem conteúdo exclusivo, e assim reter a atenção do público.

Para Silveira (2016), «a ciência brasileira necessita de ações que aproximem a produção científica da sociedade, por isso a divulgação, por meio das redes sociais, ou a elaboração de entrevistas com os autores das publicações, entre outras, são ações importantes para diminuir essa distância entre a ciência e a sociedade.»

Dessa forma, as ferramentas da *web* têm a contribuir para a visibilidade científica, sendo que «o *marketing* científico digital pode ser um aliado nessa tarefa, sobretudo para os periódicos eletrônicos.» (Araújo, 2015, p. 71).

Essa informação é complementada por Carvalho (2003), que afirma que o *marketing* pode ser utilizado para tornar as imagens da ciência, pesquisadores e instituições mais sólidas, pois estimula e aumenta a confiança e os investimentos em pesquisas.

O fato é que o uso dos recursos de *marketing* aproxima a sociedade da ciência, visto que esse é o objetivo principal do *marketing* científico digital. Além do mais, consolida a imagem da ciência, dos cientistas e das instituições que fazem pesquisa, trazendo reconhecimento, prestígio, visibilidade, e estimulando a confiança do público leigo.

## **Considerações finais**

Entender a tecnologia envolvida é uma das premissas básicas para explorar de forma satisfatória o ambiente digital. Gabriel (2010) ressalta a combinação de três elementos para o desenvolvimento de uma boa estratégia de presença digital: *marketing*, tecnologia e design.

E, para além da presença online que atende o ‘existir’, o movimento a ser seguido em qualquer estratégia de *marketing* científico digital é o de ‘atrair’, ‘relacionar’ e ‘engajar’.

Sendo o Portal de Periódicos um local que reúne um grupo de publicações institucionais, este tem também maior força para promover ações que visem ainda mais o sucesso desses veículos em um ambiente que está se tornando cada vez mais exigente, seja no sentido de qualidade, seja para alcançar diversos espaços oriundos da *web* 2.0.

Salienta-se que o ideal é a criação de um plano de *marketing* científico digital como um documento norteador não apenas das ações relacionadas a canais específicos, mas de

*marketing* digital como um todo, alinhando o Portal com as estratégias necessárias também para a promoção de seus periódicos, atendendo aos critérios dos indexadores, realizando divulgação científica, analisando feedbacks e monitorando os ambientes digitais periodicamente com a finalidade de atingir melhores resultados.

A gestão da presença digital de um Portal de Periódicos por meio do *marketing* vai ao encontro dos princípios do Acesso Aberto, a medida que favorece a disseminação das publicações e garante visibilidade e acessibilidade à informação científica.

### Referências bibliográficas

ARAÚJO, Ronaldo Ferreira de (2015) – Marketing científico digital e métricas alternativas para periódicos: da visibilidade ao engajamento. *Perspectivas em Ciência da Informação* [em linha]. Vol. 20, n.º 3, p. 67–84. [Consult. 06 Abr. 2016]. Disponível na internet: <<http://www.scielo.br/pdf/pci/v20n3/1413-9936-pci-20-03-00067.pdf>>. ISSN 1981–5344

BIZZOCCHI, Aldo (2001) – *Marketing científico: o papel do marketing na difusão da cultura em geral e da ciência em particular* [em linha]. [consult. 26 Abr. 2015] Disponível na internet: <<http://www.aldobizzocchi.com.br/artigo22.asp>>.

BOMFÁ, Cláudia Regina Ziliotto [et al.] (2009) – Marketing Científico Eletrônico: um novo conceito voltado para periódicos eletrônicos. *Estudos em Comunicação* [em linha]. n.º 5, p. 193–215. [Consult. 06 Abr. 2016]. Disponível na internet: <<http://www.ec.ubi.pt/ec/05/pdf/10-bomfa-marketing.pdf>>. ISSN 1646–4974

BOMFÁ, Cláudia Regina Ziliotto (2003) – *Revistas Científicas em Mídia Digit@l: critérios e procedimentos para publicação*. Florianópolis: Visual Books. ISBN 8575021257

BUDAPEST OPEN ACCESS INITIATIVE (2002) – Read the Budapest Open Access Initiative. [Em linha]. [Consult. 06 Abr. 2016]. Disponível na internet: <<http://www.budapestopenaccessinitiative.org/read>>.

CARVALHO, Teila de Oliveira (2014) – *Portais de periódicos científicos em bibliotecas acadêmicas: uma nova função no contexto do acesso aberto à informação científica*. 2014. Brasília: Universidade de Brasília. Monografia.

CARVALHO, Alessandra Pinto de (2003) – *A divulgação e o marketing da ciência: uma análise do documentário como instrumento híbrido de comunicação científica pública, por encomenda de instituições públicas de pesquisa à empresa Videociência*. São Paulo: Universidade Metodista de São Paulo. Tese de Doutorado.

FACHIN, Gleisy Regina Bories; HILLESHEIM, Araci Isaltina de Andrade (2006) – *Periódico Científico: padronização e organização*. Florianópolis: UFSC. ISBN 8532803415.

FAUSTO, Sibebe (2013) – Evolução do Acesso Aberto: breve histórico. [em linha]. [Consult. 06 Out. 2016]. Disponível na internet: <<http://blog.scielo.org/blog/2013/10/21/evolucao-do-acesso-aberto-breve-historico/#.WA0NnPorLIU>>.

FERREIRA, Sueli Mara Soares Pinto (2008) – Repositórios versus revistas científicas: convergência e convivências. In *Mais sobre revistas científicas: em foco a gestão*. São Paulo: Senac/Cengage Learning, 2008. ISBN 9788573597523. p. 111 – 137.

GABRIEL, Martha (2010) – *Marketing na era digital: conceitos, plataformas e estratégias*. São Paulo: Novatec. ISBN 9788575222577

GARRIDO, Isadora dos Santos; RODRIGUES, Rosângela Schwarz (2010) – Portais de periódicos científicos online: organização institucional das publicações. *Perspectivas em Ciência da Informação* [em linha]. Vol. 15, n.º 2, p. 56–72. [Consult. 06 Abr. 2016]. Disponível na internet: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/articlo/view/943>> . ISSN 1981–5344

GULKA, Juliana Aparecida (2016) – *Análise de presença digital: um estudo do Portal de Periódicos UFSC*. Florianópolis: Universidade do Estado de Santa Catarina. 203 p. Dissertação de mestrado.

HARNAD, Stevan [et al.] (2004) – The green and the gold roads to Open Access. *Nature*, [em linha]. [Consult. 15 Set. 2016]. Disponível na internet: <<http://www.nature.com/nature/focus/accessdebate/21.html>>

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (2016) – Tecnologias para Informação. [Em linha]. [Consult. 15 Set. 2016]. Disponível na internet: <<http://www.ibict.br/pesquisadesenvolvimento-tecnologico-e-inovacao>>.

MARQUES, Fabrício (2014) – Retuíte ou perea: Estudo indica que o Twitter é a rede social mais usada para divulgar artigos científicos de revistas brasileiras. *Revista Pesquisa FAPESP*, [em linha] ed. 221, jul. [Consult. 15 Set. 2016]. Disponível na internet: <<http://revistapesquisa.fapesp.br/2014/07/15/retuite-ou-pereca/>>.

MARRA, Patrícia dos Santos Caldas; WEITZEL, Simone da Rocha – Portais de periódicos de acesso aberto nas universidades brasileiras: a utilização do Open Journal Systems. In ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16, João Pessoa, 2015 – Actas. João Pessoa: ANCIB, 2015.

MEADOWS, Arthur Jack (1999) – *A comunicação científica*. Brasília: Briquet de Lemos. ISBN 85–85637–15–3

MUELLER, Suzana Pinheiro Machado (2006) – A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, [em linha]. Vol. 35, Nº 2, p. 27–38.

OLIVEIRA, Alexandre Pedro; CUNHA, Miriam Vieira da (2016) – As profissões da informação e os portais de periódicos. Não publicado.

PRINCIPE, Eloisa (2013) – Comunicação científica e redes sociais. In *Fronteiras da Ciência da Informação*. Brasília: IBICT. p.196–216.

RODRIGUES, Rosângela Schwarz; FACHIN, Gleisy Regina Bories – Portal de Periódicos científicos: um trabalho multidisciplinar. *TransInformação* [em linha]. Vol. 22, n.º 1 (2010), p. 33–45. [Consult. 10 Set. 2015]. Disponível na internet: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/483>>. ISSN 2318–0889

SILVEIRA, Lúcia da (2016) – *Portais de periódicos das universidades federais brasileiras: documentos de gestão*. Florianópolis: Universidade do Estado de Santa Catarina. Dissertação de mestrado.

STRUTZEL, Tércio – *Como a Presença Digital impacta o seu negócio?* [em linha]. [S.l.: s.n.], act. 2015b. [consult. 26 Jun. 2015] Disponível na internet: <<http://www.terciostruzel.com.br/como-a-presenca-digital-impacta-o-seu-negocio/>>.

STRUTZEL, Tércio – *Presença Digital*. Rio de Janeiro: Alta Books, 2015a. ISBN 978–85–7608–902–5

STUMPF, Ida Regina Chitto (1996) – Passado e futuro das revistas científicas. *Ciência da Informação*, Vol. 25, N° 3.

SUAIDEN, Emir (2008) – Prefácio. In *Mais sobre revistas científicas: em foco a gestão*. São Paulo: Senac/Cengage Learning.

TRISKA, Ricardo; CAFE, Lúgia (2001) – Arquivos abertos: subprojeto da Biblioteca Digital Brasileira. *Ciência da Informação*, [em linha]. Vol. 30, N° 3, p. 92–96. [Consult. 10 Out. 2016]. Disponível na internet: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652001000300012](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652001000300012)>

VICENTE, Natalí Ilza (2015) – *O uso do Twitter e Facebook para divulgação científica: um estudo netnográfico em perfis de bibliotecas universitárias federais do sul do Brasil*. Florianópolis: Universidade do Estado de Santa Catarina. 184 p. Dissertação de mestrado.

---

## Peças de teatro manuscritas (séculos XVIII e XIX) em acesso aberto da Escola Superior de Teatro e Cinema: projeto apoiado pela Fundação Calouste Gulbenkian

---

Luísa Maria Lousã Marques

Instituto Politécnico de Lisboa, Escola Superior de Teatro e Cinema, Portugal

[luisamarques@estc.ipl.pt](mailto:luisamarques@estc.ipl.pt)

### Resumo

O acervo de peças de teatro manuscritas, oriundas do arquivo histórico do Conservatório Nacional, é constituído por cerca de 2300 espécies (séc. XVIII–XX) entre originais, traduções e imitações. Estas espécies são consideradas, pelo seu valor documental e artístico, edições raras e únicas. O projeto de conservação apoiado pela Fundação Calouste Gulbenkian permitiu que 812 destas espécies (séc. XVIII–XIX) fossem digitalizadas e toda a informação bibliográfica introduzida no catálogo *online* da Biblioteca da ESTC, com acesso, total ou parcial, ao documento digital. Acrescenta-se ainda que o facto de não se encontrarem cópias destes documentos em outros arquivos ou noutros suportes, e a necessária colocação dos mesmos em domínio público, vem justificar este projeto que permite o acesso dos documentos pela comunidade em geral e pela comunidade académica e científica em particular, o que transcende a História do Teatro e inclui a História das Instituições, da Sociologia da Cultura e da História das Literaturas Dramáticas em Portugal e no Brasil.

**Palavras-chave:** Escola Superior de Teatro e Cinema, Instituto Politécnico de Lisboa, Reportório de Teatro, Séculos XVIII–XIX, Manuscritos, Fundação Calouste Gulbenkian

## Manuscript plays (XVIII–XIX centuries) belonging to the Lisbon Theatre and Film School in open access: a project supported by Calouste Gulbenkian Foundation

### Abstract

The collection of manuscript plays, belonging to the historical archive of Conservatório Nacional, consists of about 2300 documents (XVIII–XX centuries), originals, translations and imitations. These species are considered, due to their documentary and artistic value, rare and unique editions. The conservation project was supported by the Calouste Gulbenkian Foundation and allowed that 812 of these species to be scanned (XVIII–XIX centuries) and all the bibliographic information entered in the online catalog of the library with access to all or part of the digital document. The absence of these materials in other documentary supports, and making them available to the public domain, has given access both to the general community and, most particularly, to the academic and scientific community, in a way that transcends the history of theatre and includes the history of institutions, the sociology of culture and the history of performing arts in Portugal and Brazil.

**Keywords:** Lisbon Theatre and Film School, Instituto Politécnico de Lisboa, Theatre repertoire, XVIII–XIX centuries, Manuscript plays, Calouste Gulbenkian Foundation

«Outras constatações interessantes radicam, por exemplo, no facto de, nos teatros públicos que funcionavam em Lisboa nestes primeiros anos do século (Teatro do Bairro Alto, Teatro da Rua dos Condes, Teatro do Salitre e Teatro de S. Carlos), não se falar português (mas, exclusivamente, espanhol, francês e italiano) e também haver falta de actores e de actrizes. Nos anos de 1818 e 1827, no Teatro do Bairro Alto, há notícia de espectáculos de equilíbrio e de títeres, por *troupes* inglesas. [...] Este é, porém, um momento privilegiado da nossa História e no decurso do qual se afirma a consciência da importância da cena na «escrita de teatro» e da necessidade do teatro público para o desenvolvimento de uma dramaturgia nacional. Graças ao ideário liberal reformista, o ensino artístico é criado e encarado como parte integrante, mas autónoma, de uma reforma profunda de todo o sistema público de ensino.» (Vasques, 2012, p. 17 e 44).

### Introdução

A Biblioteca da Escola Superior de Teatro e Cinema herdou da antiga Escola de Teatro do Conservatório Nacional um espólio histórico que foi iniciado pelo próprio Almeida Garrett, no então Conservatório Real de Lisboa (a partir de 1836), espólio esse composto por livros, revistas, documentos de funcionamento da instituição, correspondências pessoais e institucionais, bem como por peças de teatro manuscritas, datadas dos séculos XVIII, XIX e XX. Este património de peças de teatro manuscritas tem origens várias: as que os autores submetiam a Concurso para prémio e as que se submetiam ao parecer censório da Inspecção

Geral dos Theatros, que funcionava no Conservatório. Sabemos que esta documentação passou por diversas vicissitudes entre as quais incêndios, fazendo com que muitas delas estejam completa ou parcialmente queimadas, coladas pela água e ainda rasgadas e incompletas. As peças deste espólio valioso sobreviveram e chegaram até aos nossos dias porque a Professora Doutora Eugénia Vasques, conhecida no antigo Conservatório por ser amante de “papéis velhos”, as salvou do destino do caixote do lixo. Por conseguinte, as peças que conseguiram chegar até aos nossos dias em bom estado de conservação requerem, agora, o nosso empenho em as tratar, recuperar e colocar ao serviço da comunidade e dos investigadores.

Percebe-se, hoje, a importância deste património do teatro português, pois temos espécies que não estão referenciadas, ou estudadas, ou inseridas na bibliografia especializada, ou em bases de dados de teatro e que, nalguns casos, estavam dadas como perdidas na literatura publicada sobre o teatro em Portugal. A sua digitalização é indispensável para que possamos divulgar parte da história do teatro dessa época e, acima de tudo, porque é uma documentação única e de grande valor histórico.

### **Breve contextualização histórica**

Para um melhor enquadramento do acervo de peças manuscritas, que estamos a tratar e a disponibilizar em livre acesso a todos aqueles que investigam sobre este período do teatro (século XVII–XIX) em Portugal, vamos fazer um breve percurso pelo tipo de espetáculos que se produziam e os locais onde eram representados.

A cena teatral desta época, em Portugal, vivia um notório atraso em relação ao que se passava no resto da Europa. «[E]m termos de oferta de espectáculo (muitos dos quais importados do estrangeiro, sobretudo de Espanha e de Itália)», (Vasques, 2009, p. 8) vivia-se uma fase em que o gosto do público, a plebe lisboeta mais endinheirada e frequentadora de um teatro de comédia, só muito lentamente se vai habituando às adaptações e imitações de grandes dramaturgos italianos e franceses. A ópera, destinada a uma elite de público selecionado, foi o género que recebeu maior apoio da monarquia, contribuindo para a construção de «uma rede [de foro privado] de teatros reais [n]o século XVIII» (Vasques, 2009, p. 12). Em paralelo, foram sendo construídos teatros públicos tais como o Teatro do Salitre, o Teatro da Graça, o Teatro Rua dos Condes, o Teatro do Bairro Alto, e por último o Teatro do Ginásio. Deste tempo, fica a ideia de um teatro que vivia sem identidade cultural e nacional, com maioritária expressão em espetáculos musicais, de género sério ou jocoso, na sua maioria traduzidos e adaptados para a cena portuguesa.

Este é o contexto artístico nacional que Almeida Garrett vem encontrar quando regressa «dos exílios de Londres e de Paris, e da Legação de Bruxelas, onde fora Encarregado de Negócios» (Castro, 1954, p. 17). Ao ser escolhido pelo Ministro do Reino, Passos Manuel, para propor «um plano para a fundação de um teatro nacional nesta capital,



o qual, sendo uma escola de bom gosto, contribuí[sse] para a civilização e aperfeiçoamento moral da nação portuguesa» (Castro, 1954, p. 19), Almeida Garrett, em 1836, manda publicar o decreto que dá início à restauração do teatro português. No entanto, este percurso feito em defesa do teatro em Portugal não foi fácil:

«Apesar dos louváveis esforços de Almeida Garrett em defesa de uma arte dramática que contribuiu para a civilização e aperfeiçoamento moral da nação, que deram muitos e importantes frutos, nomeadamente a criação da Inspeção-Geral dos Teatros e Espectáculos (1836) e do Conservatório Real de Arte Dramática (1837) e a construção de raiz de um Teatro Nacional (1846), seria ainda longo o caminho até que o teatro deixasse de ser encarado como simples divertimento ou espaço de convívio social.» (Magalhães, 2007, p. 23)

### **Descrição dos espaços teatrais e das espécies do Arquivo Histórico (peças de teatro manuscritas)**

As peças de teatro manuscritas herdadas do antigo Conservatório Real de Arte Dramática, como disse anteriormente, e que se encontram, agora, no Arquivo Histórico da Escola Superior de Teatro e Cinema, estão, na sua maioria, em mau estado de conservação. As espécies selecionadas, nesta fase e, que foram objeto de um processo de digitalização, obedeceram a um critério: por serem as mais procuradas por investigadores e por terem permanecido desconhecidas, até este momento. Nestes 812 documentos escolhidos, vamos encontrar informações detalhadas sobre nomes dos atores, pareceres censórios, distribuição, marcações de cena, entre outras. «[D]elineiam-se contornos que aludem a tempos e vivências históricas que podem bastas vezes pôr em causa algumas certezas vigentes e obrigar a redefinir conceitos sobre as práticas teatrais de outros tempos» (Filipe, 2014, p. 13). Encontramos ainda, na maioria destas peças, a informação dos espaços teatrais onde foram representadas: Teatro da Rua dos Condes, Teatro do Salitre (mais tarde Teatro das Variedades), Teatro do Ginásio e Teatro Nacional D. Maria II.

O Teatro da Rua dos Condes foi inaugurado em 1766 e nele representavam-se comédias e dramas em francês, italiano e espanhol. Era considerado o único “teatro nacional” e era o local onde se representavam espetáculos decentes e com os melhores atores nacionais. Em 1836, passa a funcionar como “teatro nacional” enquanto se inicia a construção de um novo edifício para este fim. Não durou muito o período de glória deste espaço que é encerrado e reaberto em 1852. No entanto, não foi por muito tempo: «1881–82 foi a última temporada de vida do velho teatro onde se fizera a história da restauração do teatro português» (Vasques, 2009, p. 25).

O Teatro do Salitre foi construído em 1782 e o seu repertório era composto por comédias e entremezes. A partir de 1840 «os alunos das várias escolas do Conservatório Real representavam no Teatro do Salitre um espetáculo de homenagem à rainha D. Maria II no seu aniversário» (Vasques, 2009, p. 36). Por causa da febre-amarela, encerra as suas portas, reabrindo em 1858 com o nome de Teatro das Variedades Dramáticas ou Teatro das Variedades já com um novo ensaiador. Em 1880 passaria a ter o nome de Teatro do Rato.



O Teatro do Ginásio ou Teatro do Ginásio Dramático, sito na Rua Nova da Trindade, «foi um dos mais populares de entre os quatro teatros do Chiado durante a segunda metade do século XIX» (Vasques, 2009, p. 42). Inaugurado em 1846, demolido em 1852 e reaberto em 1852–53, «[o] Teatro do Ginásio vocacionou-se, pois, para os géneros mais populares do teatro (farsa, comédia, revista)» (Vasques, 2009, p. 44).

O Teatro Nacional D. Maria II foi inaugurado em 1846, no vigésimo sétimo aniversário da Rainha D. Maria II que lhe deu o nome e depois de muita polémica sobre o local onde iria ser construído.

«Durante um longo período de tempo, o Teatro Nacional foi gerido por sociedades de artistas que, por concurso, se habilitavam à sua gestão. Após a implantação da República, passou a chamar-se Teatro Nacional de Almeida Garrett. A gestão mais duradoura foi a de Amélia Rey Colaço / Robles Monteiro, que permaneceu no teatro de 1929 a 1964, mas a mais célebre terá sido a da companhia Rosas e Brasão, entre 1881 e 1898, durante a qual foi ousada uma mudança de repertório (primeiras criações de peças de Shakespeare em Portugal)» (sítio Teatro Nacional D. Maria II, 2016).

Neste conjunto de espécies pudemos identificar, na sua maioria, o lugar onde os espetáculos foram representados:

Espaços teatrais	Quantidade espécies tratadas
Teatro do Ginásio	404
Teatro da Rua dos Condes	25
Teatro do Salitre (até 1858)	5
Teatro das Variedades (sucessor do Teatro do Salitre)	26
Teatro Nacional D. Maria II	11
Sem identificação do espaço	341

Tabela 1

Mais pudemos identificar que as espécies agora digitalizadas e disponibilizadas no catálogo *online* da Biblioteca da ESTC são na sua maioria traduções e imitações, embora também encontremos originais de autores portugueses.

Designação	Quantidade espécies tratadas
Traduções	cerca de 172
Imitações	cerca de 137
Originais	cerca de 245

Sem referência

cerca de 258

Tabela 2

Segundo Guilherme Filipe:

«[Existe] um número inequivocamente grande de obras originais, traduzidas e imitadas, tanto por dramaturgos de reconhecido mérito literário, como por autores que a história dramática não perpetuou, mas que, todavia, sustentaram a indústria teatral dos teatros públicos de Lisboa, Porto e restantes cidades de província, quer das companhias profissionais, quer dos grupos amadores que mimetizavam aquelas.» (Filipe, 2014, p. 13).

Estas peças abrangem uma linha temporal de escrita que vai de 1811 a 1899. O número de páginas por espécie também é muito variável, indo de 8 a 378 páginas.

### Tipologia teatral das espécies segundo designação dos autores

As 812 peças manuscritas objeto deste projeto, permitem, a partir do texto dramático, fazer um levantamento dos autores, atores, atos e tipologia do que se representava nesta época específica. Temos também a possibilidade de ler os pareceres censórios a que todos os espetáculos estavam sujeitos e que eram elaborados pela Inspeção-Geral dos Theatros e Espectáculos.

### Levantamento por géneros

Depois do tratamento documental das espécies verificamos que a maioria dos autores especifica o género dramático da peça depois do título. Assim, encontramos as seguintes referências:

Tipologia	Datas	Quantidade espécies tratadas
Brincadeira dramática	1836	1
Comédia	1813-1899	600
Comédia drama	1866,1866	2
Comédia revista	1861	1
Destempero	1860	1
Disparate	1859, 1866,1867	3
Drama	1811-1897	115
Entreacto	1856, 1856, 1864, 1867, 1867	5
Fantasia	1854, 1860	2

Farsa	1815-188?	21
Intervallo jocoso	1848	1
Intervallo lírico	1867	1
Improviso dramático	1841	1
Melodrama	[s.d.]	1
Mistério	1854	1
Ópera-cômica	1848-1869	8
Poesia	1866	1
Poesia cômica	1866	1
Propósito cômico	1870	1
Provérbio	1867	2
Quadro marítimo	188?	1
Revista	1851-1869	6
Scena cômica	1859	1
Scena de família	1864	1
Tragédia	1839-1857	6
Tolice	1859	1
Vaudeville	1854, 1860	2
Sem referência	1829-1883	25

Tabela 3

Com efeito, neste levantamento percebemos que as comédias, os dramas e as farsas são os géneros mais representados.

### Levantamento por atos

Nos dados apurados no levantamento do número médio de atos obtivemos a seguinte informação:

Atos	Quantidade espécies tratadas
1 ato	454
2 atos	95
3 atos	134
4 atos	37
5 atos	76

Sem identificação do número de atos	16
-------------------------------------	----

Tabela 4

Na contagem obtida podemos observar que as peças em 1 ato são as que estão em maior quantidade nas 812 espécies.

### Levantamento por atores

As anotações a lápis, ao lado do nome das personagens, com os nomes dos atores, permitiram-nos recuperar alguma da história do teatro desta época, ao proporcionar preencher algumas lacunas de informação. Neste levantamento obtivemos os seguintes dados:

	Datas	Quantidade
Atores	1811-1899	cerca de 194

Tabela 5

### Levantamento por autores

Este levantamento permitiu perceber que muitos dos autores destas peças de teatro são totalmente desconhecidos nas bases de dados existentes sobre teatro português e que são preenchidas com o contributo de trabalhos de investigação. Também se verificou que muitas das peças eram submetidas a concurso e por essa razão o autor não era identificado. Nesta contagem obtivemos os seguintes dados:

	Datas	Quantidade
Autores	1811-1899	cerca de 286

Tabela 6

Por conseguinte, este projeto de digitalização de 812 espécies que se encontram no Arquivo Histórico da Escola Superior de Teatro e Cinema permite, para além de tornar acessível o texto teatral, o acesso “à identificação da distribuição de intérpretes, ou ao estudo dos problemas de tradução e adaptação de obras estrangeiras. São múltiplas as perspetivas que se abrem para a organização deste espólio dramático, na recolha de informação para-teatral” (Filipe, 2014, p. 13).

### Metodologias no tratamento documental e digitalização

A Biblioteca da ESTC, enquanto unidade responsável pelo tratamento técnico, preservação e divulgação dos seus acervos documentais, histórico e contemporâneo, em permanente construção e atualização, decidiu apresentar uma candidatura ao Concurso de Recuperação, Tratamento e Organização de Acervos Documentais da Fundação Calouste Gulbenkian em fevereiro de 2013, que foi aprovada a 31 de maio do mesmo ano. O apoio da Fundação Calouste Gulbenkian permitiu a criação da “Biblioteca Digital” de parte do espólio das espécies manuscritas, por ser considerado o meio tecnológico mais eficaz para a preservação da documentação e também o meio que permite disponibilizar aos utilizadores a informação em livre acesso.

A metodologia adotada para o tratamento documental foi o das Normas Internacionais de Descrição Bibliográfica, completada por um conjunto de técnicas específicas de conservação e digitalização que recorrem às mais recentes tecnologias e equipamentos.

O projeto contemplou a captação em suporte digital de sessenta mil cento e quatro imagens e abarcou as espécies manuscritas dos séculos XVIII e XIX. O procedimento digital foi feito por uma empresa com competência para esta intervenção em arquivos de documentação histórica.

O trabalho foi desenvolvido em diferentes fases de concretização:

#### **Digitalização e introdução dos metadados**

- Os indicadores de avaliação foram efetuados de acordo com o número de imagens produzidas, com os requisitos técnicos e de qualidade definidos;
- As reproduções digitais constituíram uma cópia fiel do original;
- A organização das imagens digitalizadas foi feita pela ordem pela qual os documentos foram entregues;
- Na digitalização foi representada a totalidade do documento e não apenas a área útil, com a aplicação de uma margem de dois a cinco mm em torno do mesmo;
- As imagens foram fornecidas em TIFF (imagem de arquivo) e multi-PDF por cada peça de teatro, com a separação por atos;
- Foram atribuídos critérios iguais de orientação vertical, resolução de cor e área útil para cada lote de imagens;
- As imagens foram obtidas com uma resolução mínima de 300 dpi e com uma profundidade mínima de 24 bits de cor, independentemente do tamanho e características dos originais.

#### **Indexação (metacodificação) e edição digital**

- As imagens em TIFF foram nomeadas da seguinte forma:  
MAN\_A1 (nº do armário)\_P1 (nº da prateleira)\_0038 (nº da peça)\_001 (nº da página);
- As imagens em multi-PDF por peça foram nomeadas da seguinte forma:  
MAN\_A1 (nº do armário)\_P1 (nº da prateleira)\_0038 (nº da peça);

– Os metadados foram preenchidos nas matrizes digitais das imagens (ficheiros TIFF) de acordo com o quadro abaixo:

ImageWidth
ImageLenght
BitsPerSample
Compression
Photometric Interpretation
Document name
Make
Model
Samplesperpixel
Xresolution
Yresolution
Resolution unit
Software
DateTime
Artist
Copyright

**Tabela 7**

As imagens foram gravadas num disco externo e organizadas em pastas por Armário, por prateleira, pela cota do documento manuscrito e pela seguinte ordem:

Pasta 1 – imagens TIFF (conforme requisitos atrás indicados);

Pasta 2 – imagens em PDF (multi-PDF, por espécie), com marca de água da instituição;

Pasta 3 – imagens em PDF (multi-PDF, por peça), sem marca de água da instituição;

Pasta 4 – imagens em thumbnails;

Pasta 5 – imagem PDF da folha de rosto, com marca de água da instituição;

Pasta 6 – imagem PDF da folha de rosto, sem marca de água da instituição.

A informação bibliográfica das peças manuscritas foi introduzida no catálogo *online*, com o acesso, total ou parcial, ao documento digital, ficando disponíveis aos investigadores e à comunidade em geral que, cada vez mais procuram o Arquivo Histórico desta Biblioteca –

vejam-se, por exemplo, os dados patentes no Google Analytics (27420 visualizações) – para trabalhos de pesquisa na área da história do teatro e das artes cénicas em Portugal e no Brasil, assim como, outros países distribuídos por outros continentes.

Toda a informação anotada nas espécies foi transcrita para o campo de notas e sistematizada da seguinte forma:

- Alterações textuais manuscritas a lápis (ou a tinta);
- Contém cortes/acrescentos/substituição/ou...substituições) textuais;
- Contém cortes/acrescentos/substituição/ou...substituições) nas didascálias;
- Contém manuscritos e dactiloscritos (caso da miscelânea);
- Contém distribuição de personagens a lápis;
- Contém didascália;
- Contém implantação cenográfica;
- Contém fotografia a p&b da cenografia;
- Contém plano de trabalhos;
- Contém projeto de orçamentos;
- Contém partitura da canção "Título";
- Contém letras de canções manuscritas/ou dactiloscritas;
- Direitos de autor/representação: "Esta cópia licenciada pertence ao Snr....";
- Distinção gráfica entre didascália e corpo de texto;...ou;
- Distinção gráfica e de cor entre didascália e corpo de texto;...ou;
- Distinção gráfica de cor entre didascália e corpo de texto;
- Do repertório de... "Academia Instrutiva do Pessoal dos Caminhos de Ferro de Leste e Norte"; do Teatro da Trindade ; do Teatro da Rua dos Condes, etc. (quando conste), ou;
- Do repertório de...[carimbo];
- Exemplar dactiloscrito com folha de rosto manuscrita;
- Exemplar dactiloscrito;
- Exemplar do ponto;
- Exemplar do ponto...(caso tenha o nome);
- Folheação (ou paginação) descendente;
- Licença de representação para..., ou;

- ▶ Licença de representação encontra-se registada em L.º 2º nº 114 = n, Inspeção Geral dos Theatros;
- ▶ Parecer censório...[assinatura de/...ou assinatura ilegível];
- ▶ Parecer censório com/sem cortes;
- ▶ Pedido de licenciamento;
- ▶ Trad. de...;
- ▶ Título do texto original...;
- ▶ Contém marcação de cena;
- ▶ Contém informação....

## Resultados

O projeto permitiu, com o valioso apoio da Fundação Calouste Gulbenkian, que 812 peças de teatro manuscritas do Arquivo Histórico que chegaram até à atualidade, ficassem acessíveis em formato digital, fazendo renascer espécies que já tinham sido declaradas como inexistentes, possibilitando desta forma um novo impulso na investigação em teatro.

Por força do seu valor histórico, este conjunto de peças de teatro manuscrito, passou a integrar o Registo Nacional de Objetos Digitais (RNOD), gerido pela Biblioteca Nacional. O sistema RNOD é um agregador de conteúdos bibliográficos digitais e digitalizados sendo que estes conteúdos estão integrados no Portal da Europeana (Biblioteca digital que reúne a maior coleção *online* de arte, cultura e ciência). No sítio <http://rnod.bnportugal.pt/rnod>, podemos ler que: este sistema está principalmente vocacionado para conteúdos de acervos de bibliotecas ou de outras organizações detentoras de materiais bibliográficos, cobrindo espécies manuscritas e publicações impressas de variada tipologia (texto, imagem, som, música, cartografia e iconografia). Mais informa que: o RNOD propicia um ponto central de pesquisa e acesso a objetos digitais disponibilizados por diferentes entidades, utilizável por profissionais e público em geral, sendo que a disponibilização da informação permite evitar a duplicação de investimentos no que respeita à digitalização de publicações impressas. Possibilita, também, uma visão global das iniciativas de digitalização de coleções bibliográficas e afins, em Portugal, permitindo registar, quer informação sobre objetos digitalizados, quer intenções de digitalização com uma flexibilidade de participação, permitindo diferentes níveis de registo, quer para itens individuais quer para grupos ou coleções ainda não descritos individualmente.

Para concluir e citando Guilherme Filipe, utilizador e investigador da Biblioteca da ESTC: «...Entre os manuscritos do Arquivo Histórico da Escola Superior de Teatro e Cinema, três armários de Pandora, muitos inéditos partilham o destino esperançoso de *Lisboa em 1850*: salvados do incêndio, aguardam o resgate do esquecimento.» (Filipe, 2012, p. 63).



## Considerações finais

O projeto teve como finalidade salvaguardar e disponibilizar parte do espólio histórico do próprio Almeida Garrett, através do tratamento documental e da digitalização das peças de teatro manuscritas dos séculos XVIII e XIX, do Arquivo Histórico do Conservatório Real de Lisboa, que estão agora acessíveis: no catálogo *online* da Biblioteca da ESTC, no RNOD e nos conteúdos bibliográficos da Europeana (Biblioteca digital).

Na opinião do investigador, já anteriormente citado, Guilherme Filipe:

«O acervo histórico da Escola Superior de Teatro e Cinema, fruto de proveniências diversas, constitui um aliciente campo de trabalho de investigação sobre o teatro profissional e amador, desde a análise da escolha de repertórios /.../, à identificação da distribuição de intérpretes, ou ao estudo dos problemas de tradução e adaptação de obras estrangeiras. São múltiplas as perspectivas que se abrem para a organização deste espólio dramático, na recolha de informação para-teatral. Entre as obras ditas menores, de entretenimento comercial, tantas vezes criticadas pela crítica coeva e contemporânea, encontram-se inéditos até agora apenas referenciados na imprensa do tempo, de que se desconheciam os originais.» (Filipe, 2014, p. 13).

O projeto abre, também, uma nova linha de investigação a estudantes de mestrado e doutoramento, que têm agora acesso a documentos que poderão levar à descoberta de novas fontes de pesquisa e a completar a informação existente sobre este período da história do teatro.

## Referências bibliográficas

BIBLIOTECA NACIONAL DE PORTUGAL. Biblioteca digital. *Registo Nacional de Objectos Digitais* [Em linha]. Lisboa : BNP. [Consult. 13 set. 2016]. Disponível na Internet <URL: <http://rnod.bnportugal.pt/rnod/>>.

CASTRO, Augusto de (1954) – *Garrett e o teatro português: conferência*. Lisboa : Bertrand. 51 p.

FILIFE, Guilherme (2012) – Quando as revistas eram do ano. *Sinais de Cena*. ISSN 1646-0715 N° 18, p. 63.

FILIFE, Guilherme (2014) – Um arquivo histórico enquanto baú de memórias. *Politecnia*. ISSN 1645-006X. N° 28, p. 12-13.

MAGALHÃES, Paula (2007) – *Os dias alegres do Ginásio: memórias de um teatro de comédia*. Lisboa : Universidade de Lisboa, Faculdade de Letras. 166 p. Dissertação de mestrado.

PICCHIO, Luciana Stegagno (s.d.) – *História do teatro português*. Lisboa : Portugália Editora. 486 p.

TEATRO NACIONAL D. MARIA II. O teatro. *História* [Em linha]. Lisboa : TNDMII. [Consult. 13 set. 2016]. Disponível na Internet <URL: <http://www.teatro-dmaria.pt/pt/o-teatro/historia/>>.

VASQUES, Eugénia (2009) – *Espaços teatrais da Lisboa do barroco aos séculos XVIII e XIX*. 2ª ed. Amadora : Escola Superior de Teatro e Cinema. (Sebentas – Coleção História do Teatro Português, n° 1). 47 p.

---

## Repositórios de dados de pesquisa para as ciências da saúde

---

Lucas Paganine

Universidade de Brasília (UnB)

[lnpaganine@hotmail.com](mailto:lnpaganine@hotmail.com)

Michelli Costa

Universidade de Caxias do Sul (UCS) e Universidade de Brasília (UnB)

[michelli@unb.br](mailto:michelli@unb.br)

### Resumo

O estudo tem por objetivo elaborar um panorama geral dos repositórios de dados de pesquisa na área das ciências da saúde. Para tanto, foi realizada uma pesquisa descritiva, na qual foram analisados vinte repositórios de dados cadastrados e certificados no diretório re3data.org. A análise metodológica foi desenvolvida a partir das características essenciais apontadas por Rodrigues *et al.* (2010) para um repositório de dados, que foram sistematizadas em categorias. As categorias exploram aspectos acerca do armazenamento, descrição e apresentação dos itens. Os resultados do estudo indicam que os repositórios analisados se encontram em diferentes estados de desenvolvimento. No entanto, foi possível observar que não existe tendência para o uso de software de repositório específico, embora haja tendência estruturada para o padrão de metadados e os tipos de recursos disponibilizados. De forma geral, grande parte dos repositórios analisados possuem características de repositórios de acesso aberto, no entanto apenas 30% deles distribuem seus conteúdos acompanhados de licenças abertas. Percebe-se, portanto, uma incipiente e crescente preocupação com o desenvolvimento de sistemas dessa natureza.

**Palavras-chave:** Repositórios de dados de pesquisa. Ciências da saúde. Open access. Comunicação científica.

## Research data repositories for health sciences

### Abstract

The study aims to show an overview of research data repositories in the area of health sciences. To achieve it, a descriptive research was held, where twenty registered data and certificated repositories were analysed based in the directory re3data.org. The methodology was based in the essential data repositories characteristics mentioned by Rodrigues *et al.* (2010). The characteristics are related to storage, description and presentation of items. The study`s results indicate that repositories analysed has different stages of development. However, it was observed that there was no trend towards the use of specific repository software, although there is a tendency for the metadata standard and the types of available resources. In general, most repositories have open access characteristics, however only 30% of them distribute their contents based in open license. As conclusion, we can say that the results suggest an incipient and growing concern about the development of such systems.

**Keywords:** Repositories of research data. Health sciences. Open access. Scientific communication.

### Introdução

O advento das novas tecnologias da comunicação e informação provoca mudanças nas formas como as pessoas interagem e se comunicam. Conseqüentemente, as mudanças têm gerado processos significativos de transformação da comunicação científica. Nesse contexto de mudanças, surge o movimento Open Access, para apresentar uma nova alternativa ao sistema de mercado tradicional dos periódicos científicos comerciais. Dentre as iniciativas que formalizam e orientam o movimento Open Access, destaca-se aqui a declaração de Bethesda que evidenciou, ainda em 2003, a importância do acesso aberto para as ciências da saúde (SARMENTO E SOUZA *et al.*, 2005). A declaração de Bethesda formalizou a demanda para o tratamento dos dados de pesquisa no contexto do acesso aberto.

O compartilhamento de dados de pesquisa é uma questão de crescente importância devido aos seus diversos benefícios para otimização da ciência. Os repositórios de dados de pesquisa têm sido apresentados pela literatura pertinente como uma das ferramentas adequadas para tal objetivo. Porém ainda se apresentam como desafios a padronização dos sistemas e a curadoria dos dados envolvidos. Considerando o contexto brevemente levantado e a importância do tema, o presente estudo buscou identificar a situação atual de desenvolvimento dos repositórios de dados de pesquisa no campo das ciências da saúde.

## Metodologia

Para execução da análise proposta, foi realizada uma pesquisa de levantamento e análise descritiva acerca dos repositórios de dados na área das Ciências Saúde, registrados no re3data.org (Registry of Research Data Repositories) e que possuíam algum certificado de sistema. O levantamento dos dados ocorreu no último semestre do ano de 2015.

A partir dos critérios definidos para a pesquisa foram identificados 20 repositórios de dados de pesquisa no campo das Ciências da Saúde (Quadro 1). Dentre os repositórios que compõem o conjunto de análise 50% deles são originários dos Estados Unidos da América, 10% do Canadá e 30% de países da Europa. Apenas dois deles originam-se de países do Sul do mundo, um é de origem indiana e outro de provem de Gana.

**QUADRO 1: REPOSITÓRIOS SELECIONADOS PARA ANÁLISE**

Nome do repositório	URL
American Type Culture Collection	<a href="http://www.lgcstandards-atcc.org/">http://www.lgcstandards-atcc.org/</a>
ArrayExpress	<a href="http://www.ebi.ac.uk/arrayexpress/">http://www.ebi.ac.uk/arrayexpress/</a>
Bacterial Carbohydrate Structure DataBase	<a href="http://csdb.glycoscience.ru/bacterial/index.html">http://csdb.glycoscience.ru/bacterial/index.html</a>
Bii	<a href="http://isa-tools.org/">http://isa-tools.org/</a>
Canadian Epigenetics, Environment and Health Research Consortium Platform	<a href="http://www.epigenomes.ca/">http://www.epigenomes.ca/</a>
Centers for Disease Control and Prevention, Data & Statistics	<a href="http://www.cdc.gov/DataStatistics/">http://www.cdc.gov/DataStatistics/</a>
ClinicalTrials.gov	<a href="http://www.clinicaltrials.gov/">http://www.clinicaltrials.gov/</a>
Collaborative Psychiatric Epidemiology Surveys	<a href="http://www.icpsr.umich.edu/icpsrweb/CPES/">http://www.icpsr.umich.edu/icpsrweb/CPES/</a>
Danish Data Archive	<a href="https://www.sa.dk/en/services/danish-data-archive">https://www.sa.dk/en/services/danish-data-archive</a>
DiversityData.org	<a href="http://www.diversitydata.org/">http://www.diversitydata.org/</a>
Domino	<a href="http://mint.bio.uniroma2.it/domino/">http://mint.bio.uniroma2.it/domino/</a>
Hardin.MD	<a href="http://hardinmd.lib.uiowa.edu/">http://hardinmd.lib.uiowa.edu/</a>
HomoMINT	<a href="http://mint.bio.uniroma2.it/HomoMINT/Welcome.do">http://mint.bio.uniroma2.it/HomoMINT/Welcome.do</a>
Human Proteinpedia	<a href="http://www.humanproteinpedia.org/">http://www.humanproteinpedia.org/</a>
INDEPTH Data Repository	<a href="http://www.indepth-ishare.org/index.php/home">http://www.indepth-ishare.org/index.php/home</a>
InnateDB	<a href="http://www.innatedb.com/">http://www.innatedb.com/</a>
MaizeGDB	<a href="http://www.maizegdb.org/">http://www.maizegdb.org/</a>
Mentha	<a href="http://mentha.uniroma2.it/">http://mentha.uniroma2.it/</a>
NeuroMorpho	<a href="http://neuromorpho.org/">http://neuromorpho.org/</a>
Neuroscience Information Framework	<a href="http://neuinfo.org/">http://neuinfo.org/</a>

Fonte: Elaboração própria

A análise teve como referência as três características essenciais para um repositório de dados proposta por Rodrigues *et al.* (2010). As características forma utilizadas como categoria de análise dos dados coletados. Portanto, os resultados foram sistematizados em três grandes categorias: armazenamento, descrição e apresentação dos itens (Quadro 2).

QUADRO 2: CATEGORIAS E ITENS ANALISADOS

Categorias	Elementos analisados
Armazenamento	Software: programa(s) utilizados no repositório
	Sistema de preservação: métodos ou práticas de preservação utilizadas
	URL persistente: se o repositório apresenta URL persistente ou não
	Instituição do pesquisador: se o repositório faz referência a instituição à qual o pesquisador pertence ou não
	Licença: tipo de licença(s) utilizada
	Repositório aberto: se o repositório de dados é um repositório de acesso aberto ou não
Descrição	Metadados (padrão): qual padrão de metadados utilizado.
Apresentação dos itens	Tipo de recurso: caracterização dos dados apresentados no repositório, se dados audiovisuais, textuais, etc.
	Quantidade de itens: número de entradas de itens no repositório
	Relação do item com o conjunto: se há indicação da relação do dado com o todo
	Relação do conjunto com a publicação: se há indicação da relação do conjunto de dados com a publicação onde se encontra

Fonte: Elaboração própria

## Armazenamento dos dados de pesquisa em repositórios

A primeira categoria de análise tratou de aspectos referentes ao armazenamento dos dados de pesquisa e contemplou questões sobre o *software* dos repositórios, os sistemas de preservação utilizados, as licenças concedidas aos conteúdos armazenados e a política acerca do acesso aberto dos repositórios de dados de pesquisa.

### *Software dos repositórios de dados de pesquisa nas ciências da saúde*

O estudo identificou que existe uma ampla variedade de software utilizado no contexto e que nenhum deles apresenta-se como majoritário em relação aos outros. Dentre os sistemas selecionados para a análise identificou-se que cada repositório apresentou informação singular acerca do software utilizado, conforme indicado no Quadro 3.

**QUADRO 3:** SOFTWARE UTILIZADOS PELOS REPOSITÓRIOS DE DADOS DE PESQUISA NAS CIÊNCIAS DA SAÚDE

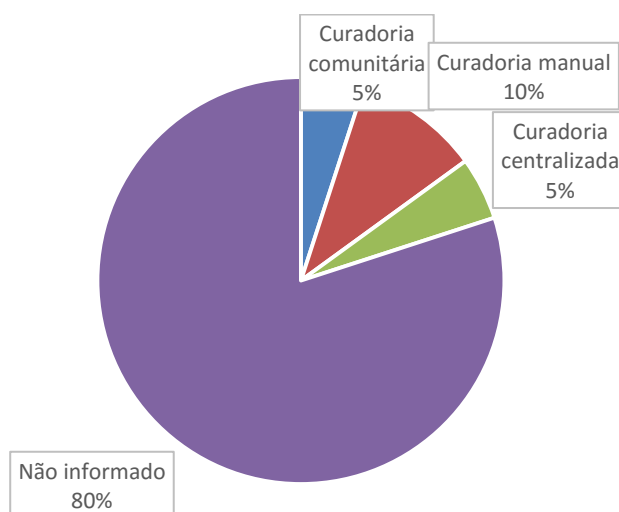
Repositório	Software
HomoMINT	MINT Viewer
Bacterial Carbohydrate Structure DataBase	SOAP
InnateDB	Cerebral program (Java), o ProbeLynx e o INVEX
Mentha	REST
Bii	BioInvIndex
Canadian Epigenetics, Environment and Health Research Consortium Platform	OpenCEMT

Fonte: Elaboração própria

O HomoMINT com seu software MINT viewer permite uma visualização gráfica no contexto de dados de alto rendimento. Já o software SOAP usado pelo Bacterial Carbohydrate Structure DataBase é uma base de dados relacional baseado no MySQL com uma tabela de aproximação para correlacionar os dados em estruturas. O software REST do Mentha é uma interface que oferece uma aplicação gráfica e ferramentas que representam interações e opções de caminhos de navegação. O BioInvIndex do Bii é uma aplicação web, modelo database persistente e pacote de serviços e web serviços. O OpenCent é um software de análise, manuseio e publicação de dados epigenéticos. O repositório InnateDB possui três programas para o desempenho de suas funcionalidades: Cerebral program, ProbeLynx e INVEX. O Cerebra program é um plugin Java para visualização de dados sobre moléculas. O INVEX é utilizado para análise e visualização de parâmetros e metadados dos dados. O ProbeLynx para anotações sobre microarranjos.

#### *Sistemas de preservação dos repositórios de dados de pesquisa nas ciências da saúde*

A maior parte dos repositórios analisados (80%) não declararam utilizar sistema de preservação. A falta de informação, no entanto, não permite concluir a ausência de estratégia para tal finalidade, uma vez que essa informação pode apenas ter sido negligenciada no momento do registro do repositório. Dentro os 20% dos repositórios que declararam possuir alguma estratégia de preservação dos conteúdos que armazenam, metade informou que realiza uma curadoria manual, na qual não conta com um sistema de preservação em rede com outras instituições ou iniciativas similares (Figura 1).

**FIGURA 1:** SISTEMA DE PRESERVAÇÃO DOS REPOSITÓRIOS DE DADOS DE PESQUISA NAS CIÊNCIAS DA SAÚDE

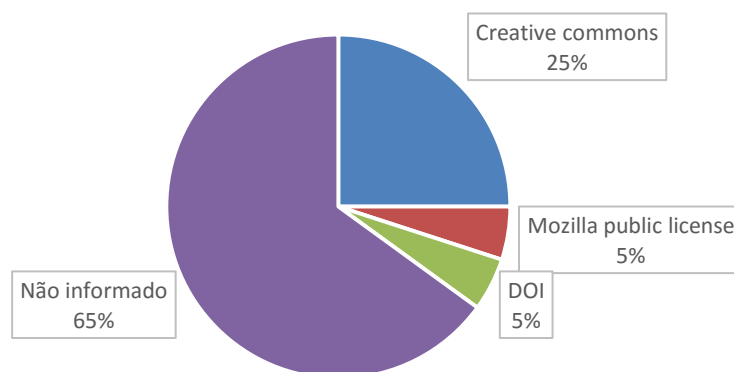
**Fonte:** Elaboração própria

A ausência de informações acerca do sistema de preservação dos repositórios não é suficiente para afirmar sua inexistência. No entanto, a falta de dados sobre o tema demonstra que ainda não há a preocupação devida com a preservação dos dados de pesquisa.

#### *Licenças de acesso e uso concedidas aos conteúdos armazenados nos repositórios de dados de pesquisa nas ciências da saúde*

Apenas seis dos vinte repositórios analisados informaram utilizar alguma licença de acesso e uso para o conteúdo que armazenam e disponibilizam. Dentre o conjunto de licenças previstas pelo re3data.org, a licença Creative Commons foi a mais utilizada, presente em um quarto da totalidade dos repositórios. Menos representativas, as licenças Mozilla Public License e a política DOI, foram utilizadas em pelo menos um dos repositórios selecionados (Figura 2).

FIGURA 2: LICENÇAS DE ACESSO E USO CONCEDIDAS AOS CONTEÚDOS ARMAZENADOS



Fonte: Elaboração própria

De forma geral foi observado que apenas 30% dos repositórios de dados de pesquisa nas ciências da saúde utilizam licenças adequadas para o armazenamento, disponibilização e reutilização dos dados no contexto da ciência colaborativa. Frente aos dados obtidos no Diretório é possível perceber que as preocupações acerca do uso de licenças abertas para a distribuição dos conteúdos ainda não alcançaram um nível desejável para os objetivos da ciência aberta.

#### *Política de acesso aberto dos repositórios de dados de pesquisa nas ciências da saúde*

Embora tenha sido considerada baixa a utilização das licenças para a ciência aberta, cerca de 85% dos repositórios declararam ser uma iniciativa de acesso aberto. Apenas 3 repositórios, o American Type Culture Collection, Collaborative Psychiatric Epidemiology Surveys e o Danish Data Archive não se declararam como repositórios de acesso aberto. O American Type Culture Collection disponibiliza tanto os dados quanto o próprio material biológico para compra, o Collaborative Psychiatric Epidemiology Surveys apresenta datasets tanto de uso público quanto de uso restrito e o Danish Data Archive possui alguns dados que só podem ser acessados fisicamente.

#### **Descrição dos dados de pesquisa em repositórios**

A segunda categoria de análise abordou os padrões de metadados para a descrição dos conteúdos disponíveis nos repositórios de dados de pesquisa em ciências da saúde.

#### *Metadados dos repositórios de dados de pesquisa nas ciências da saúde*

Dentre o conjunto analisado, apenas 8 repositórios apresentaram informações acerca do padrão de metadados utilizado para descrever seus recursos. Em alguns casos notou-se



que um mesmo repositório utiliza mais de um padrão de metadados para a finalidade da descrição. Portanto, foram identificados mais padrões de metadados do que repositórios que apresentavam essa informação. No total foram elencados 9 padrões de metadados, conforme elencado no Quadro 4.

**QUADRO 4:** PADRÕES DE METADADOS UTILIZADOS PELOS REPOSITÓRIOS DE DADOS DE PESQUISA

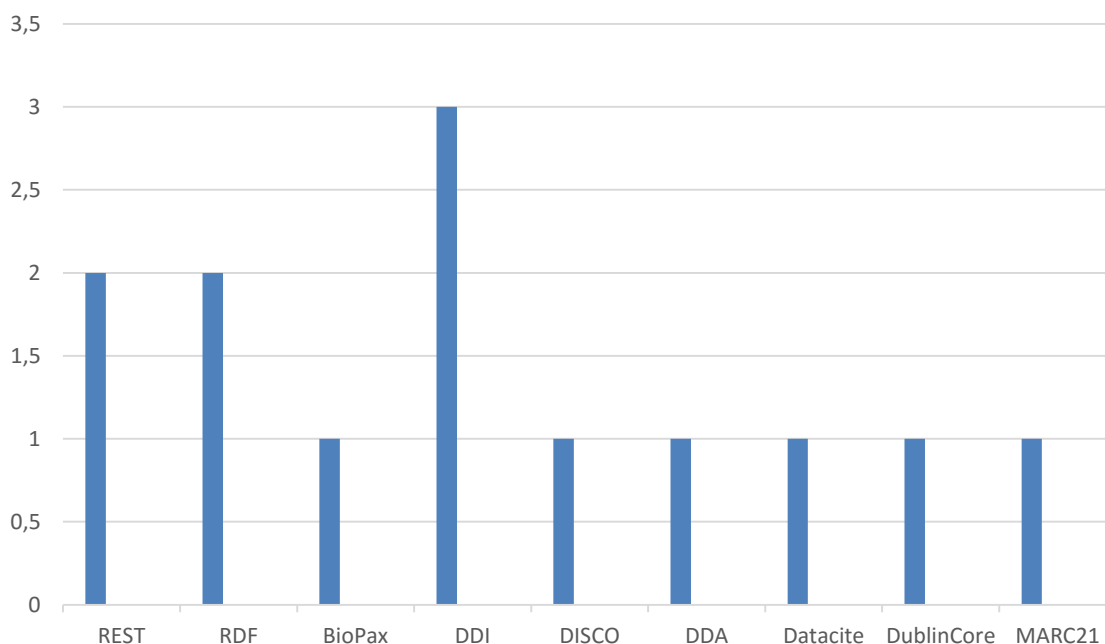
Repositório	Padrão de metadados
ArrayExpress	REST
ClinicalTrials.gov	
Bacterial Carbohydrate Structure DataBase	RDF
INDEPTH Data Repository	DDI, RDF
Danish Data Archive	DDI, DDA, Datacite
Collaborative Psychiatric Epidemiology Surveys	DDI, Dublin Core, MARC21
InnateDB	BioPax
Neuroscience Information Framework	DISCO

Fonte: Elaboração própria

De forma geral, foi identificado que o padrão de metadados mais utilizado pelos repositórios é o Data Documentation Initiative (DDI). O padrão foi desenvolvido pela DDI Alliance para descrição de recursos informacionais relacionados a pesquisas socioeconômicas, censos e outras coleções de microdata. Seu objetivo é ser facilmente compreendido por humanos e computadores, de forma a ter seu uso ampliado (KRAMER; LEAHEY, 2012).

Outros padrões que se destacaram no conjunto analisado foram o RDF e o REST (Figura 1). O RDF é um padrão voltado para facilitar a junção de dados criando modelos simples, já o padrão REST de metadados foca nas interações e restrições destacando dados significantes.

FIGURA 3: FREQUENCIA DE USO DOS PADRÕES DE METADADOS



Fonte: Elaboração própria

Além dos três formatos já descritos, mostraram-se relevantes para o contexto dos repositórios de dados de pesquisa em ciências da saúde os seguintes padrões de metadados: BioPax, DISCO, DDA, Datacite, DublinCore e Marc 21, que serão descritos brevemente. O BioPax faz parte de um projeto aberto e colaborativo e é uma linguagem padronizada que visa possibilitar a interação, troca, visualização e análise de dados biológicos. Já DISCO é um método de integração de informação baseado no protocolo de descoberta da Microsoft para facilitar interoperação entre recursos da internet. O DDA é o padrão próprio do Danish Data Archive. O Datacite é o esquema padrão da organização Datacite para publicação e citação de dados de pesquisa. O DublinCore é um esquema de metadados que objetiva a descrição de objetos digitais a partir de 15 elementos, que podem ser expandidos por seus especificadores. Por fim o MARC21, ou *machine readable cataloging* 21 é uma linguagem que permite a leitura e processamento por máquinas de registros catalográficos, ele se propõe a ser utilizado como formato padrão para troca de registros tanto bibliográficos quanto catalográficos.

### Apresentação dos itens nos repositórios

A última categoria de análise sistematizou informações referentes aos tipos de recursos disponíveis nos repositórios, o volume de itens e a relação dos itens como seus conjuntos de dados.

*Tipos de recurso disponíveis nos repositórios de dados de pesquisa nas ciências da saúde*

De forma geral foram identificados sete tipos de recursos disponíveis nos repositórios analisados. Os tipos em questão correspondem à classificação previamente estabelecida pelo próprio diretório re3data.org. A seguir é apresentada a relação entre os tipos de recursos e os repositórios onde eles estão disponíveis (Quadro 5).

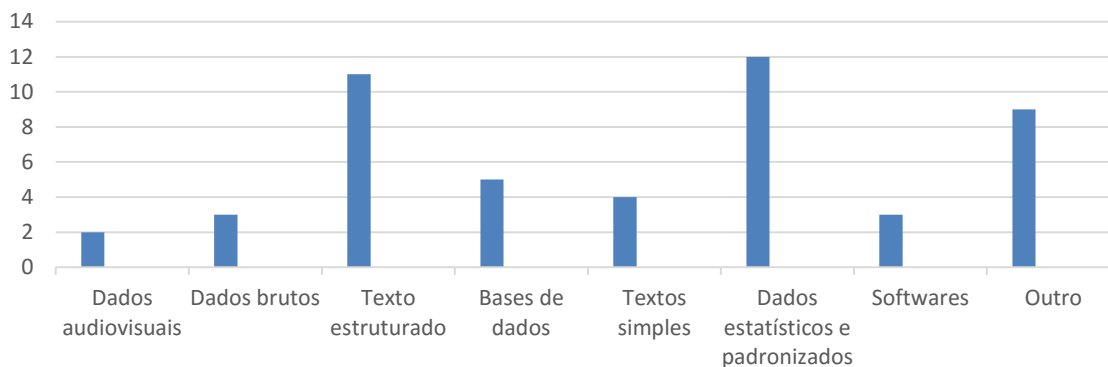
**QUADRO 5:** TIPOS DE RECURSOS DISPONÍVEIS NOS REPOSITÓRIOS DE DADOS DE PESQUISA EM CIÊNCIAS DA SAÚDE

Tipo de recurso	Repositório de dados de pesquisa
Dados audiovisuais	HomoMINT e NeuroMorpho
Dados brutos	HomoMINT, American Type Culture Collection e DiversityData.org
Texto estruturado	HomoMINT, MaizeGDB, Bacterial Carbohydrate Structure DataBase, Mentha, Human Proteinpedia, Bii, Canadian Epigenetics, Environment and Health Research Consortium Platform, ClinicalTrials.gov, Domino, NeuroMorpho e Neuroscience Information Framework
Bases de dados	American Type Culture Collection, ArrayExpress, INDEPTH Data Repository, Hardin.MD e Neuroscience Information Framework
Textos simples	MaizeGDB, InnateDB, Mentha, e Human Proteinpedia
Dados estatísticos e padronizados	MaizeGDB, ArrayExpress, Bacterial Carbohydrate Structure DataBase, InnateDB, Mentha, INDEPTH Data Repository, Bii, Canadian Epigenetics, Environment and Health Research Consortium Platform, Centers for Disease Control and Prevention, Data & Statistics, Collaborative Psychiatric Epidemiology Surveys, Danish Data Archive e Neuroscience Information Framework
Softwares	InnateDB, Mentha e Neuroscience Information Framework
Outro	MaizeGDB, InnateDB, INDEPTH Data Repository, Human Proteinpedia, Hardin.MD, ClinicalTrials.gov, Collaborative Psychiatric Epidemiology Surveys, DiversityData.org e Domino

Fonte: Elaboração própria

Em relação aos tipos de recursos disponibilizados pelos repositórios foi possível perceber a predominância dos dados estatísticos e padronizados (Figura 4). Este tipo de recurso engloba dados numéricos passíveis ou resultantes de análises estatísticas e dados numéricos estruturados em padrões específicos. Os dados textuais também padronizados em padrões específicos apareceram como o segundo tipo com maior ocorrência dentro do conjunto analisado. Estes resultados apontam que os tipos de recursos pertinentes para as ciências da saúde demandam padrões específicos não previstos pelo diretório re3data.org.

FIGURA 4: FREQUÊNCIA DOS TIPOS DE RECURSOS DISPONIBILIZADOS PELOS REPOSITÓRIOS DE DADOS DE PESQUISA



Fonte: Elaboração própria

Além dos dados estatísticos, padronizados e os textos estruturados, outros tipos de recursos foram identificados com menor relevância para o conjunto de repositórios, são eles: os dados audiovisuais, os dados brutos, as bases de dados, os textos simples e os softwares.

#### *Volume de itens nos repositórios de dados de pesquisa nas ciências da saúde*

A união de todos os recursos disponíveis nos repositórios analisados e que possuam essa informação somam cerca de 840 bilhões de itens. Com relação à quantidade de itens observou-se grande disparidade entre os repositórios, conforme demonstrado no Quadro 6.

QUADRO 6: QUANTIDADE DE ITENS POR REPOSITÓRIO ANALISADO

Nome do repositório	Quantidade de itens
Neuroscience Information Framework	829.679.866
Human Proteinpedia	7.038.972
ArrayExpress	1.962.932
InnateDB	888.753
Mentha	700.745
HomoMINT	343.277
ClinicalTrials.gov	206.902

American Type Culture Collection	74.000
Bacterial Carbohydrate Structure DataBase	21.904
Domino	15.981
Collaborative Psychiatric Epidemiology Surveys	6.274
INDEPTH Data Repository	4.239
Bii	133

Fonte: Elaboração própria

Os repositórios MaizeGDB, Hardin.MD, Canadian Epigenetics, Environment and Health Research Consortium Platform, Centers for Disease Control and Prevention, Data & Statistics, Danish Data Archive, DiversityData.org e o NeuroMorpho não informam o número de itens em seus bancos de dados.

A distância tão discrepante no número de itens entre os repositórios deve-se, em parte, ao escopo e as políticas adotadas. Enquanto alguns consideram um conjunto de dados completo como um único item, outros tratam cada molécula e cada interação como um item diferente, como nos casos dos Neuroscience Information Framework e da Human Proteinpedia. Considerando as peculiaridades ressaltadas, o Neuroscience Information Framework, é responsável por 98% do universo, mostrando-se assim como o repositório mais expressivo do universo analisado.

#### *Relação dos dados com seu conjunto*

A maioria dos repositórios (70%) apresentam informações sobre a relação do item com o seu conjunto (dataset). De acordo com Padilla Navarro *et al.* (2013) o conjunto de dados é uma representação sistemática parcial do objeto que está sendo pesquisado. Além disto, Rodrigues *et al.* (2010) ressaltam que eles são caracterizados por reunir informações ou fatos relacionados entre si e registrados num formato comum. Portanto, a existência de conjunto de dados no repositório demanda que a visão sobre objetos individuais seja extrapolada e que o conjunto receba tratamento diferenciado em casos específicos (RODRIGUES *et. al.*, 2010).

Foram sete os repositórios que não apresentam informações sobre a relação dos itens com suas coleções, a saber: American Type Culture Collection, Mentha, Hardin.MD, Bii, Canadian Epigenetics, Environment and Health Research Consortium Platform e o Domino. Destaca-se que nos casos onde não foi possível identificar essa relação não é possível concluir que a relação é inexistente, ela apenas pode não ter sido devidamente informada no registro do repositório.

#### **Considerações finais**

Os resultados encontrados demonstram que os repositórios de dados de pesquisa na área das ciências da saúde encontram-se em diversos estados de desenvolvimento. De um lado, foram observados repositórios em estados avançados como o Bacterial Carbohydrate Structure DataBase, que apresentou interface bem trabalhada e estruturação na descrição e apresentação de seus itens. De outro lado, foram identificados repositórios menos desenvolvidos, como o Danish Data Archive, que apresenta problemas sérios de navegação, impedindo muitas vezes a coleta de dados para análise proposta. No entanto, o volume dos itens nos repositórios indica uma crescente preocupação com a disponibilização dos dados de pesquisa em repositórios.

Ressalta-se ainda que a pesquisa se limitou aos repositórios de dados de pesquisa classificados como da área das ciências da saúde, que possuíam algum certificado de sistema e que estavam cadastrados no diretório re3data.org. Além disto, as fontes de informações consideradas para a análise limitaram-se às informações disponibilizadas pelos próprios portais dos repositórios e aquelas disponíveis no diretório re3data.org. Apesar das limitações declaradas sobre a pesquisa, considera-se que o retrato apresentado sobre o cenário tem condições de instrumentalizar as discussões sobre o compartilhamento e o reuso dos dados de pesquisa para a promoção da ciência aberta e produzir orientações para o aprimoramento dos sistemas, em especial na área das ciências da saúde.

## Referências bibliográficas

DDI ALLIANCE. Document, discover, and interoperate. Disponível em: <<http://www.ddialliance.org/>>. Acesso em: 25 de abr. de 2016.

DECLARAÇÃO DE BETHESDA. Meeting on Open Access Publishing, Bethesda. Abril. 2003. Disponível em: <<http://legacy.earlham.edu/~peters/fos/bethesda.htm>>. Acesso em: 03 de fev. de 2016.

KRAMER, S.; LEAHEY, A. Using RDF to describe and link social science data to related resources on the Web: leveraging the Data Documentation Initiative (DDI) model – Parade@Portsmouth. . In: DATA DOCUMENTATION INITIATIVE. Ann Arbor, Michigan: 2012 Disponível em: <<http://eprints.port.ac.uk/9029/>>. Acesso em: 12 maio. 2016

OECD. OECD Principles and Guidelines for Access to Research Data from Public Funding, 2007. Disponível em: <<http://www.oecd.org/dataoecd/9/61/38500813.pdf>>

PADILLA NAVARRO, P. A. *et al.* Acceso a datos de investigación e información científica en Chile. Revista española de Documentación Científica, v. 36, n. 3, 2013.

PANTON PRINCIPLES, Principles for Open Data in Science. About. Disponível em: <<http://pantonprinciples.org/>>. Acesso em: 22 de dez. de 2015.

REGISTRY OF RESEARCH DATA REPOSITORIES, RE3DATA. About. Disponível em: <<http://service.re3data.org/about>>. Acesso em: 22 de dez. de 2015.

RODRIGUES, E. *et al.* Os repositórios de dados científicos: estado da arte. [s.l: s.n.]. Disponível em: <<http://repositorium.sdum.uminho.pt/handle/1822/10830>>. Acesso em: 17 abr. 2015.

SARMENTO E SOUZA, M. F. *et al.* Algumas considerações sobre as principais declarações que suportam o movimento Acesso Livre. 2005 Disponível em: <<http://eprints.rclis.org/8512/>>. Acesso em: 30 mar. 2016

W3C. Resource Description Framework (RDF). Disponível em: <<https://www.w3.org/RDF/>>. Acesso em: 25 de abr. de 2016.

---

## Os repositórios das instituições de ensino superior portuguesas : Estudo comparativo

---

**Maria Eduarda Pereira Rodrigues**

Instituto Politécnico de Castelo Branco– ESACB/ESART – CERNAS–IPCB

[erodrigues@ipcb.pt](mailto:erodrigues@ipcb.pt)

**Maria João Amante**

ISCTE – Instituto Universitário de Lisboa

[maria.amante@iscte.pt](mailto:maria.amante@iscte.pt)

**Clarisse Pais**

Instituto Politécnico de Bragança

[clarisse@ipb.pt](mailto:clarisse@ipb.pt)

**Susana Lopes**

ISCTE – Instituto Universitário de Lisboa

[susana.lopes@iscte.pt](mailto:susana.lopes@iscte.pt)

**Teresa Segurado**

ISCTE – Instituto Universitário de Lisboa

[teresa.segurado@iscte.pt](mailto:teresa.segurado@iscte.pt)

**António Moitinho Rodrigues**

Instituto Politécnico de Castelo Branco–ESACB – CERNAS–IPCB

[amrodrig@ipcb.pt](mailto:amrodrig@ipcb.pt)

### **Resumo**

O Movimento de Acesso Aberto ao Conhecimento provocou alterações no modelo de difusão e publicação de conhecimento científico, sendo os repositórios institucionais de acesso aberto os sistemas de informação que, ao nível das instituições de ensino superior, melhor concretizam este objetivo. O presente estudo pretende contribuir



para o conhecimento do modo de funcionamento dos repositórios das instituições de ensino superior portuguesas presentes no Portal RCAAP<sup>i</sup>. O estudo foi realizado mediante aplicação de um inquérito por questionário aos gestores dos repositórios. Verificou-se que os bibliotecários desempenham um papel preponderante na organização e funcionamento dos Repositórios Institucionais. A maioria possui política de depósito de documentos e, destas, a maioria é mandatória. 80% dos repositórios não possuem controlo sobre o depósito. De um modo geral o incentivo ao depósito está relacionado com a avaliação dos docentes. 60% consideram fundamental ter política de depósito de documentos e 80 % consideraram que incrementaria o depósito. Apenas 25% dos Repositórios possuem plano de preservação digital publicado.

### **Abstract**

The Open Access movement has changed the scientific diffusion and publication. The repositories materialized that change. Present study aims to contribute for better knowledge about the informational structure and functioning of Portuguese Institutional Repositories (RI) listed in Portal RCAAP. The survey was done by the application of a questionnaire to the repositories administrators. The results show that librarians develop a very important role in the repositories, being in most cases its administrators. Most of the RI possesses archival policy, in most cases mandatory but in 80% of the RI it is not enough to ensure the deposit of documents. The encouragement to archive mostly is done through the staff evaluation system. 60% of the inquiries consider that having a deposit policy is fundamental to increase the repository. Only 25% of the RI possess digital preservation plan.

### **Introdução**

Nos últimos anos vem-se assistindo a uma pressão cada vez maior no sentido da disponibilização em acesso aberto da totalidade da literatura científica. Uma das facetas mais visíveis dessa pressão está relacionada com a criação do Movimento de Acesso Aberto. Este movimento preconiza o acesso à literatura científica, em texto integral, de forma gratuita e livre de quaisquer constrangimentos (Cullen, Chawner, 2011). Este existe há cerca de 25 anos e, a partir de 2009, passou a celebrar-se anualmente em outubro, a Semana Internacional do Acesso Aberto ao Conhecimento (Lawton, 2016). O mesmo autor considera que esta semana constitui uma oportunidade ímpar para divulgar o valor do acesso aberto no domínio da literatura produzida em meio científico.

As instituições de ensino superior acompanharam toda esta dinâmica tendo ocorrido, nas últimas décadas progressos significativos no domínio da aceitação do acesso livre ao conhecimento que se traduziram na criação de diversos repositórios digitais (Saraiva e Rodrigues, 2010) na sequência do lançamento, pelo MIT, do DSpace em 2002 (Covey, 2011). O mesmo autor refere que, acompanhando o Movimento de Acesso Aberto, os repositórios responderam à mudança que se começava a operar em termos de comunicação em meio científico. Também Hanh e Wyatt (2014) referem que é inquestionável o facto de os

repositórios estarem a crescer, em número, em todo o mundo. Nesse sentido, os repositórios surgiram como sistemas de informação que recebem dados, metadados e documentos que, em termos genéricos, procuram refletir a dinâmica de produção científica das respetivas organizações promovendo a sua difusão em modo de acesso aberto (Rodrigues e Rodrigues, 2014) assumindo-se, desde logo, como via alternativa de difusão e publicação de literatura científica. São ainda inclusivos aceitarem documentos de tipologias diversas (Cassela, 2010) e em formatos muito diferenciados.

Alguns autores consideram os Repositórios Institucionais (RI) como um dos componentes mais importantes das bibliotecas digitais (Cassela, 2010). Estes têm, geralmente, objetivos concretos comuns bem definidos e identificados com a estratégia organizacional. Procuram agregar todos os *outputs* científicos das suas instituições e procuram fazê-lo em acesso aberto promovendo o acesso fácil e global aos resultados da investigação fundamental ou aplicada (Saraiva e Rodrigues, 2010). Para cumprirem esse desiderato, na proporção o mais próximo possível dos 100%, socorrem-se de várias estratégias, utilizando quer os instrumentos institucionais disponíveis numa perspetiva formal, tais como os objetivos conjunturais e estruturais das instituições, quer outros mecanismos de motivação como atribuição de recompensas, demonstrações, formação desenvolvida especificamente, eventos científicos, entre outros (Zuccala, Oppenheim e Dhienza, 2008) tentando aproximar o número de *outputs* científicos depositados anualmente nos RI do número de *outputs* científicos efetivamente produzidos e publicados anualmente por cada instituição/organização. A este propósito também Hahn e Wyatt (2014) referem que o contributo voluntário dos investigadores para os repositórios é ainda bastante fraco, conquanto considerem que varia de disciplina para disciplina.

A literatura publicada sobre esta temática revela que a missão dos repositórios não é fácil e em muitas instituições/organizações os resultados apresentados indicam que o número de *outputs* científicos produzidos e publicados no âmbito de cada instituição para um determinado período está ainda muito afastado do número de *outputs* depositados nos repositórios para o mesmo período. Xia (2008) indica mesmo a relutância de alguns autores em disponibilizarem a sua produção científica para o repositório. Por outro lado, estudos diversos atestam a importância da existência de documentos de referência no que toca à execução e grau de sucesso dos objetivos delineados para os RI destacando-se as políticas de depósito de documentos, incluindo políticas mandatárias e mandatos (Xia, 2008 ; Cassela, 2010 ; Chant, 2016). Consideram ainda outros referenciais tais como mecanismos de avaliação de docentes e investigadores, ou outros instrumentos supra-institucionais que contribuam para aumentar os níveis de depósito de documentos nos RI. Também Lyons e Booth (2011) verificaram que as taxas de implementação e arquivo de documentos nos repositórios são muito variáveis e a aceitação dos mesmos pela respetiva comunidade por vezes é muito lenta. Aliás Zuccala, Oppenheim e Dhienza (2008) referem a necessidade imperiosa de coordenador institucionalmente esforços de modo a obter sucesso com a

estratégia de implementação do repositório. Na mesma linha Cassela e Morando (2011) defendem a importância das competências do *staff* do repositório como um fator chave no sucesso do mesmo.

Segundo Saraiva e Rodrigues (2010) o desenvolvimento e incremento dos repositórios portugueses varia de instituição para instituição. Os mesmos autores associam o sucesso de repositórios como o RepositoriUM da Universidade do Minho com as políticas seguidas na instituição relativamente a esta faceta da gestão da produção de literatura científica. Por outro lado Zhang, Book e Wirth (2015) também referem que sem a mediação dos bibliotecários os investigadores têm dificuldade em arquivar a sua produção científica nos repositórios.

Considerando as diversas abordagens, muitos estudos neste domínio são realizados com a finalidade de compreender quais os mecanismos que mais contribuem para a motivação dos investigadores no sentido de estes procederem ao arquivo da sua produção científica ou de a disponibilizarem para ser arquivada através de depósito mediado.

### **Âmbito e objetivos**

Constituiu objetivo deste trabalho verificar a existência de referenciais para o depósito de documentos nos RI portugueses, nomeadamente políticas de depósito de documentos, e de que forma é que estes influenciam o depósito da produção científica das instituições nos RI, bem como identificar o grau de interoperabilidade do repositório com outros sistemas. O seu âmbito foram os repositórios das Instituições de Ensino Superior (IES) portuguesas presentes no Portal do Repositório Científico de Acesso Aberto de Portugal (RCAAP), num total de 38 repositórios à data de distribuição do questionário (maio de 2016).

### **Material e métodos**

O instrumento de recolha de dados utilizado para a realização do presente estudo foi o inquérito por questionário, modelado com recurso ao *Google docs*. O inquérito foi distribuído *online* a todos os administradores/gestores dos repositórios das IES portuguesas.

A estrutura do questionário permitiu obter dados relativamente a:

- a) tipo de IES, dimensão em estudantes e docentes/investigadores;
- b) data de criação, dimensão, crescimento anual efetivo, dados de utilização, composição da equipa do repositório;
- c) políticas do repositório; quem propôs e aprovou a política, se a política é mandatária, se está disponível publicamente, cumprimento da política, verificação do cumprimento da política, interação do repositório com outros sistemas de informação internos (nomeadamente, o sistema de avaliação de desempenho docente), interoperabilidade, efeitos sobre o depósito, importância da existência de política no repositório;
- d) política de preservação digital.

A análise dos resultados foi efetuada com recurso ao Programa Estatístico SPSS. Foram efetuadas contagens, calculou-se a média e o desvio padrão.

## Resultados

Obteve-se um total de 17 respostas, sendo 10 de universidades e institutos universitários (U+IU) e 7 de institutos politécnicos (IP), correspondendo a 44,7% do universo dos inquiridos. Não se obtiveram respostas de nenhuma das Escolas Não Integradas (Escolas de ensino superior não integradas em Universidades ou Institutos Politécnicos).

No momento de aplicação do questionário as datas de criação dos RI variavam entre 2003, a mais antiga, e 2014, a mais recente sendo que o RI mais antigo tinha 12,5 anos e o mais recente tinha 2,6 anos de idade. A idade média dos RI das U+IU (8,4 anos) é superior à idade média dos repositórios dos IP (7,0 anos), o que se reflete na sua dimensão em termos de documentos depositados com um total de 121.146 documentos depositados nos RI das U+IU e 29.960 documentos depositados nos IP. Determinou-se também uma relação positiva não significativa ( $R=0,436$ ;  $P>0,05$ ) entre a idade do RI e o número de documentos depositados. Verificou-se, também que o número médio de documentos depositados por ano nas U+IU foi de 1406,0 e nos IP foi de 786,9.

Verificou-se que 82,4% dos administradores/gestores dos RI são bibliotecários. Parece assim poder afirmar-se que os bibliotecários desempenham um papel muito importante em termos de gestão dos RI.

Relativamente ao número médio de documentos depositados verificou-se que os RI das U+IU possuem, em média, 12.114,9 documentos e os RI dos IP possuem, em média, 4.280,0 documentos (Tabela 1). No que diz respeito ao número de utilizadores com permissão para arquivar e ao número de utilizadores registados as U+IU apresentaram valores muito mais elevados do que os IP o que parece denotar um maior interesse pelo arquivo no repositório destes docentes relativamente aos dos IP.

	Tipo instituição	Média	Mínimo	Máximo
Idade do repositório (anos)	IP	7,0	2,6	10,4
	U+IU	8,4	5,4	12,5
	Média total	7,8	2,6	12,5
Documentos depositados	IP	4.280,0	1.102	11.442
	U+IU	12.114,9	3.458	39.296
	Média total	8.888,8	1.102	39.296
Utilizadores com permissão para arquivar	IP	161,7	2	429
	U+IU	310,6	10	863
	Média total	245,4	2	863
Utilizadores registados	IP	367,3	30	965
	U+IU	3.346,6	222	13.576
	Média total	2.119,8	30	13.576

**Tabela 1.** Idade do RI, número de documentos depositados e número de utilizadores por tipo de instituição

Como seria de esperar, determinou-se uma correlação positiva elevada entre o número de documentos depositados e o número de utilizadores registados no RI ( $r=0,943$ ;  $P<0,01$ ), reforçando a ideia de que quanto mais utilizadores estiverem registados no RI maior será o número de documentos depositados no mesmo. Tendência idêntica ( $r=0,583$ ;  $P<0,05$ ) foi identificada quando se relacionou o número de documentos depositados e o número total de docentes/investigadores da instituição de ensino superior.

Não se identificou uma tendência evidente entre o número de documentos depositados e o número de *downloads* ( $r=0,458$ ;  $P>0,05$ ) embora tenha sido identificada uma correlação positiva elevada entre o número de documentos depositados e o número de consultas ( $R=0,941$ ;  $P<0,01$ ). Na tabela 2 apresentam-se os números médios de *downloads* e de consultas observados nos IP e nas U+IU analisados neste trabalho. Considerando os valores totais, verificamos que as U+IU apresentam valores médios muito mais elevados de *downloads* (6.532.844,4) e de consultas (3.502.867,4). No entanto, ao analisarmos o número de *downloads* por ano de idade do RI, verificamos que os IP apresentam um valor médio mais elevado (750.964,2 *downloads/ano*) (Tabela 2).

		Média	Mínimo	Máximo
Total de <i>downloads</i>	IP	4.267.001,3	185.119,0	14.189.045,0
	U+IU	6.532.844,4	12.125,0	15.620.718,0
	Média total	5.626.507,2	12.125,0	15.620.718,0
Total de <i>downloads por ano</i>	IP	750.964,2	17.873,3	2.598.060,3
	U+IU	725.348,2	1.449,9	1.978.250,9
	Média total	735.594,6	1.449,9	2.598.060,3
Total de consultas	IP	1.296.783,8	369.252,0	3.600.688,0
	U+IU	3.502.867,4	539.636,0	11.424.690,0
	Média total	2.620.434,0	369.252,0	11.424.690,0
Total de consultas por ano	IP	226.241,6	35.651,4	574.902,2
	U+IU	368.426,6	64.528,4	940.403,2
	Média total	311.552,6	35.651,4	940.403,2

**Tabela 2.** Número total de *downloads* e consultas

Verifica-se que 82,4% ( $n=14$ ) dos RI possuem Política de Depósito de Documentos (PDD) e, num número significativo destes ( $P<0,05$ ), a PDD é mandatória (64,3%). A PDD mais antiga data de 2005 e a mais recente de 2016. A totalidade das PDD estão disponíveis publicamente, 71,4% das quais nas páginas *web* dos RI e 28,6% nas páginas *web* institucionais. Verifica-se que 92,9% das PDD aplicam-se a todas as publicações. No entanto em 4 casos as respostas indicam que está prevista a implementação de mandatos, sendo que em 1 caso o mandato é para todas as publicações e em outros 2 respeita apenas a teses

e dissertações. Neste ponto e relativamente às políticas de depósito de documento importa referir que em 9 dos 14 RI foi o bibliotecário quem propôs a elaboração e aprovação da política. Tal vem na linha da importância reconhecida ao bibliotecário neste contexto e já referida anteriormente.

Relativamente ao crescimento anual total dos repositórios verificou-se que em 5 U+IU o número total de documentos depositados anualmente é superior a 1000, enquanto nos IP este valor apenas é atingido em 2 repositórios, sendo que destes dois RI, apenas 1 possui política de depósito de documentos aprovada. Aliás não se encontrou uma relação ( $R=0,035$ ;  $P>0,05$ ) entre a existência de política mandatória e o número de documentos depositados no repositório.

Dos RI que têm PDD, apenas 28,6% referiram possuir mecanismos de verificação de cumprimento da política de depósito de documentos pelo RI. Apesar destes valores 78,6% ( $n=11$ ) dos respondentes referem a existência de incentivos institucionais ao depósito de documentos indicando, a maioria, o sistema de avaliação dos docentes/investigadores. No entanto, em apenas 35,7% ( $n=14$ ) das instituições o sistema de avaliação dos docentes/investigadores está ligado ao RI.

Dos resultados obtidos constata-se ainda que 42,9% dos RI se articulam com sistemas internos de gestão da informação científica. Verifica-se, igualmente, que a totalidade dos respondentes considera a interoperabilidade vantajosa para o incremento do RI. No entanto os dados obtidos não permitiram verificar se esta articulação produz efeitos em termos de crescimento do RI. Em todo o caso, uma percentagem elevada de respondentes (82,4%) referiu que a PDD pode contribuir para o incremento do depósito no RI mas apenas 57,1% dos RI que possuem PDD indicaram que esta contribuiu efetivamente para o aumento do depósito de documentos.

Do total de respondentes apenas 35,3% refere que o RI tem um plano de preservação digital.

## Conclusões

Como principais conclusões verifica-se que a maioria dos respondentes pertence a IES universitárias sendo que a maioria dos administradores/gestores dos RI são bibliotecários. Verificou-se também que a um maior número de docentes/investigadores corresponde um maior número de documentos depositados. Este facto não se reflete no número de *downloads* efetuados sobre esses documentos.

Embora se constate que a maioria dos repositórios possui Política de Depósito de Documentos (PDD) o seu cumprimento apenas é monitorizado em 28,6% dos RI o que significa que ainda existe margem para melhoria dos mecanismos de alimentação dos RI das IES portuguesas. Todas as políticas estão disponíveis num sítio *web*, o que denota preocupação com a informação ao utilizador final. Ainda no domínio das políticas de depósito de documentos verifica-se alguma heterogeneidade, na medida em que nem todas

são mandatórias e nem todas se referem a toda a produção científica da respetiva IES existindo alguns mandatos específicos para certos documentos.

Verifica-se também que apesar de existirem incentivos ao depósito, alguns dos respondentes indicaram que o não depósito não acarreta qualquer consequência facto que constitui um obstáculo ao depósito no RI.

As respostas obtidas pela aplicação dos questionários permitiram recolher dados sobre aspetos internos de funcionamento dos repositórios que poderão ajudar no delineamento de medidas que contribuam para melhorar o desempenho dos RI. Em face dos resultados constata-se que, entre outras, as medidas poderão passar pelo desenvolvimento e implementação efetiva de sistemas de incentivo e monitorização concomitante do depósito de documentos no RI por parte da sua comunidade de interesse. Poderão passar igualmente, pela criação de mecanismos de interoperabilidade efetiva entre os vários sistemas internos de registo e validação da informação relativa à publicação científica no sentido de agregar o esforço de trabalho necessário à disponibilização de metadados e documentos. Igualmente se considera que é indispensável a implementação de mecanismos de monitorização do cumprimento das PDD tendo como objetivo último garantir que, num futuro próximo, a totalidade da produção científica de cada instituição está depositada no seu RI.

### Referências bibliográficas

CASSELLA, M. e MORANDO, M. – Fostering new roles for librarians : skills set for repository managers : results of a survey in Italy. *Liber quarterly*, 2011, 21:3/4.

CASSELLA, M. – Institutional Repositories: an internal and external perspective on the value of IRs for researchers' communities. *Liber Quarterly*, 2010, 20:2, p. 210-225.

CHANT, I. - Increasing IR participation: best practices for repository buy-in. *Library Journal*, 2016, 15, p. 20-21.

CULLEN, R. ; CHAWNER, B. – Institutional repositories, open access, and scholarly communication : a study of conflicting paradigms. *Journal of Academic Librarianship*, 2011, 37:6, p. 460-470.

COVEY, D.T. – Recruiting content for the institutional repository: the barriers exceed the benefits. *Journal of Digital Information*, 2011, n.º 12, p.1-18.

HAHN, S. e WYATT, A. – Business faculty's attitudes: open access, disciplinary repositories, and institutional repositories. *Journal of Business & Finance Librarianship*, 2014, 19, p. 93-113.

LAWTON, A. – Communicating the open access message: a case study from Ireland. *New Review of Academic Librarianship*, 2016, 22:1, p. 60-77.

LYONS, C. e BOOTH, H.A. – An overview of open access in the fields of business and management. *Journal of Business and Finance Librarianship*, 2011, 16:22, p. 108-124. DOI:10.1080/08963568.2011.554786.

RODRIGUES, M.E.P.; RODRIGUES, A.M. (2014) - Os autores e o repositório científico – estudo de caso. *RECIIS – Rev. Eletron. de Comun. Inf. Inov. Saúde*, 2014, 8:2, p. 107-121.

SARAIVA, R. ; RODRIGUES, E. – Open access in Portugal. In *Open Access in Southern European countries*. Lluís Anglada e Ernest Abadal, eds. Madrid : FEYCT, 2010. p. 83-99.

XIA, J. – A comparison of subject and institutional repositories in self-archiving practices. *The Journal of Academic Librarianship*, 2008,6, p.489-495.

ZHANG, H., BOOCK, M. e WIRTH, A. – It takes more than a mandate: factors that contribute to increased rates of article deposit to an institutional repository. *Journal of Librarianship and Scholarly Communication*, 2015, 3:1.

ZUCCALA, A., OPPENHEIM, C. e DHIENSA, R. – Managing and evaluating digital repositories. *Information Research*, 2008, 13:1.

---

<sup>i</sup> <http://www.rcaap.pt/>



---

## Repositórios de dados de pesquisa no mundo

---

Michelli Costa

Universidade de Caxias do Sul (UCS) e Universidade de Brasília (UnB)

[michelli@unb.br](mailto:michelli@unb.br)

Tiago Braga

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e  
Universidade de Brasília (UnB)

[tiagobraga@ibict.br](mailto:tiagobraga@ibict.br)

### Resumo

O estudo analisa o desenvolvimento dos repositórios de dados no mundo, a partir de uma abordagem qualitativa e quantitativa. A pesquisa utilizou como fonte de informação os registros dos repositórios disponíveis no diretório re3data.org. Foram definidas categorias e características de análise e os dados obtidos foram utilizados para a identificação do perfil dos repositórios catalogados. Como resultados do estudo observou-se que a área de ciências da vida é responsável pela maioria dos repositórios, sendo que estes estão usualmente hospedados nos países da América do Norte e Europa. Do mesmo modo, constatou-se que a maioria dos repositórios de dados de pesquisa são temáticos, focados no armazenamento de textos e imagens, com pouca representatividade no que diz respeito a certificação e padrões de metadados. No entanto, foi possível observar uma tendência de desenvolvimento de repositórios de dados baseados em três tecnologias dominantes. De forma geral, a maioria dos repositórios de dados distribuem seus conteúdos sob uma licença *Copy Right*, embora seus dados estejam caracterizados como de acesso aberto. Ao final dessa pesquisa conclui-se que para atingir os objetivos da ciência aberta para os dados de pesquisa ainda há avanços a serem alcançados.

**Palavras-chave:** repositórios de dados, dados de pesquisa, ciência aberta.

### Repositories of research data in the world

#### Abstract

This paper presents a qualitative and quantitative approach aiming to analyse how the data repositories around the world are being developed. To do this, the directory

re3data.org was assessed so the repository could be categorized. Some categories and characteristics were defined and the collected data was the source to catalogue the repositories` profile. As result, the research identified that majority of repositories are linked to Life Science and is hosted by North American and European countries. The most part of repositories is theme based, focused in text and image storage, without much certification and metadata standardization concerns and based in three main technologies. As a final result, the analyses showed that most repositories has Copy Rights, but are also described as open access. As a conclusion, this paper presents that some advances are needed to achieve the research data availability, an essential element to the Open Science.

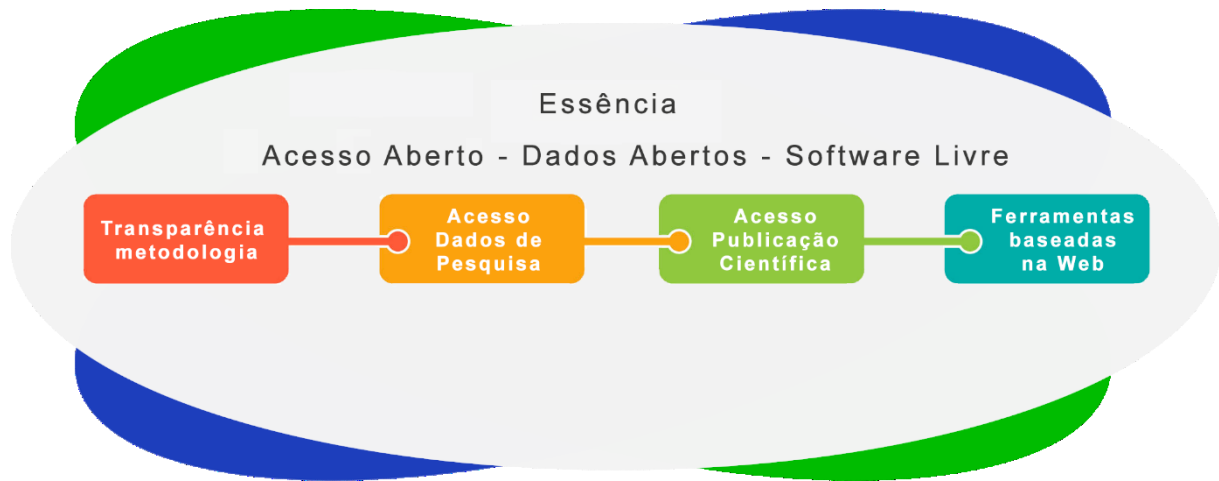
**Keywords:** data repositories, research data, open science.

## Introdução

A emergência de novas formas de comunicação científica, intensificou a demanda por compartilhamento e reuso de dados de pesquisa, um tema relevante e contemporâneo nas discussões acerca da ciência aberta. Segundo Boulton (2013), os princípios que sustentam a ideia da ciência aberta são os mesmos que serviram de base para as revoluções científicas do século XVIII e XIX. Isto porque, em ambos os contextos a motivação seria tornar a ciência pública. O autor cita como gênese destas ideias o pioneirismo de Henry Oldenburg, que solicitou que suas cartas, enviadas à sede da associação científica Royal Society, fossem publicadas com o objetivo de atingir volumoso número de interessados sobre o tema. Oldenburg também solicitou que a publicação das cartas fosse feita em sua língua vernácula, e que fossem publicadas suas evidências para comprovar suas teorias. Neste caso, entendem-se as evidências como os dados coletados pela pesquisa e utilizados para sustentação teórica de sua argumentação. Conforme o autor, a prática de tornar públicas as evidências que sustentam uma teoria favorecem o princípio científico da refutação e do que ele chamou de *scientific self-correction*.

Os dados de pesquisa também são elementos essenciais para o cumprimento dos quatro objetivos da ciência aberta apontados por Gezelter (2009). O primeiro deles trata da transparência da metodologia, observação e coleta dos dados. O segundo objetivo pontua a disponibilidade pública dos dados de pesquisa e permissão para sua reutilização. O terceiro diz respeito a disponibilizar abertamente as publicações científicas. O quarto objetivo evidencia a necessidade do uso de ferramentas baseadas na web com vistas a facilitar a colaboração científica (Figura 1). De acordo com o autor, tais objetivos representam a essência dos projetos relacionados ao acesso aberto, dados abertos e softwares livres.

FIGURA 1: OBJETIVOS DA CIÊNCIA ABERTA



Fonte: Elaboração própria

Segundo definição apresentada pela OECD (2004) dados de pesquisa são registros de fatos usados como fonte primária da pesquisa e que, geralmente, são usados na comunidade científica como necessários para validar os resultados de uma pesquisa. De acordo com Walport e Brest (2011), os repositórios de dados de pesquisa têm sido apontados como uma estratégia eficiente para sua organização, preservação e compartilhamento. No entanto, destacam que estes sistemas necessitam estar bem estabelecidos e dispor de ferramentas que sejam capazes de descrever e divulgar os dados de pesquisa de modo a promover seu amplo acesso e reutilização.

São evidentes as discussões sobre a relevância dos sistemas para o melhoramento dos processos relacionados a comunicação científica e o próprio avanço da ciência, mas estudos sobre o desenvolvimento dos repositórios em nível global ainda são incipientes. Portanto, o presente estudo teve por objetivo analisar o desenvolvimento dos repositórios de dados no mundo.

## Metodologia

A pesquisa é de natureza mista por envolver técnicas e métodos de análise dos dados dentro da perspectiva qualitativa e quantitativa. O universo pesquisado foi limitado aos repositórios cadastrados no diretório re3data.org (Registry of Research Data Repositories). O diretório se apresenta como um registro global de repositórios de dados de pesquisa e cobre diversas áreas do conhecimento. Os dados sobre os sistemas foram coletados no diretório no primeiro semestre de 2016, período em que haviam 1558 repositórios registrados.

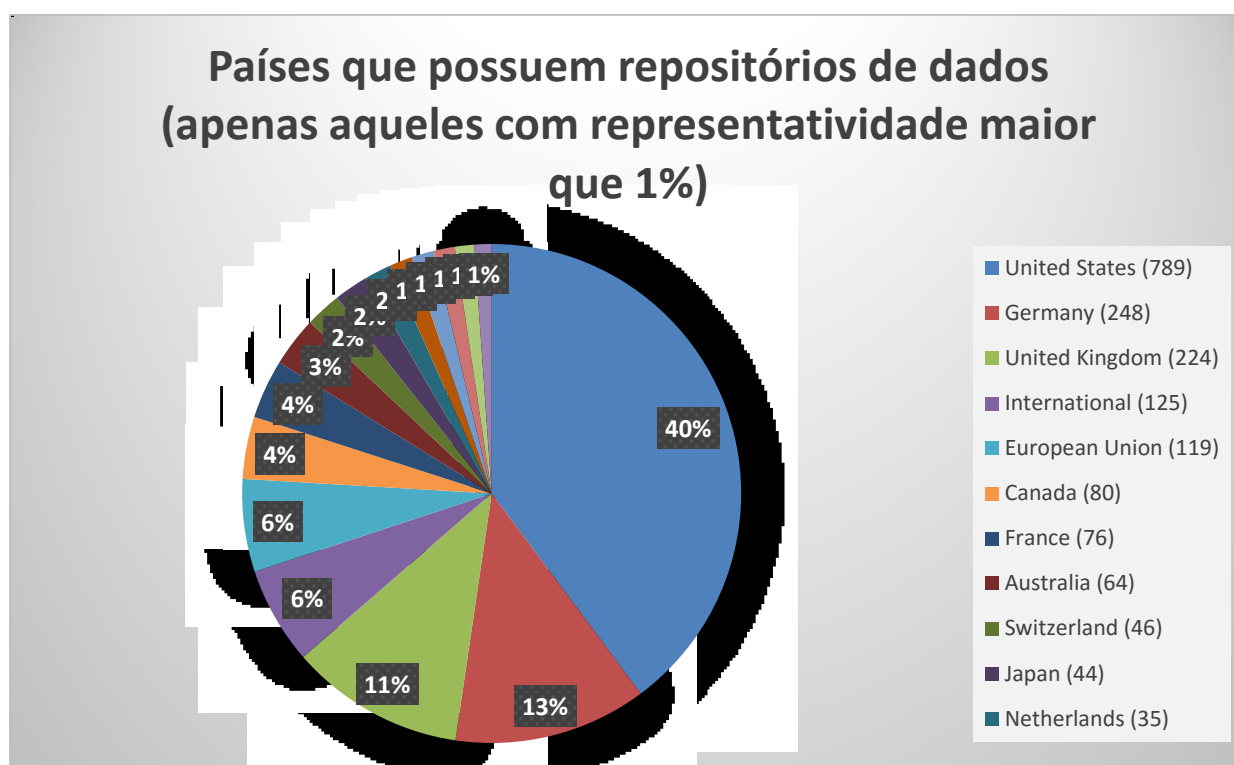
Para a coleta dos dados foram utilizados os sistemas de busca e de descoberta do diretório, que permite, entre outras funcionalidades, a organização dos dados segundo

critérios previamente determinados. Os resultados obtidos foram sistematizados em quatro categorias: abrangência geográfica, abrangência temática, características dos sistemas e promoção da ciência aberta.

### Abrangência geográfica dos repositórios de dados de pesquisa

Na categoria abrangência geográfica foram identificadas e organizadas informações acerca dos países pelos quais os repositórios estão registrados. O país com maior quantidade de repositórios foram os Estados Unidos da América, de onde originam-se 40% dos repositórios de dados do mundo. A Alemanha e o Reino Unido também tiveram presença relevante no universo, representando 13% e 11% respectivamente. Além dos países da América do Norte e da Europa, apenas a Austrália, o Japão e a Índia obtiveram representação entre os países com pelo menos 1% do conjunto total de repositórios (Figura 2). Com isto, observou-se que nenhum país da América Latina e da África apresentaram expressão em quantidade de repositórios de dados de pesquisa.

FIGURA 2: ABRANGÊNCIA GEOGRÁFICA DOS DEPOSITÓRIOS DE DADOS DE PESQUISA

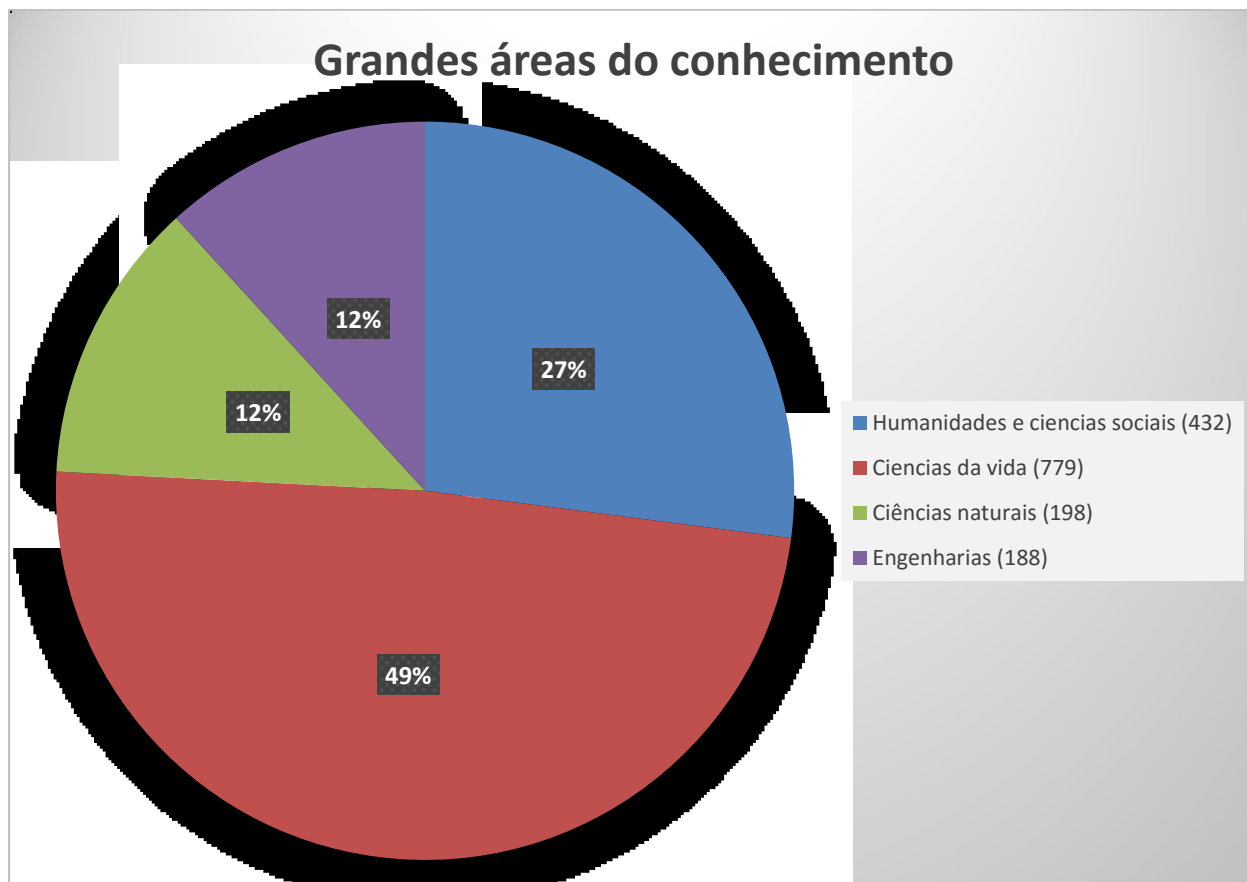


Fonte: Elaboração própria

### Abrangência temática dos repositórios de dados de pesquisa

A segunda variável tratou sobre a cobertura temática dos repositórios analisados. Nesse aspecto, foi identificado que aproximadamente metade dos repositórios são relacionados às ciências da vida e a outra metade dos repositórios, dividem-se entre ciências naturais (12%), engenharias (12%), ciências sociais e humanidades (27%) (Figura 3).

FIGURA 3: ÁREAS DO CONHECIMENTO DOS REPOSITÓRIOS DE DADOS DE PESQUISA



Fonte: Elaboração própria

### Características dos repositórios de dados de pesquisa

Na terceira categoria de análise foram exploradas cinco características dos sistemas, a saber: tipo de repositório, tipo de recursos armazenados, certificados dos repositórios, software dos repositórios e metadados.

#### *Tipo de repositório*

De forma geral, identificou-se que 69% dos repositórios de dados são temáticos, ou seja, são limitados por área do conhecimento e não por instituições (Figura 4).

FIGURA 4: TIPOS DE REPOSITÓRIOS DE DADOS DE PESQUISA

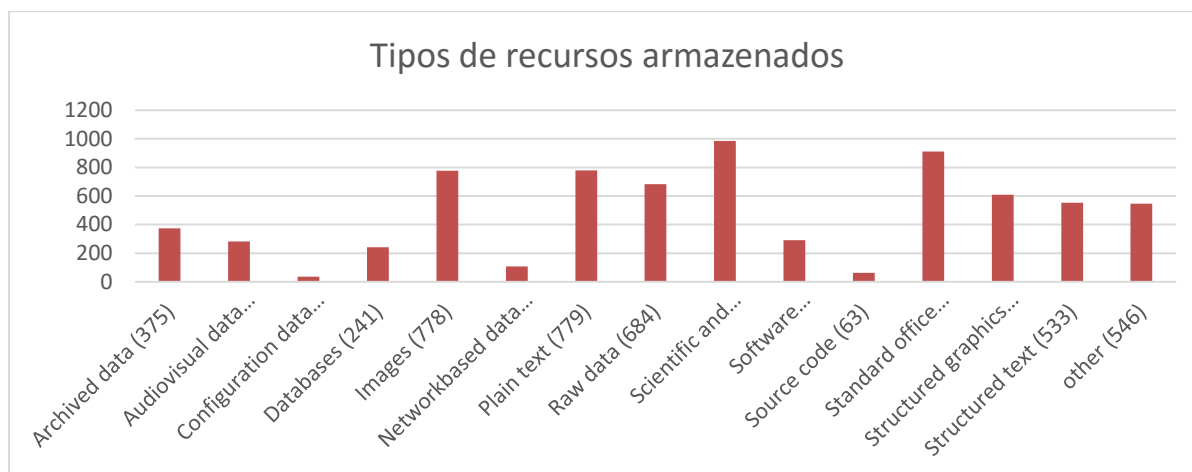


Fonte: Elaboração própria

**Tipos de recursos armazenados**

Dentre os documentos que os sistemas armazenam e disponibilizam, destacam-se quatro tipos: os dados científicos em formatos estatísticos (14%), documentos de texto (*office*) padrão (13%), imagens (11%) e textos simples (11%). Os outros tipos representam menos de 10% do total de documentos, inclusive os dados brutos, que são recorrentemente citados como necessários para a reutilização dos dados no contexto da ciência colaborativa (Figura 5).

FIGURA 5: TIPOS DE RECURSOS ARMAZENADOS NOS REPOSITÓRIOS DE DADOS DE PESQUISA

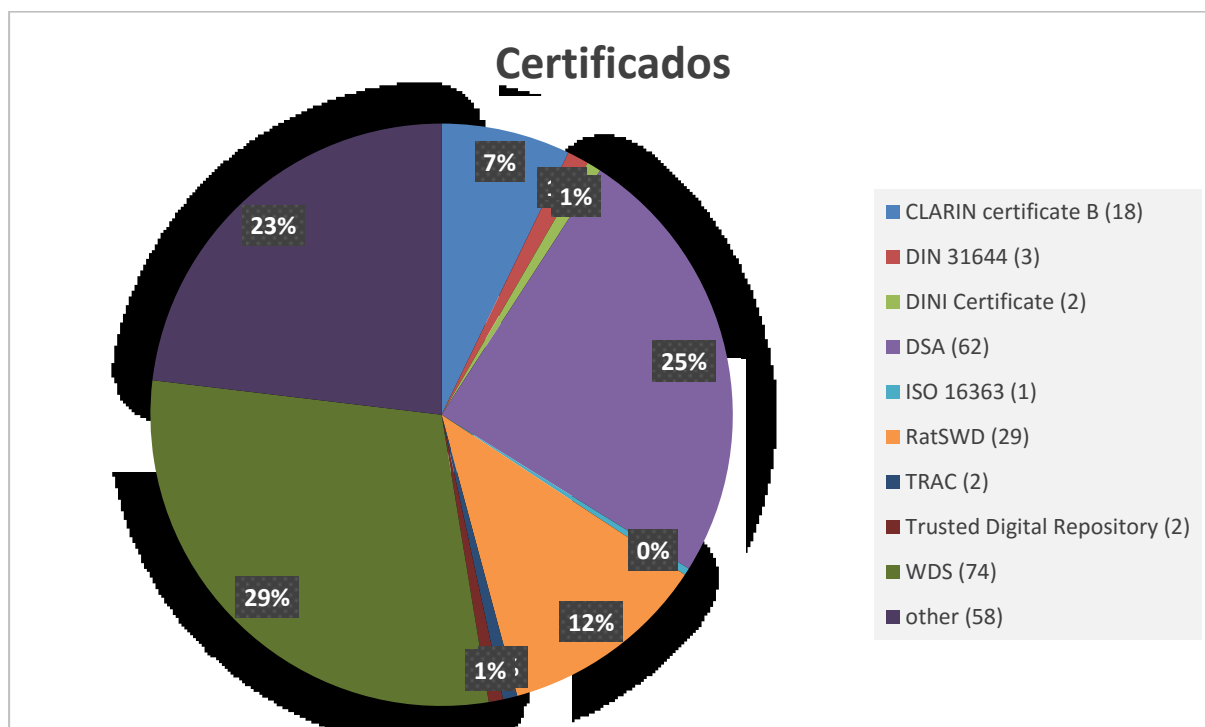


Fonte: Elaboração própria

### *Certificados dos sistemas*

Cerca de 17% dos repositórios apresentavam algum tipo de certificado de sistema. Dentre eles, destacaram-se o World Data System (WDS), em 29%; o Data Seal of Approval (DSA), em 25%; e o German Data Forum (RatSWD), em 12% destes repositórios que possuem certificados. O conjunto total de certificados utilizados pelos repositórios é apresentado na Figura 6.

FIGURA 6: CERTIFICADOS APRESENTADOS PELOS REPOSITÓRIOS DE DADOS DE PESQUISA



Fonte: Elaboração própria

O WDS é uma iniciativa do International Council for Science (ICSU) e foi criado em 2008. Seu objetivo é promover a governança de longo prazo e acesso universal e equitativo a dados científicos com garantia de qualidade, além de serviços de dados, produtos e informação.

O DSA é um selo desenvolvido pela Data Archiving and Network Services (DANS), um instituto criado em 2009 pela Academia Real de Arte e Ciência dos Países Baixos (KNAW) com apoio da Organização para Pesquisa Científica dos Países Baixos (NWO). Seu objetivo é resguardar dados, garantindo sua qualidade, bem como garantir o gerenciamento confiável dos dados para uso futuro sem a necessidade de implementar novos padrões ou onerar sua utilização.

O RatSWD é um conselho independente de diversos pesquisadores de universidades, institutos de pesquisa ligados a universidades e independentes e produtores de dados da

Alemanha. Foi criado em 2004 pelo Ministério Alemão de Educação e Pesquisa e seu objetivo é melhorar a pesquisa em infraestrutura de dados e sua competitividade internacional.

Ao se analisar os três certificados mais utilizados é possível perceber que eles possuem características bem peculiares, apesar de algumas similaridades. Do ponto de vista de abrangência, o RatSWD certifica apenas repositórios que possuem participação Alemã em sua constituição. Já o DSA e o WDS certificam repositórios independente dos países em que estes foram criados. Outra peculiaridade pode ser identificada ao se analisar o processo de criação destes certificados. Enquanto o WDS foi criado a partir de um consórcio internacional, o DSA e o RatSWD possuem sua criação baseada nos Países Baixos e na Alemanha respectivamente. O DSA foi fomentado por entidades de governo e o RatSWD, apesar de possuir respaldo do ministério alemão para Educação e Pesquisa também possui instituições privadas ligadas a sua criação.

Por fim, pode-se analisar os três certificados quanto às temáticas dos repositórios depositados. Enquanto 92% dos repositórios que utilizam o WDS armazenam dados relacionados às ciências naturais ou ciências exatas e da terra o DSA certifica 94% de repositórios cujo foco é em ciências humanas, 45% em Letras e Linguística e 29% em ciências naturais. Já os repositórios certificados pelo RatSWD possuem seu foco principalmente em ciências humanas, 72%, sociais aplicadas, 59%, e saúde, 21%.

### ***Software***

As informações acerca do *software* dos repositórios de dados de pesquisa cadastrados no re3data.org foram consideradas as menos abrangentes obtidas nesse estudo. Apenas 12% dos registros apresentaram alguma indicação a respeito dos programas computacionais utilizados para gerenciar a base de dados de seus sistemas. Parte dos sistemas que não apresentaram informações relacionadas ao software (22% do total), indicaram que não haviam opções pertinentes sobre o tema no momento do registro e portanto, responderam utilizar “outro” software entre os listados.

Dentre os repositórios que indicaram o software utilizado, a maioria (25%) está relacionada ao uso do DSpace (Figura 7). O software DSpace permite o acesso a todos os tipos de conteúdo digital e possui sua utilização focada em academia e organizações comerciais e sem fins lucrativos. Ele é gratuito e de código aberto e pode ser facilmente adaptado para atender as especificidades do repositório. Ele é mantido pela DuraSpace, uma organização sem fins lucrativos.

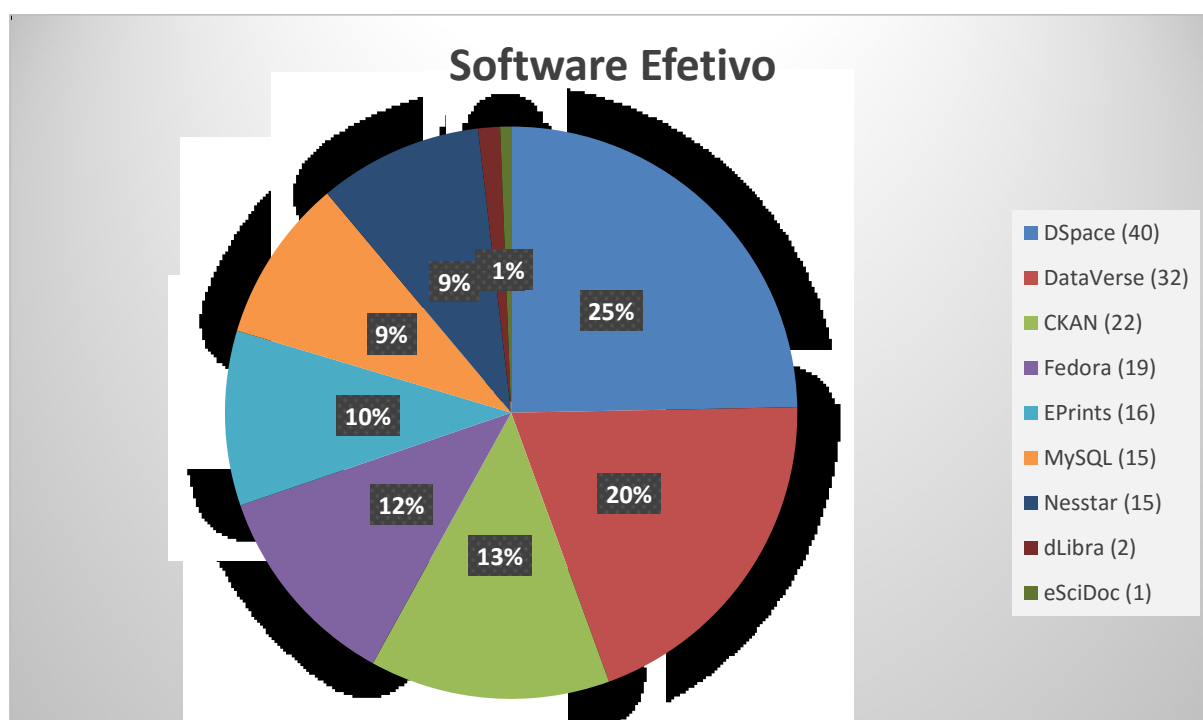
O segundo software mais utilizado para criação de repositórios de dados de pesquisa identificado pelo estudo foi o DataVerse, que está presente em 20% dos sistemas que apresentaram informação acerca do software. O DataVerse é um software de código aberto voltado para preservação, citação, exploração e análise de dados de pesquisa. Foi idealizado



através da parceria entre o Instituto para Ciência Social Quantitativa e a Universidade de Harvard. Além do software essa parceria permitiu a criação de um programa de suporte na utilização do sistema.

O software CKAN foi identificado como o terceiro mais utilizado entre os sistemas analisados, estando presente em 13% dos repositórios de dados de pesquisa. O CKAN é mantido pela Open Knowledge Foundation e possui o propósito de compartilhar, publicar, encontrar e utilizar dados. Sua utilização é focada em governos e iniciativa privada.

**FIGURA 7: SOFTWARE UTILIZADOS PELOS REPOSITÓRIOS DE DADOS DE PESQUISA**



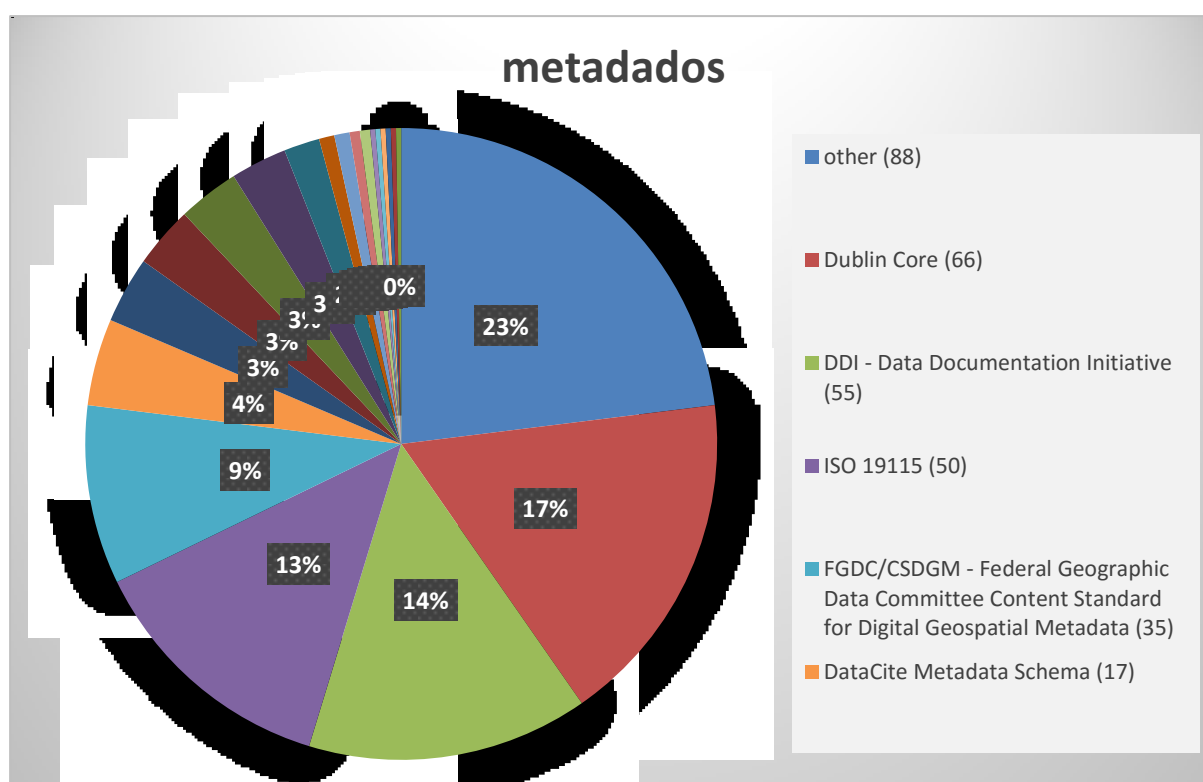
Fonte: Elaboração própria

De forma geral, o software realiza basicamente a mesma função de gestão de dados de pesquisa. Enquanto o DSpace se coloca claramente focado no ambiente acadêmico, CKAN e DataVerse não enfatizam este uso de forma tão ativa. No que se refere à mantenedora do software, tanto o DSpace quanto o CKAN são mantidos por instituições sem fins lucrativos, enquanto o DataVerse possui uma grande instituição mantenedora, a Universidade de Harvard. Uma das grandes características do DSpace é a possibilidade de adaptação do sistema para as necessidades do usuário e sua relação com uma grande comunidade de usuários e desenvolvedores.

## Metadados dos repositórios

No que diz respeito à característica metadados, foi possível compreender, de forma não-abrangente, os padrões utilizados pelos repositórios. Do total de repositórios disponíveis no diretório re3data, apenas 24% deles apresentam informações relativas aos metadados utilizados para a descrição dos seus recursos. Destes destacam-se o Dublin Core, Data Documentation Initiative (DDI) e a International Organization for Standardization (ISO) 19115. Destaca-se aqui que 23% dos repositórios que possuem categorização por metadados não apresentam a definição exata de qual padrão utilizam, sendo categorizados como outros (Figura 8).

FIGURA 8: PADRÕES DE METADADOS DOS REPOSITÓRIOS



Fonte: Elaboração própria

O padrão Dublin Core está presente em 17% dos repositórios que apresentam informações sobre seus metadados. Segundo informações apresentadas no DCMI (2016), o padrão possui cinco grandes princípios: construção de forma aberta e sob consenso; participação e escopo internacional; neutralidade em seus propósitos e modelos de negócio; neutralidade na tecnologia e foco transdisciplinar. A organização e operação da iniciativa é formalizada através de uma série de regras que definem como os membros da iniciativa atuam no que diz respeito a suas responsabilidades e os processos de decisão sobre o padrão.

Já o padrão DDI é focado na descrição de dados provenientes das ciências sociais, comportamentais e econômicas (KRAMER; LEAHEY, 2012) e está presente em 14% dos repositórios que possuem esta característica. Este padrão é expresso em XML e contempla todo o ciclo de vida dos dados de pesquisa, desde a etapa de conceituação, passando pela análise e arquivamento dos dados. O projeto de criação do padrão DDI remonta ao ano de 1995, sendo que a primeira versão foi apresentada à comunidade em idos de 2000. Desde então o padrão passou por diversos processos de redesenho e a versão atual é a DDI Lifecycle 3.2, que possui características modulares e a capacidade de ser expandida (DATA DOCUMENTATION INITIATIVE, 2013).

Por fim, o padrão ISO 19115 é aplicado em 13% dos repositórios. O padrão foi atualizado pela última vez no ano de 2014 e seu foco é na descrição de informações geográficas. Segundo a ISO (2014), o padrão é aplicável a catalogação de todos os tipos de recursos, bem como serviços geográficos, conjunto de dados e características geográficas.

Ao se analisar os três padrões de metadados percebe-se que o Dublin Core possui características mais abrangentes em relação a sua utilização. O DDI também possui esta característica, mas o seu processo de criação foi focado em áreas específicas do conhecimento. Por fim, o padrão ISO 19115, embora possua a indicação de uso para qualquer tipo de documento, seu desenvolvimento foi focado em dados geográficos, o que indica que o padrão é o mais específico entre os três.

## **Repositórios de dados de pesquisa e a ciência aberta**

A última categoria de análise avaliou o comportamento dos repositórios em relação aos princípios da ciência aberta declarados pelo Panton Principles (2011). O Panton Principles: Principles for Open Data in Science, foi elaborado por quatro pesquisadores, três do Reino Unido e um dos Estados Unidos, e é assinado por mais de 265 pesquisadores. O documento parte da ideia geral que a ciência é baseada na reutilização e discussão do conhecimento científico já publicado, e que este processo é mais efetivo se os dados das publicações estiverem abertos, ressaltando assim os princípios da ciência moderna. De acordo com Molloy (2011), a abertura dos dados significa a disponibilização livre na Internet com permissões para download, cópia, análise, reprocessamento e uso em software. Concluindo que os dados precisam ser disponibilizados em um local de domínio público. Após a definição de contexto e de conceitos, são propostos quatro princípios para a abertura dos dados de pesquisa (Quadro 1)

QUADRO 1: PANTON PRINCIPLES

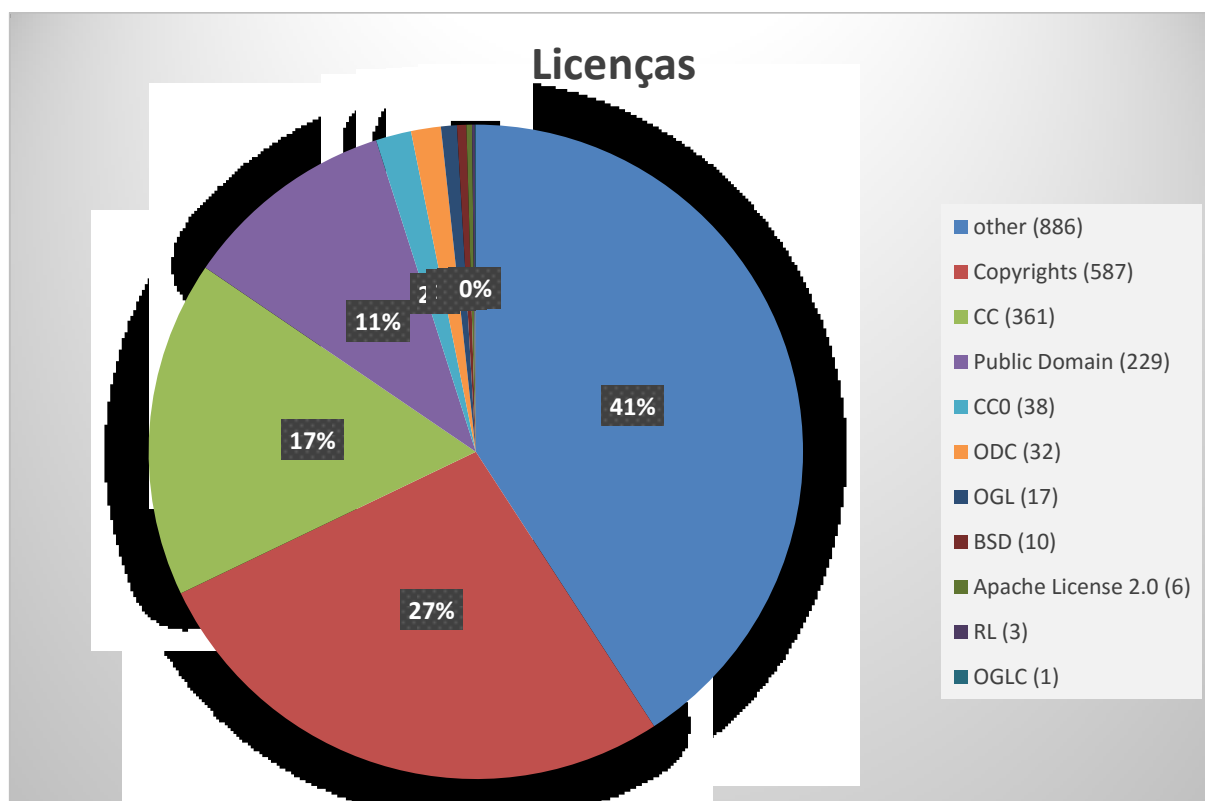
Princípios
1. É livre para uso e reuso.
2. É distribuído sem restrições
3. Utiliza a licença que determina que dados derivados compartilha a mesma licença de abertura.
4. É disponível em todo seu conjunto.
5. Quando possuir custo de reprodução, este deve ser razoável.
6. Preferencialmente deve ser apto para o acesso por meio da Internet e sem custos.
7. É disponível em formato conveniente e modificável.

Fonte: Elaboração própria

Nos princípios enunciados sobressai a perspectiva de abertura dos dados, no qual implica a disponibilização livre na Internet com permissões para *download*, cópia, análise, reprocessamento e uso em *software*. Para tanto, as permissões devem ser explicitamente declaradas por meio de uma licença adequada, conforme destacou Molloy (2011).

De forma geral, os repositórios analisados não foram satisfatórios em relação a apresentação de licença para a ciência aberta, pois 41% deles não apresentaram informações sobre a licença utilizada e 27% declararam utilizar a licença *copyright*. Apenas um quarto dos repositórios cumpriram a adequação proposta pelo Panton Principles, 17% utilizando licenças Creative Commons e 11% empregando licenças Public Domain (Figura 9).

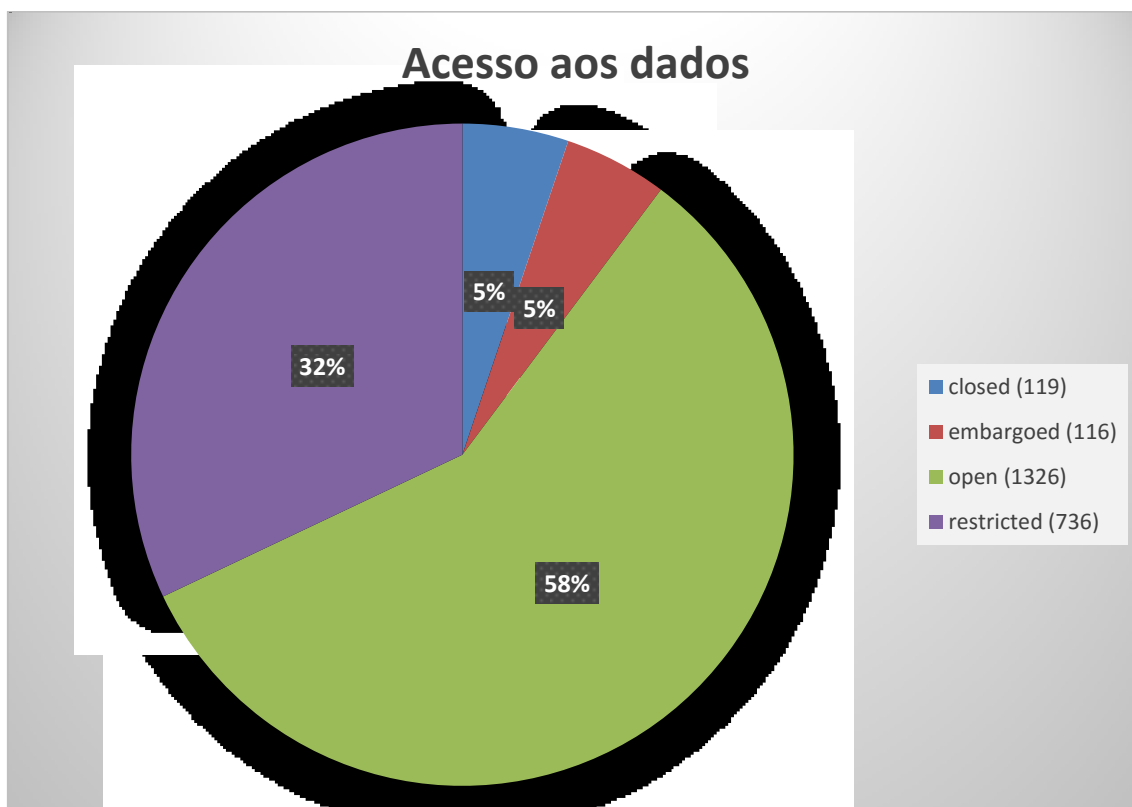
FIGURA 9: LICENÇAS UTILIZADAS NOS REPOSITÓRIOS DE DADOS DE PESQUISA



Fonte: Elaboração própria

Apesar das deficiências relacionadas às licenças de acesso e uso aos conteúdos, 58% dos repositórios analisados declararam permitir o acesso aberto direto e 5% somente após um período de embargo. Na contramão da ciência aberta, 37% dos repositórios declararam não promover o acesso aberto aos dados de pesquisa que armazenam (Figura 10).

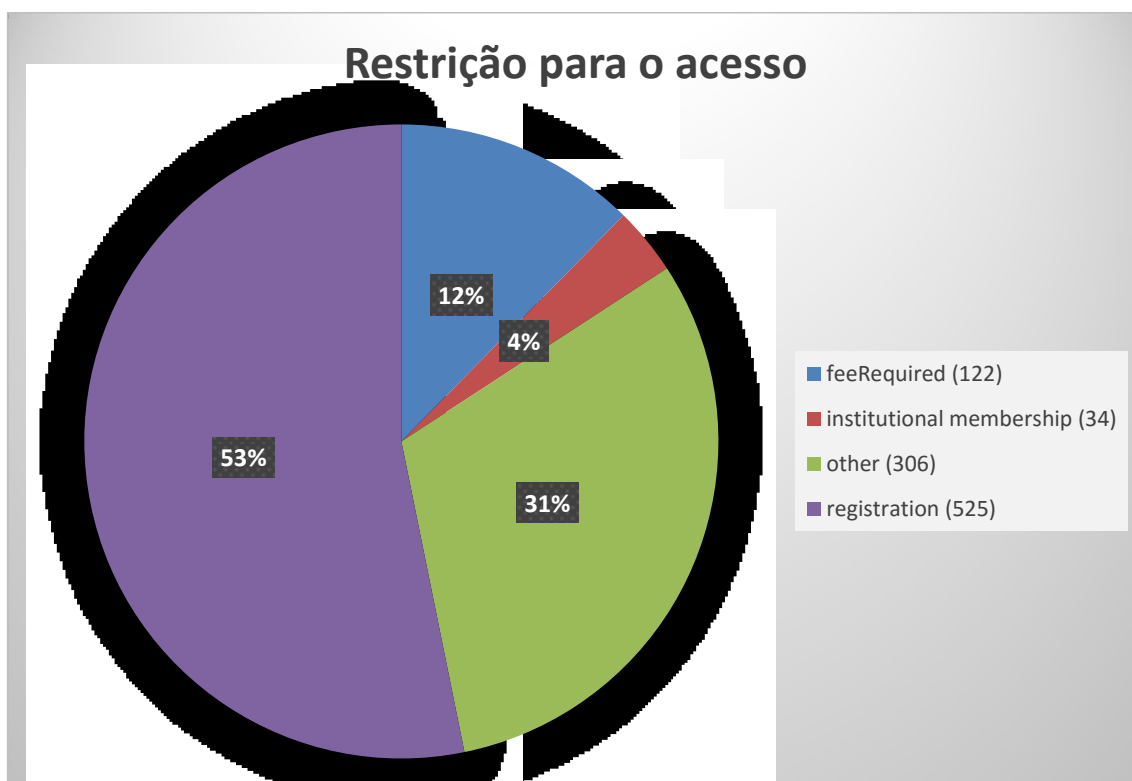
FIGURA 10: POLÍTICA DE ACESSO AOS DADOS ARMAZENADOS NOS REPOSITÓRIOS DE DADOS DE PESQUISA



Fonte: Elaboração própria

Parte das restrições impostas pelos repositórios relacionam-se ao registro dos usuários no sistema (53%) e à verificação da filiação institucional do usuário (4%). Além destas, observou-se que 12% das restrições implicam no pagamento de taxas para o acesso aos dados de pesquisa armazenados (Figura 11).

FIGURA 11: RESTRIÇÕES PARA O ACESSO AO CONTEÚDO DISPONÍVEL NOS REPOSITÓRIOS DE DADOS DE PESQUISA



Fonte: Elaboração própria

### Considerações finais

Conclui-se, portanto, que o cenário mundial dos repositórios de dados de pesquisa aponta uma tendência para criação de repositórios temáticos, em vez de institucionais e aplicação de certificados específicos de acordo com as áreas do conhecimento, embora a certificação não seja necessariamente uma tendência. A ampla expressividade das Ciências da Saúde indica que suas práticas de financiamento, produção e comunicação de conhecimento científico são favoráveis ao contexto da ciência colaborativa e ao compartilhamento e reutilização de dados de pesquisa. No entanto, observou-se que metade dos repositórios analisados, independente da área do conhecimento, não licenciam seus registros de forma adequada para a ciência aberta, comprometendo sua ampla disseminação e reutilização. Por fim, a visão ampla acerca da disponibilidade de dados apresentada nesse artigo indica que para atingir os objetivos da ciência aberta há, ainda, que se obter avanços nessa área que é considerada um dos elementos essenciais para a democratização da ciência.

### Referências bibliográficas

BOULTON, G. Reinventing Open Science for the 21st Century. In: Uma década de acesso

aberto na UMinho e no Mundo. Lisboa: Universidade do Minho, 2013. p. 239-250.

DATA DOCUMENTATION INITIATIVE. DDI Lifecycle 3.2. Disponível em: <<http://www.ddialliance.org/Specification/DDI-Lifecycle/3.2/>>. Acesso em: 6 out. 2016.

DCMI. Dublin Core Metadata Initiative (DCMI): About. Disponível em: <<http://dublincore.org/>>. Acesso em: 13 out. 2016.

GEZELTER, D. What, exactly, is Open Science?The OpenScience Project, 2009. Disponível em: <<http://www.openscience.org/blog/?p=269>>. Acesso em: 4 out. 2016

ISO. ISO 19115-1:2014 – Geographic information -- Metadata -- Part 1: Fundamentals. Disponível em: <[http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=53798](http://www.iso.org/iso/catalogue_detail.htm?csnumber=53798)>. Acesso em: 13 out. 2016.

KRAMER, S.; LEAHEY, A. Using RDF to describe and link social science data to related resources on the Web: leveraging the Data Documentation Initiative (DDI) model – Parade@Portsmouth. . In: DATA DOCUMENTATION INITIATIVE. Ann Arbor, Michigan: 2012Disponível em: <<http://eprints.port.ac.uk/9029/>>. Acesso em: 12 maio. 2016

MOLLOY, J. C. The Open Knowledge Foundation: Open Data Means Better Science. PLoS Biol, v. 9, n. 12, p. e1001195, 6 dez. 2011.

OECD. Declaration on Access to Research Data from Public FundingOrganization for Economic Co-operation and Development (OECD), , 2004. Disponível em: <[http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=157#\\_ftn0](http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=157#_ftn0)>

Panton Principles. Disponível em: <<http://pantonprinciples.org/>>. Acesso em: 12 maio. 2016.

WALPORT, M.; BREST, P. Sharing research data to improve public health. The Lancet, v. 377, n. 9765, p. 537-539, 18 fev. 2011.

---

## Panorama do autoarquivamento nos repositórios institucionais portugueses

---

Viviane Santos de Oliveira Veiga

ICICT. Fundação Oswaldo Cruz

[viviane.veiga@icict.fiocruz.br](mailto:viviane.veiga@icict.fiocruz.br)

Luis Guilherme Gomes de Macena

ICICT. Fundação Oswaldo Cruz

[guilhermelg2004@gmail.com](mailto:guilhermelg2004@gmail.com)

Cícera Henrique da Silva

ICICT. Fundação Oswaldo Cruz

[cicera.henrique@icict.fiocruz.br](mailto:cicera.henrique@icict.fiocruz.br)

Maria Manuel Borges

Universidade de Coimbra

[mmb@fl.uc.pt](mailto:mmb@fl.uc.pt)

### Resumo

Este trabalho almejou obter um panorama do compartilhamento de artigos científicos através de repositórios institucionais (RIs) em Portugal e a percepção do gestor de RI quanto a esta forma de compartilhamento. Foi realizada pesquisa exploratória, utilizando como instrumento de coleta de dados questionário eletrônico semiestruturado que foi aplicado a gestores de RIs em Portugal. A seleção dos repositórios que entrariam na amostra foi realizada a partir das informações disponibilizadas no OpenDOAR e no RCAAP. Retiradas as duplicidades, foram encontrados 48 RIs com coleção de artigos em Portugal, sendo que 1 estava com acesso inativo. O questionário online foi enviado para os 47 correios eletrônicos ou “Fale Conosco” dos RIs. Foram retornados 27 questionários com respostas válidas. Constatou-se que apenas 19% dos repositórios institucionais portugueses não possuem o autoarquivamento habilitado no sistema. Nestes repositórios verificou-se que a maioria dos gestores (81%) não acreditam que o autoarquivamento funcionaria com seus autores. Entre os RIs com autoarquivamento habilitado, alguns gestores



(27%) afirmam que 91% a 100% dos documentos disponibilizados são frutos do autoarquivamento. Porém, 41% dos repositórios afirmam que apenas 6 a 30% dos materiais no repositório foram autoarquivado e a maioria, 59%, declara que menos de 30% do material disponível no RI foi autoarquivado. Concluiu-se que os gestores dos repositórios portugueses, em sua grande maioria, viabilizam o autoarquivamento de seus autores, habilitando esta função no sistema. Porém, é imprescindível conhecer as especificidades das áreas, estudar as barreiras e os estímulos ao compartilhamento de artigos para ampliar a adesão ao autoarquivamento por parte dos pesquisadores em Portugal.

**Palavras-chave:** Acesso aberto ao conhecimento, Autoarquivamento, Repositório Institucional, Compartilhamento de informação em acesso aberto.

## Introdução

### Contexto do problema

#### *O papel do periódico científico como veículo fundamental na comunicação do conhecimento científico*

O Movimento de Acesso Aberto ao Conhecimento, desencadeado em 2002, propõe dois caminhos para se atingir o objetivo de acesso aos resultados da investigação: (i) o acesso aberto (AA) através do autoarquivamento em repositórios institucionais e/ou temáticos, também conhecido pela via verde ou (ii) a publicação em periódicos em AA também conhecido por via dourada.

Após 14 anos do AA pretende-se verificar os avanços da primeira estratégia desta declaração, o autoarquivamento, em Portugal. Migueis (2012) analisando os documentos depositados no Repositório Institucional (RI) da Universidade de Coimbra, verificou que apenas 6% dos documentos em livre acesso foram inseridos através do autoarquivamento. Ela afirma que existem indícios que um conjunto importante dos autores portugueses não deposite diretamente a sua produção no RI.

Apesar de saber que existem vários fatores que podem influenciar positivamente ou negativamente na decisão do pesquisador em compartilhar os seus artigos em repositórios, nesta pesquisa nos limitaremos a verificar a percepção do gestor dos repositórios sobre o tema e o índice de autoarquivamento alcançado nos RIs de Portugal.

O gestor do RI tem uma função importante na condução do Acesso Aberto pela via verde na instituição. Normalmente é ele quem conhece as possibilidades e objetivos do sistema e tem em suas mãos o poder de decidir sobre as configurações do RI. Na maioria das instituições o gestor do RI é ouvido pela direção da instituição sobre a condução do RI. É claro que em alguns casos o gestor pode ver suas propostas para condução do RI na instituição sendo ignoradas, mas de modo geral ele ocupa um lugar de influência nesta gestão. O entendimento e a visão do gestor do repositório institucional são importantes para que o movimento ganhe força e atraia a comunidade científica do país.

## **Antecedentes**

O ato de comunicar é fundamental para a ciência, seja através de canais formais ou informais, ao permitir a disseminação tão ampla quanto possível do conhecimento científico. “A realização de pesquisas e a comunicação de seus resultados são atividades inseparáveis” (MEADOWS, 1999, p. 61). Desde os primórdios da comunicação científica existe a necessidade de se comunicar, debater, colocar à prova as descobertas científicas e construir coletivamente o conhecimento. Desde a publicação do primeiro periódico científico em 1665, o artigo científico vem se consagrando como o veículo mais prestigiado para comunicação científica, na maioria das áreas do conhecimento. O aumento no custo das assinaturas dos periódicos causou um retrocesso nesta expansão. Este fenômeno, conhecido por “crise dos periódicos”, impediu que as instituições pudessem manter suas assinaturas, criando uma barreira entre o pesquisador e o conhecimento certificado.

A Internet e a World Wide Web permitem que os autores ultrapassem as várias limitações do meio impresso e comecem a testar alternativas para facilitar essa disseminação e aumentar o impacto do conhecimento publicado.

Tais iniciativas congregaram-se em torno do que conhecemos como Acesso Aberto (AA). O AA, segundo Harnad (2012), é o acesso online aos resultados de pesquisa publicados em periódicos com avaliação por pares. Harnad afirma que existem outros autores que defendem que o AA é mais do que isto, mas que todos querem ao menos essa tipologia documental em AA. O AA foi desencadeado pela comunidade científica através da Declaração de Budapeste, em 2002, e as Declarações de Berlim e de Bethesda em 2003. Nestas declarações são propostas duas estratégias para o acesso aos artigos científicos, a primeira: o autoarquivamento em repositórios; a segunda, a publicação de artigos em periódicos de AA.

Repositório Institucional (RI) é um sistema para disseminação, armazenamento e preservação da produção científica e intelectual de uma instituição. O autoarquivamento é incentivado desde a Declaração de Budapeste para todos os pesquisadores, mas tem encontrado barreiras em vários países na sua execução. Portugal tem desempenhado um papel de destaque no cenário mundial como um grande incentivador na criação de repositórios e de políticas mandatórias.

## **Via verde em Portugal**

O Movimento de Acesso Aberto em Portugal obteve grandes conquistas ao longo dos anos. Saraiva e Rodrigues (2010) relatam que as primeiras iniciativas ocorreram em 2002, mas a criação do RepositoriUM (Repositório Institucional da Universidade do Minho) em 2003 foi um dos fatores predominante para alavancar o crescimento do AA no território português (RODRIGUES, SARAIVA, 2013). Uma outra iniciativa importante foi a elaboração da Declaração de Acesso Aberto do Conselho de Reitores Portugueses (CRUP), em 2006. Para consolidar e

ampliar a visibilidade do AA, no ano de 2008 foi criado o “meta-repositório nacional e de um serviço de alojamento de novos repositórios” (SARAIVA, RODRIGUES, 2010, p. 3), então denominado de Repositório Científico de Acesso Aberto de Portugal (RCAAP). O RCAAP potencializou a visibilidade dos repositórios portugueses criando uma rede de colaboração onde possibilitou que em um único local pudesse acessar o conteúdo dos demais repositórios portugueses. A Fundação para a Ciência e Tecnologia (FCT) criou a Política sobre Acesso Aberto a Publicações Científicas resultantes de Projetos de I&D financiados pela FCT. A partir dessa política ocorreu um controle sobre as publicações científicas financiadas com recursos públicos. Uma das obrigações dos autores com financiamento (integral ou parcial) era o depósito em repositórios dos resultados de pesquisa em versão final em pelo menos um dos repositórios integrante do RCAAP (FUNDAÇÃO PARA A CIÊNCIA E TECNOLOGIA, 2014).

### **Procedimentos metodológicos**

Para obter um panorama do autoarquivamento em Portugal e conhecer a visão dos gestores dos repositórios foi realizada pesquisa exploratória, utilizando como instrumento de coleta de dados questionário eletrônico semiestruturado que foi aplicado a gestores de repositórios institucionais em Portugal. A seleção dos repositórios que entrariam na amostra foi realizada a partir das informações disponibilizadas no OpenDOAR e no Repositório Científico de Acesso Aberto de Portugal (RCAAP). O corpus da pesquisa foi constituído mediante os seguintes critérios de seleção: área geográfica do repositório, Portugal; tipo do conteúdo, artigo; e característica do repositório, institucional. O levantamento dos dados foi realizado entre 30 de março e 28 de abril de 2016. Foram identificados 43 repositórios institucionais portugueses no OpenDOAR. Destes, 39 repositórios declaravam possuir a coleção de artigos. No RCAAP foram identificados 49 RIs portugueses, destes 48 com coleção de artigos. Verificamos neste levantamento que o RCAAP possui mais informações sobre os repositórios em Portugal do que o OpenDoar. Retiradas as duplicidades foram encontrados 48 RIs com coleção de artigos, sendo que 1 estava com acesso inativo.

O questionário online, criado utilizando o aplicativo Google Forms, foi enviado para os 47 correios eletrônicos ou “Fale Connosco” dos RIs, cadastrados no OpenDOAR, no RCAAP ou nas páginas do repositório em duas sessões, a primeira entre os dias 04 e 30 de abril de 2016 e a segunda entre os dias 1 a 18 de agosto de 2016. Foram retornados 27 questionários com respostas válidas.

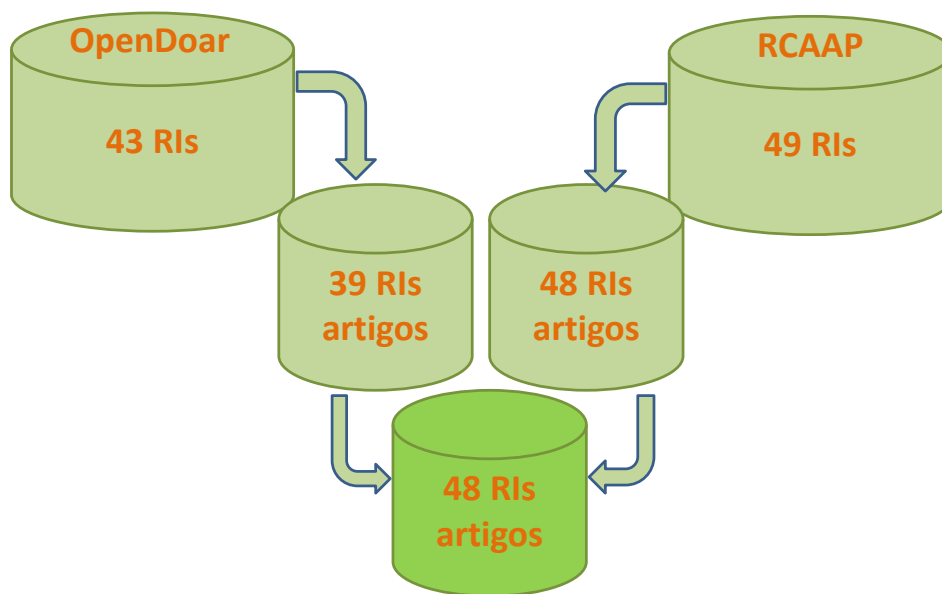
### **Resultados e discussões**

De acordo com a BOAI os RIs foram criados objetivando o compartilhamento de artigos científicos, isto porque estes eram (e ainda são) considerados como o principal veículo para a comunicação do conhecimento científico na maioria das áreas do conhecimento. Vários autores (CROW, 2002; LYNCH, 2003; SWAN 2008; HARNAD, 2012) definem repositórios como veículo utilizado pelo pesquisador para compartilhar seus artigos

científicos em acesso aberto. Com o passar dos anos os repositórios foram agregando outras tipologias documentais importantes e complementares, como os dados de pesquisa, os recursos educacionais, trabalhos apresentados em congressos, relatórios de pesquisa, teses e dissertações, entre outros. Apesar do momento de transição que vivemos, e da importância que novas tipologias vem adquirindo no processo de comunicação científica, o artigo científico ainda tem papel de destaque e, portanto ainda deve ser usado como o principal indicador para o avanço da via verde, cumprindo os objetivos do Movimento de AA.

Alguns países como o Brasil (VEIGA, MACENA, 2015) têm avançado na criação de repositórios sem a coleção de artigos. São Repositórios de teses e dissertações, ou de outras coleções importantes no processo de produção de conhecimento, mas não o coração dela.

Em Portugal percebe-se um entendimento quanto a importância do acesso aos artigos científicos na via verde e o papel dos RIs neste contexto. Verifica-se isto no levantamento realizado no OpenDOAR onde 39 repositórios, entre um total de 43 RIs, declaram possuir a coleção de artigos. O mesmo se percebe no levantamento no RCAAP, onde foram identificados 49 RIs portugueses, destes 48 com coleção de artigos.



Quadro 1 – Levantamentos no OpenDoar e RCAAP

Um outro princípio na criação dos repositórios, inspirado pela proposta subversiva de Stevan Harnard, e nos primeiros repositórios Citeseer e Arxiv, é que o próprio autor possa compartilhar em acesso aberto seus artigos, sem intermediários, o que na BOAI é chamado de *Selfarchiving* (autoarquivamento). Apesar dos softwares mais utilizados para construção de RIs (Dspace e EPrints) possuírem como *default* o autorquívamento, algumas instituições vêm desabilitando esta função no sistema para que o depósito seja feito por outros funcionários da instituição, normalmente bibliotecários. Em Portugal verificou-se que isto ocorre apenas em 19% dos repositórios institucionais, conforme se pode observar no gráfico

1, que indica que a maioria dos repositórios de Portugal está em consonância com os princípios da BOAI.

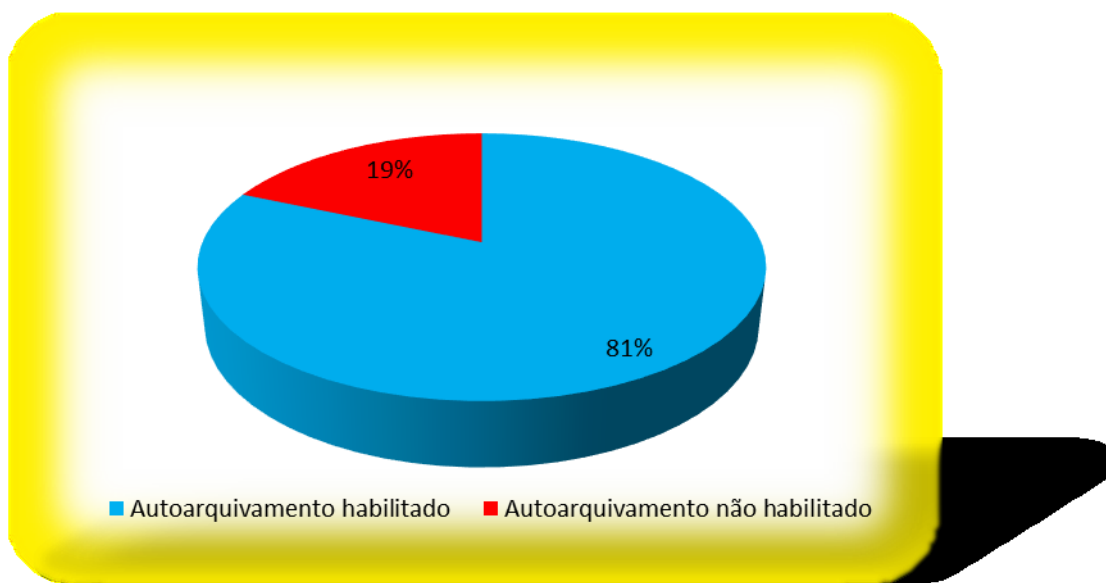


Gráfico 1 – Autoarquivamento habilitado nos Ris de Portugal (N=27)

Para os repositórios em que esta função está desabilitada perguntou-se o principal motivo para não habilitar o autoarquivamento no sistema (gráfico 2). A maioria dos gestores (60%) respondeu que os pesquisadores não o fariam. Estes gestores não acreditam que o autoarquivamento funcionaria com seus autores. Outros gestores creditam a não habilitação do autoarquivamento a questões de direitos de autor (20%), à uniformização dos conteúdos e à utilização pela instituição de outro sistema onde é dada a entrada da produção que envia os dados para o repositório (20%). Estes achados coincidem com os de Henning (2013) que verificou que os gestores de RIs em hospitais portugueses preferem fazer os depósitos por meio de seus bibliotecários do que deixar que os médicos autores o façam.

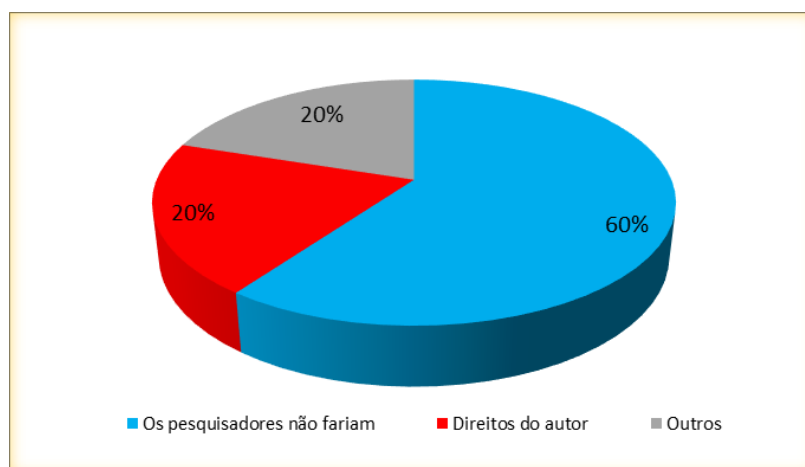
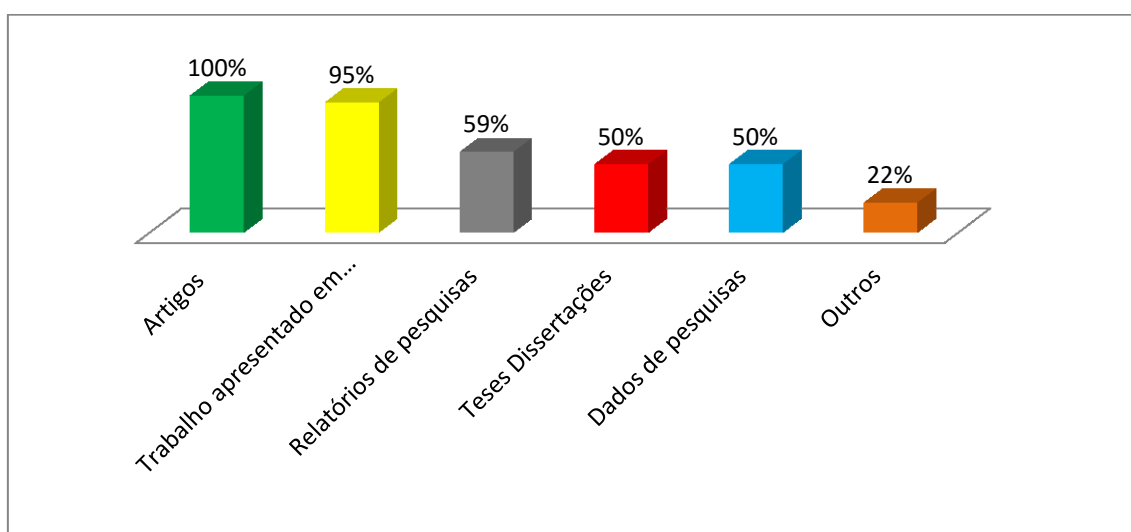


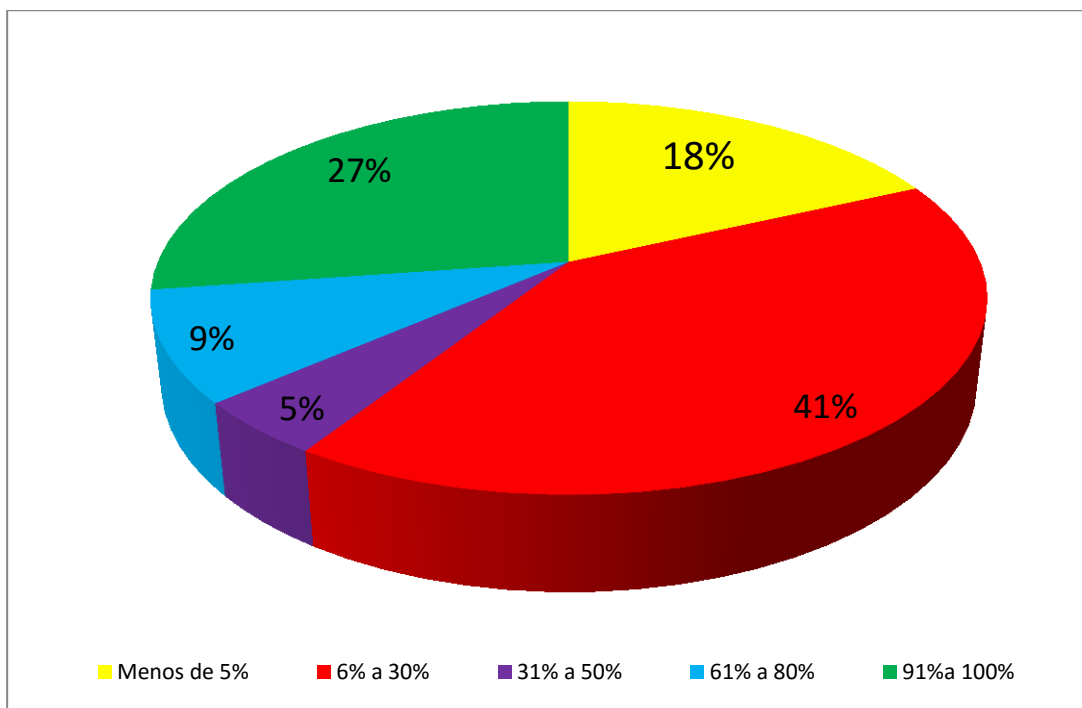
Gráfico 2 – Motivos para não habilitar o autoarquivamento nos RIs de Portugal (N=5)

Foi perguntado para os gestores que afirmaram que o autoarquivamento estava habilitado no RI em quais tipologias de documentos o autoarquivo está disponível (gráfico 3). Verificou-se que os artigos estão liberados para autoarquivamento em 100% dos repositórios que tem esta função habilitada no sistema. 95 % possibilitam o autoarquivamento para trabalhos apresentados em eventos e 59% para relatórios de pesquisa. Percebeu-se que em 50% dos repositórios que possibilitam o autoarquivamento, as dissertações e teses podem ser autoarquivadas e o mesmo percentual também disponibiliza a função para os dados de pesquisa. A tipologia de teses e dissertações tem tradição no mundo acadêmico e de pesquisa e muitas instituições de ensino recebem as teses nos repositórios através da biblioteca ou outro departamento da instituição, como a Secretaria acadêmica. Por outro lado, a tipologia de Dados de pesquisa não possui a mesma história de tratamento técnico nas bibliotecas, mas o seu acesso ganha cada vez mais importância no mundo científico. Enquanto muitos repositórios ao redor do mundo ainda se preparam para receber os dados de pesquisa nos repositórios, Portugal demonstra este pioneirismo quando possibilita o autoarquivamento de dados de pesquisa em grande número de seus repositórios. Acredita-se que as políticas governamentais estabelecidas auxiliam neste avanço, como as diretrizes da Comissão Europeia para o acesso aberto aos dados de pesquisa no Horizonte 2020 (COMISSÃO EUROPEIA, 2016).



**Gráfico 3** - Tipologias autorizadas para o autoarquivamento nos RIS de Portugal (N=22)

Entre os RIs com autoarquivamento habilitado 41 % dos gestores de repositórios afirmam que apenas 6 a 30% do material depositado foi autoarquivado. Isto demonstra que ainda há um caminho longo a percorrer, embora uma parcela ((27%) dos gestores de RIS com autoarquivamento habilitado, afirmar que 91% a 100% do material disponibilizado é fruto do autoarquivamento, conforme demonstrado no gráfico 4.



**Gráfico 4** – Percentual de documentos autoarquivados nos Ris de Portugal (N=22)

Estes dados coincidem com os achados de Migueis (2012) que confrontou seus dados com a investigação de Borges (2006). Migueis afirma que há indícios que um conjunto importante dos autores portugueses não deposita diretamente a sua produção no RI, o que é confirmado através da presente pesquisa.

Se juntarmos os percentuais que declaram autoarquivamento a menos de 5% e de 6% a 30% verificaremos que a maioria, 59%, declara que menos de 30% do material disponível no RI foi autoarquivado. Neste gráfico podemos verificar também que 36% dos RIs portugueses conseguem que a maioria dos registros disponibilizados no RI seja autoarquivado (61% a 100%).

### **Considerações finais:**

Este trabalho almejou obter um panorama do compartilhamento de artigos através de repositórios institucionais (RIs) em Portugal e a percepção do gestor de RIs quanto a esta forma de compartilhamento.

Concluiu-se que os gestores dos repositórios portugueses, em sua grande maioria, viabilizam o autoarquivamento de trabalhos por seus autores, habilitando esta função no sistema, o que demonstra um entendimento dos objetivos da criação de repositórios. Verificou-se também que os RIs Portugueses estão atentos para a importância do acesso aos dados de pesquisa na comunicação científica, preparando seus repositórios para o autoarquivamento desta tipologia. Quanto à adesão dos pesquisadores ainda é preciso

avançar. Ainda existe uma baixa adesão ao autoarquivamento pelos pesquisadores. É imprescindível envolver a comunidade científica no movimento de acesso aberto para uma ciência aberta, conhecer as especificidades das áreas, verificar as barreiras e os estímulos ao compartilhamento de artigos científicos e dados de pesquisa para, assim, ampliar a adesão ao autoarquivamento por parte dos pesquisadores em Portugal.

## Referências bibliográficas

BORGES, M. M. (2006) – *A esfera: comunicação académica e nova media*. Coimbra Faculdade de Letras da Universidade de Coimbra. 787 p. Tese de doutoramento. Disponível na internet: <URL: [https://estudogeral.sib.uc.pt/bitstream/10316/8557/5/MMBorges\\_tese.pdf](https://estudogeral.sib.uc.pt/bitstream/10316/8557/5/MMBorges_tese.pdf)>.

CROW, R. (2002). The case for institutional repositories: a SPARC position paper. *ARL* [Em linha] n° 223. [Consult. 04 set. 2016]. Disponível na internet: <URL: [http://www.sparc.arl.org/sites/default/files/media\\_files/instrepo.pdf](http://www.sparc.arl.org/sites/default/files/media_files/instrepo.pdf)>.

COMISSÃO EUROPEIA. *EC's Guide on Open Access to Scientific Publications and Research Data in Horizon 2020* (updated August 25, 2016). [Consult. 10 out. 2016]. Disponível na internet: <URL: [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)>.

FUNDAÇÃO PARA A CIÊNCIA E TECNOLOGIA (2014) – *Política sobre Acesso Aberto a Publicações Científicas resultantes de Projetos de I&D Financiados pela FCT*. Disponível na internet: <URL: [http://www.fct.pt/documentos/PoliticaAcessoAberto\\_Publicacoes.pdf](http://www.fct.pt/documentos/PoliticaAcessoAberto_Publicacoes.pdf)>.

HARNAD, Stevan. *Why the UK should not heed the Finch report* [Em linha] . London: The London School of Economics and Political Science. Department of Government. [Consult. 18 ago 2016] Disponível na internet: <URL: <http://blogs.lse.ac.uk/impactofsocialsciences/2012/07/04/why-the-uk-should-not-heed-the-finch-report/>>.

HENNING, Patrícia Corrêa (2013) – *Micro e macropolíticas de informação: o acesso livre à informação científica no campo da saúde no Brasil e em Portugal*. Rio de Janeiro: Curso de Informação e Comunicação em Saúde, Instituto de Comunicação e Informação Científica e Tecnológica em Saúde. (ICT/IOC/CRUZ) 234 p. Tese de doutoramento. Disponível na internet: <URL: [http://www.arca.fiocruz.br/bitstream/icict/6998/1/TESE\\_PATRICIA\\_HENNING.pdf](http://www.arca.fiocruz.br/bitstream/icict/6998/1/TESE_PATRICIA_HENNING.pdf)>.



LYNCH, C. A. (2003) – Institutional repositories: essential infrastructure for scholarship in the digital age. *ARL* [Em linha], nº 226. [Consult. 04 set. 2016]. Disponível na internet: <URL: <http://www.arl.org/storage/documents/publications/arl-br-226.pdf>>.

MEADOWS, A. J. (1999) – *A comunicação científica*. Brasília: Brinquet de Lemos/Livros. 268 p.

MIGUÉIS, A. M. E. (2012) – *Atitudes e percepções do autores depositantes do repositório científico da Universidade de Coimbra*, Coimbra: Universidade de Coimbra. 95 p. Dissertação. Disponível da Internet: <URL: <<https://estudogeral.sib.uc.pt/jspui/handle/10316/21116>>>.

RODRIGUES, Eloy; SARAIVA, Ricardo (2013) – RepositóriUM: 10 anos de Acesso Aberto ao Conhecimento. In: *Uma década de acesso aberto no UMinho e no mundo*. Braga : Publito. ISBN 978-989-98704-1-3. p. 25-47.

SARAIVA, Ricardo; RODRIGUES, Eloy (2010) – O Acesso Livre à literatura científica em Portugal: a situação actual e as perspectivas futuras. *Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas* [Em linha] nº 10. [Consult. 28 ago. 2016] . Disponível na internet: <URL: <http://bad.pt/publicacoes/index.php/congressosbad/article/view/192>>.

SWAN, A. (2008) – The business of digital repositories. In: *A DRIVER'S guide to european repositories*. Amsterdam: Amsterdam University Press. p. 15-48. [Consult. 04 set. 2016]. Disponível na internet: <URL: <http://dare.uva.nl/cgi/arno/show.cgi?fid=93898>>.

VEIGA, Viviane; MACENA, Luis Guilherme (2015) – O autoarquivamento nos repositórios institucionais brasileiros: um estudo exploratório. *Ponto de acesso* [Em linha] v. 9, n. 3 [Consult. 15 ago. 2016]. Disponível na internet: <URL: <http://www.portalseer.ufba.br/index.php/revistaici/article/view/15107/10302>>

---

## Acesso aberto, direitos autorais e a inclusão sociocultural das pessoas portadoras de deficiência

---

Allan Rocha de Souza <sup>i</sup>

Universidade Federal Rural do Rio de Janeiro

[allansouza@gmail.com](mailto:allansouza@gmail.com)

Alexandre de Serpa Pinto Fairbanks<sup>ii</sup>

Universidade Federal Rural do Rio de Janeiro

[alexandre\\_spf@hotmail.com](mailto:alexandre_spf@hotmail.com)

Wemerton Monteiro Souza <sup>iii</sup>

Universidade Federal Rural do Rio de Janeiro

[wemertonmonteiro@hotmail.com](mailto:wemertonmonteiro@hotmail.com)

### Resumo

Em 2008, o Brasil ratificou a Convenção Internacional sobre os Direitos das Pessoas com Deficiência capitaneada pela Organização das Nações Unidas. Como mecanismo de efetivação da Convenção, foi promulgada a lei n. 13.146/15 para a ampla inclusão de pessoas com deficiência. Ainda em 2015, foi ratificado pelo Brasil o Tratado de Marraqueche, concluído na OMPI, que busca garantir a disponibilização de textos para portadores de deficiência visual. Ambos tratados foram internalizados com status de emenda constitucional. Diante desta nova realidade jurídica brasileira, torna-se imperioso averiguar quais seus efeitos sobre a comunicação científica e o acesso aberto, uma vez que está legalmente garantido o acesso a bens culturais e científicos em formato acessível. Os seus efeitos alcançam, dentre outros aspectos, as políticas e ações institucionais de acesso aberto a produção científica e, neste âmbito, exige-se que, além de perseguirem os próprios objetivos do movimento de acesso aberto, disponibilizem este material também em formato acessível às pessoas portadoras de deficiência. Neste sentido, vemos que, mesmo diante da fragilidade institucional do acesso aberto no Brasil, um novo desafio se apresenta para o movimento, que deverá não só institucionalizar-se como avançar na consolidação do acesso aberto de obras em formato acessível.

**Palavras-Chave:** Acesso Aberto; Direitos Autorais; Portadores de Deficiência; Tratado de Marraqueche.

## Open access, copyright and the sociocultural inclusion of persons with disabilities

### Abstract

In 2008, Brazil ratified the International United Nations Convention on the Rights of Persons with Disabilities. As a mechanism to make the Convention effective the law n. 13,146/15 was enacted, calling for the broad inclusion of people with disabilities. In 2015, Brazil ratified the Marrakesh Treaty, completed at WIPO, which seeks to ensure the availability of texts for the visually impaired. Both treaties were internalized as constitutional amendments. Faced with this new Brazilian legal reality, it is imperative to ascertain what are its effects on scientific communication and open access, since it is legally guaranteed the right of access to cultural and scientific works in accessible format. Its effects reach, among other things, institutional policies and actions of open access to scientific production and, in this context, it is required that, whilst pursuing their own goals, the open access movement must also make available the works in accessible format for people with disabilities. In this sense, we see that even with the institutional weakness of open access in Brazil, a new challenge presents itself to the movement, which should not be institutionalized but move forward towards the consolidation of open access in an accessible format.

**Keywords:** Open Access; Copyright; Persons with Disabilities; Marrakesh Treaty.

### Introdução

Em 2008, o Brasil ratificou a Convenção Internacional sobre os Direitos das Pessoas com Deficiência e seu Protocolo Facultativo<sup>iv</sup> capitaneado pela Organização das Nações Unidas. Como mecanismo de efetivação da Convenção, foi promulgada a lei federal n. 13.146 para a ampla inclusão de pessoas com deficiência, que entrou em vigor a partir de 04 de janeiro de 2016. Já em 2015, foi ratificado pelo Brasil o Tratado de Marraqueche, concluído na Organização Mundial da Propriedade Intelectual<sup>vi</sup>, que visa a disponibilização de textos em formatos acessíveis para pessoas com deficiência visual.

Ao versarem sobre direitos humanos e terem obtido aprovação de três quintos dos membros de cada casa do Congresso Nacional, ambos tratados foram internalizados com status de emenda constitucional, conforme procedimento estabelecido no § 3<sup>o</sup><sup>vii</sup> do art. 5<sup>o</sup> da Constituição Federal de 1988. A ratificação dos tratados e a promulgação do Estatuto dos Deficientes foram importantes passos no respeito e valorização da dignidade, liberdade e autonomia individual, pois buscam a máxima inclusão social, cultural, científica, educacional e tecnológica das pessoas portadoras de deficiência.

Diante desta nova realidade jurídica brasileira, torna-se imperioso averiguar quais seus efeitos sobre a comunicação científica, uma vez que está legalmente garantido o acesso

a bens culturais e científicos em formato acessível. Busca-se neste trabalho, em especial, averiguar os efeitos sobre os direitos autorais e sobre as ações e políticas de acesso aberto e as instituições que as conduzem no Brasil.

Trata-se de estudo a ser realizado mediante pesquisa teórica, documental e bibliográfica. O método de abordagem será hipotético-dedutivo para analisar as proposições teóricas e método dialético para vislumbrar os contornos e efeitos sobre o acesso aberto. Serão também empregados os métodos de procedimento histórico para avaliar a trajetória e justificativas dos tratados e funcional para revelar o papel que busca satisfazer dentro do sistema de proteção autoral.

Assim, o trabalho será desenvolvido com o objetivo de apresentar os novos desafios que surgem com a obrigatoriedade de disponibilização em formato digital, que se somam aos já existentes. Inicia-se assim com uma breve apresentação da análise sobre os repositórios brasileiros. Em seguida enfrenta-se a questão dos Tratados de inclusão das pessoas com deficiência, para adiante encarar os efeitos destes instrumentos internacionais sobre os direitos autorais, para, enfim, indicar os efeitos sobre o movimento de acesso aberto.

### **A implantação do acesso aberto no Brasil: breves considerações**

Entre as formas de licenças abertas, o acesso aberto ao material científico produzido pelas universidades e instituições públicas foi o que mais avançou, mas, ainda assim, não logrou êxito em face da legislação nacional.

Muito fora articulado e debatido sobre o movimento de acesso aberto ao conhecimento científico produzido pelas universidades e instituições públicas brasileiras, e ferramentas foram criadas pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) para incentivar a implementação desse movimento nas instituições de pesquisa. Após toda essa agitação, fora protocolado na Câmara dos Deputados o Projeto de Lei 1.120 de 2007, de autoria do Deputado Federal Rodrigo Rollemberg (PSB/DF). Tal projeto visava obrigar as instituições públicas de ensino superior a construir os repositórios institucionais para depósito de toda produção técnico-científica do corpo docente e discente. No entanto, tal projeto não vingou e foi arquivado.

Em 2011, como Senador Federal da República, Rodrigo Rollemberg submeteu um projeto semelhante no Senado (Projeto de Lei do Senado nº 387, de 2011). A passos lentos o projeto ainda tramita no Senado Federal, estando, desde 26/02/2015, na CCJ – Comissão de Constituição, Justiça e Cidadania –, pronto para a pauta na comissão. Contudo, a matéria já tem o voto contrário do Senador Álvaro Dias que pediu a rejeição do projeto de lei.

Buscando identificar e analisar os repositórios institucionais brasileiros e suas políticas institucionais de acesso aberto, realizou-se um levantamento de dados dos diretórios internacionais (*Registry of Open Access Repositories* – ROAR, *The Directory of Open Access Repositories* – openDOAR, *Repositório Científico de Acesso de Portugal* – RCAAP

e o *Registry of Open Access Repository Mandates and Policies* – ROARMAP) e do portal nacional do *Instituto Brasileiro de Informação em Ciência e Tecnologia* –IBICT.

Os dados colhidos foram analisados qualitativamente e as políticas institucionais foram classificadas<sup>1</sup> como *satisfatórias*, *significantes* e *inviáveis*, tendo como parâmetro algumas recomendações presentes na BOAI: (I) *A Política Institucional de Acesso aberto é mandatória?*, (II) *A Política orienta os autores a realizar o depósito no momento da aceitação para publicação?*, (III) *A Política permite o auto-depósito aos pesquisadores e acadêmicos membros da instituição?*, (IV) *Há necessidade de revisão por pares antes do material ser depositado?*, (V) *A política respeita a liberdade dos acadêmicos de submeter seus trabalhos em outras revistas?* (VI) *A licença utilizada pelo repositório é a CC-BY ou equivalente?*

O resultado dos dados analisados e interpretados estão dispostos nas tabelas seguintes:

Repositórios	nº	%
<b>Nº total de repositórios diferentes identificados</b>	<b>107</b>	
Nº total de repositórios que possuem algum tipo de política pública de Acesso Aberto	63	58,88% <sup>1</sup>

Quadro 1:

1. Número relativo frente ao número total de repositórios diferentes identificados.

Instituições	nº	%
<b>Nº total de Instituições que possuem repositório</b>	<b>86</b>	
Nº de universidades pública e privadas	54	62,79% <sup>1</sup>
Nº de institutos de pesquisa	13	15,11% <sup>2</sup>
Nº de instituições de outros tipos	19	22,1% <sup>3</sup>

Quadro 2:

1. Número relativo frente ao número total de instituições que possuem repositórios.
2. Número relativo frente ao número total de instituições que possuem repositórios.
3. Número relativo frente ao número total de instituições que possuem repositórios.

Instituições	nº	%
<b>Nº total de instituições que possuem repositórios com políticas públicas de Acesso Aberto</b>	<b>51</b>	

Quadro 3

Políticas	nº	%
-----------	----	---

<sup>1</sup>Após a análise e comparação minuciosa das políticas institucionais de acesso aberto, as políticas foram qualificadas em três grandes grupos: I) satisfatória, que apresentou em seu texto 5 (cinco) ou 6 (seis) dos requisitos listados; II) significativa, que cumpriu com pelo menos 4 (quatro) dos 6 (seis) requisitos; III) inviáveis, políticas que cumpriram somente 2 (dois) ou nenhum dos requisitos.

<b>Nº total de Políticas Institucionais de Acesso Aberto diferentes identificadas</b>	<b>63</b>	
Nº de políticas públicas de Acesso Aberto classificadas como SATISFATÓRIA	27	42,86% <sup>1</sup>
Nº de políticas públicas de Acesso Aberto classificadas como SIGNIFICANTES	7	11,1% <sup>2</sup>
Nº de políticas públicas de Acesso Aberto classificadas como INVIÁVEIS	29	46,04% <sup>3</sup>

Quadro 4:

1. Número relativo frente ao número de políticas institucionais diferentes.
2. Número relativo frente ao número de políticas institucionais diferentes.
3. Número relativo frente ao número de políticas institucionais diferentes.

Conforme observado nas tabelas acima, o número total de repositórios institucionais de acesso aberto identificado foi de 107 (cento e sete) repositórios, que estão dispostos em 86 (oitenta e seis) instituições de pesquisa. Contudo, esse número de instituições de pesquisa identificadas é relativamente pequeno frente ao número total de instituições de pesquisa, públicas e privadas, existentes no Brasil.

Em relação as políticas institucionais de acesso aberto que foram classificadas como satisfatórias é tímido perante o número total de repositórios identificados. Desse modo, pôde-se concluir que as políticas institucionais realmente não se encontram desenvolvidas ao nível necessário para serem sustentáveis.

Assim, é evidente que nenhuma iniciativa relativa ao movimento de acesso aberto ao conhecimento científico foi alcançada na legislação pátria e tal movimento no Brasil está distante do desejado pelo movimento internacional, bem como pela BOAI – *Budapest Open Access Initiative*.

### A Convenção de Nova York e o Tratado de Marraqueche

O Tratado de Marraqueche, primeiro documento internacional a estabelecer limitações mandatórias aos direitos autorais, entrou em vigor mundialmente em 30 de setembro de 2016, e foi ratificado pelo Brasil em 2015, com status de Emenda Constitucional. Sua finalidade é assegurar aos deficientes visuais acesso aos textos em igualdade de condições.

Em parecer da Comissão das Pessoas com Deficiência (CPD), em sua tramitação na Câmara dos Deputados, é salientado que a discriminação e exclusão histórica sofrida pelas pessoas com deficiência visual e outras deficiências que interferem leitura, em decorrência da escassez na produção e distribuição das obras em formato acessível, ressaltando que as pessoas com deficiência não querem privilégios ou tratamento diferenciado, mas “almejam, na verdade, que a sociedade lhes proporcione condições para o exercício de seus direitos de cidadania em igualdade de condições com as demais pessoas.”<sup>viii</sup>

O segundo parecer, de relatoria do Deputado Federal Leo de Brito, foi apresentado no dia 29 de maio de 2015 à Comissão de Cultura (CCULT). O parlamentar votou e sugeriu a aprovação do texto do Tratado de Marraqueche que, acima de tudo, "reconhece o direito das pessoas com deficiência em participar da vida cultural em igualdade com as demais pessoas"<sup>ix</sup>. O relator ressaltou a "notória relevância de livros na divulgação de informação e cultura"<sup>x</sup>, afirmando ser o objetivo primordial do Tratado o combate a "fome de livros"<sup>xi</sup> ocasionada pela falta ou a restrição de acesso aos materiais impressos pelos deficientes visuais, que intensifica "as restrições sociais e econômicas que pessoas com deficiência enfrentam, gerando uma situação de exclusão socioeconômica"<sup>xii</sup>.

Após aprovação na Câmara, já no Senado, a relatora Senadora Marta Suplicy sugeriu a aprovação do referido tratado com status de Emenda à Constituição, com intuito de dar maior efetividade ao Tratado,<sup>xiii</sup> cujo objetivo principal é promover a disponibilização das obras em formato acessível para que os deficientes visuais possam ter acesso "à leitura, à educação, ao desenvolvimento pessoal e ao trabalho em igualdade de oportunidades."<sup>xiv</sup>

Antes do Tratado de Marrakesh, com o Decreto nº 6.949 de 25 de agosto de 2009, o país promulgou a ratificação como Emenda Constitucional da Convenção da Organização das Nações Unidas (ONU) sobre os Direitos das Pessoas com Deficiência, conhecida como Convenção de Nova York, bem como seu protocolo facultativo. A Convenção de Nova York, que objetiva estabelecer e promover a igualdade de direitos em todas as esferas para as pessoas com deficiência, fora ratificada também com status de Emenda Constitucional.

Suas finalidades e vinculação com o Tratado de Marraqueche é explicitada no Parecer do Deputado Aeilton de Freitas na CPD na Câmara dos Deputados, quando ressaltou que a ratificação importa na adequação do Brasil aos princípios fundantes da Convenção sobre os Direitos das Pessoas com Deficiência, que são os "os princípios da não-discriminação; do respeito pela dignidade inerente; pela autonomia individual, inclusive a liberdade de fazer as próprias escolhas, e pela independência das pessoas; da plena e efetiva participação e inclusão na sociedade; da igualdade de oportunidades e da acessibilidade."<sup>xv</sup>

Relacionada, portanto, ao Tratado de Marraqueche, em razão de seu conteúdo, a Convenção de Nova York, estabelece a obrigação de, entre outros, garantir a disponibilidade de bens culturais, científicos e tecnológicos em formatos acessíveis, e, nesse sentido, estabelece deveres que vão além das metas restritas do Tratado de Marraqueche, uma vez que não se limita nem ao material impresso nem em benefício unicamente das pessoas com deficiência visual, mas alcança, basicamente, pessoas portadoras de qualquer deficiência e todos os tipos de expressões culturais.

Após a aprovação da Convenção de Nova York, sua efetivação ocorreu com a promulgação de lei federal para a ampla inclusão de pessoas com deficiência, que entrou em vigor a partir de 04 de janeiro de 2016, atingindo os domínios cultural e tecnológico. A Lei n. 13.146, de 2015, estabelece, em seu artigo 42, garantias com relação ao direito de acesso aos produtos culturais em formatos acessíveis. E, ainda mais interessante, em seu parágrafo

1º, afirma que "é vedada a recusa de oferta de obras intelectuais em formatos acessíveis às pessoas com deficiência, sob qualquer argumento, incluindo a alegação de proteção dos direitos de propriedade intelectual."

Este é um importante passo no respeito e na valorização da dignidade, liberdade e autonomia individual, pois, intentando a máxima inclusão dos deficientes, acaba por forçar os entes federativos a criar políticas públicas capazes de reprimir a discriminação e fomentar um progresso social isonômico que permita, não de forma meramente formal, o pleno desenvolvimento e inclusão sócio-cultural de todos.

### **Os direitos autorais no novo contexto jurídico**

Os efeitos dos Tratados sobre o sistema de direitos autorais têm diversas dimensões. No plano mais abstrato, a aceitação destes Tratados como Emenda Constitucional implica no reconhecimento da equivalência dos direitos de acesso à cultura, informação e conhecimento aos direitos autorais, uma vez que são direitos fundamentais de igual importância. Isto impõe a necessidade de um maior equilíbrio na harmonização da proteção autoral e o direito de acesso.

Os efeitos imediatos a sobre os direitos autorais são principalmente de duas ordens. Em primeiro lugar substancia a noção de que as limitações e exceções são não só necessárias, mas obrigatórias, em razão do seu papel de equilibrar e harmonizar dentro do sistema de direitos autorais os interesses proprietários dos titulares e não proprietários da coletividade. Além disso, consolida a interpretação extensiva das limitações expressas, já adotada pelo STJ, uma vez que estas são representações dos direitos fundamentais que devem ser compatibilizados com os direitos autorais. É expressão da função social dos direitos autorais. A ratificação dos Tratados implica na consolidação da interpretação extensiva das limitações.

Por fim, as normas dos Tratados impactam diretamente os negócios jurídicos de direitos autorais, cujos termos não podem conflitar com os interesses coletivos relevantes e constitucionalmente estabelecidos. A necessidade dos contratantes atentarem na elaboração do programa contratual para os interesses de terceiros e da coletividade encontra fulcro na função social dos contratos. Neste sentido já caminha o artigo 42 do Estatuto das Pessoas com Deficiência, quando proíbe a não disponibilização das obras em formato acessível. Os efeitos imediatos destes Tratados sobre os negócios jurídicos de direitos autorais são, portanto, dois: (i) tornam nulas as cláusulas contratuais proibitivas de disponibilização das obras em formato acessível às pessoas com deficiência; (ii) quando ausentes, são consideradas implícitas as permissões contratuais de disponibilização das obras em formato acessível às pessoas com deficiência.

### **Considerações finais: os novos desafios para o acesso aberto**

Dentre as conclusões desta investigação nota-se, em primeiro lugar, a obrigação imposta aos entes federativos de criarem políticas públicas capazes de combater a



discriminação e fomentar um progresso social isonômico das pessoas portadoras de deficiência.

Pela força normativa destes tratados, impõe-se também uma modificação no sistema de direitos autorais, com ampliação das limitações à exclusividade de forma a possibilitar, sem necessidade de autorização e remuneração, a disponibilização dos bens culturais, artísticos e científicos às pessoas portadoras de deficiência em formato acessível. Sua incorporação ao ordenamento jurídico brasileiro altera o conteúdo das limitações e promove uma revisão sistêmica da proteção aos direitos autorais, na busca por maior equilíbrio interno.

Os seus efeitos alcançam também as instituições, públicas e privadas, de produção e comunicação científica, artística ou cultural, pois as sujeita a disponibilização de sua produção em formato acessível aos portadores de deficiência. Alcança-se, por fim, as políticas e ações institucionais de acesso aberto à produção científica e, neste âmbito, exigem-se que, além de perseguirem os próprios objetivos do movimento de acesso aberto, disponibilizem este material também em formato acessível às pessoas portadoras de deficiência.

#### Referências bibliográficas:

AMARO, Bianca; LABBÉ, Gloria Carmen; LISOWSKA, Malgorzata e NAKANO, Silvia. *Rede Federada de Repositórios Institucionais de Publicações Científicas em Acesso Aberto LA Referencia: a integração da produção científica regional*. In: Eloy Rodrigues, Alma Swan e AnaAlice Baptista (editores). *Uma Década de Acesso Aberto na UMinho e no Mundo*. Braga: Universidade do Ninho, 2013, p. 123-132. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL:https://repositorium.sdum.uminho.pt/bitstream/1822/27502/1/RepositoriUM\_10anos.pdf>.

ASCENSÃO, José de Oliveira. *Direito da internet na sociedade de informação*. Rio de Janeiro: Forense, 2002.

ASCENSÃO, José de Oliveira. O Direito Intelectual em Metamorfose. In *Revista de Direito Autoral*, ano II, n. 4. Rio de Janeiro: Lumen Juris, 2006.

BABINI, D. The risk of open access becoming integrated into existing commercial publishing – the need of a global system of noncommercial open access scholarly communications. *RevEletron de ComunInflnov Saúde* [Internet]. Out-dez 2014; 8(4):438-442. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL:http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/982.en>.

COSTA, Suely Maria de Souza; KURAMOTO, Hélio, e LEITE, Fernando L. M. *Acesso Aberto no Brasil: aspectos históricos, ações institucionais e panorama atual*. In: Eloy Rodrigues, Alma

Swan e Ana Alice Baptista (editores). *Uma Década de Acesso Aberto na UMinho e no Mundo*. Braga: Universidade do Ninho, 2013, p. 133–150. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL: <https://bibliotecadigital.ipb.pt/bitstream/10198/8982/1/Biblioteca%20Digital%20do%20IPB.pdf>>.

GEIGER, Christophe. Constitutionalising intellectual property law? The influence of fundamental rights on intellectual property in the European Union. In: *International Review of Intellectual Property and Competition Law*, vol. 37, n. 4, Munique, Alemanha: Max Planck Institute for Intellectual Property, Competition and Tax Law, 2006.

GOMES, S. L. R. *O Acesso Aberto ao conhecimento científico: o papel da universidade brasileira*. RECIIS – Rev. Eletron. De Comum. Inf. Inov. Saúde. Jun 2014; 8(2) – p. 93 –106. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL: DOI: 10.3395/reciis.v8i2.936.pt>.

KURAMOTO, H. *Acesso livre à informação científica: novos desafios*. Liinc em Revista, v.4, n.2, setembro 2008, Rio de Janeiro, p. 154 – 157. [Consult. 09 Ago. 2016] Disponível na Internet: <URL: <http://repositorio.ibict.br/bitstream/123456789/154/1/Kuramoto,%20H.pdf>>.

KURAMOTO, Hélio. *Acesso livre: caminho para maximizar a visibilidade da pesquisa*. Revista de Administração Contemporânea. Curitiba, v. 12, n. 3, p. 861–872, jul./set. 2008. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL: <http://www.redalyc.org/articulo.oa?id=84012313>>.

KURAMOTO, Hélio. *Acesso Livre: uma solução adotada em todo o globo; porém, no Brasil parece existir uma indefinição*. RECIIS – Revista Eletrônica de Comunicação Informação e Inovação em Saúde, 2014, vol. 8, n. 2, p. 166–179. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL: <http://www.reciiis.icict.fiocruz.br/index.php/reciis/article/view/630/1270>>.

RODRIGUES, Eloy, e SARAIVA, Ricardo. *RepositóriUM: 10 anos de Acesso Aberto ao Conhecimento*. In: Eloy Rodrigues, Alma Swan e Ana Alice Baptista (editores). *Uma Década de Acesso Aberto na UMinho e no Mundo*. Braga: Universidade do Ninho, 2013, p. 25–48. [Consult. 09 Ago. 2016] Disponível na Internet: <URL: [https://repositorium.sdum.uminho.pt/bitstream/1822/27502/1/RepositoriUM\\_10anos.pdf](https://repositorium.sdum.uminho.pt/bitstream/1822/27502/1/RepositoriUM_10anos.pdf)>

RODRIGUES, Eloy. *O Acesso Aberto (na UMinho e no mundo): onde estamos e onde vamos?* RECIIS – Revista Eletrônica de Comunicação Informação e Inovação em Saúde, 2014, vol. 8, n. 2., p. 180–194. [Consult. 09 Ago. 2016]. Disponível na Internet: <URL: <http://www.reciiis.icict.fiocruz.br/index.php/reciis/article/view/631/1271>>.

SOUZA, Allan Rocha de. *A Função Social dos Direitos Autorais: uma interpretação civil-constitucional das limitações*. Rio de Janeiro: Editora da Faculdade de Direito de Campos, 2006.

SOUZA, Allan Rocha. *Direitos Culturais no Brasil*. Rio de Janeiro: Azougue Editorial, 2012.

SOUZA, Allan Rocha; ALMEIDA Jr., Vitor de Azevedo; SOUZA, Wemerton Monteiro. *Os direitos autorais na perspectiva civil-constitucional*. In Revista Brasileira de Direito Civil. Vol. 8. Abr-jun. Instituto Brasileiro de Direito Civil: Curitiba, PR, 2016, pp. 9-31.

---

<sup>i</sup> Professor e Pesquisador de Direito Civil no Departamento de Direito, Humanidades e Letras, Instituto Três Rios/UFRRJ. Professor e Pesquisador em Direitos Autorais e Políticas Culturais no Programa de Pós-Graduação em Políticas Públicas, Estratégia e Desenvolvimento (PPED) na UFRJ. E-mail: allansouza@gmail.com . CV Lattes: <http://lattes.cnpq.br/5178459691896082>

<sup>ii</sup> Acadêmico do Curso de Direito do Instituto Três Rios/UFRRJ. Pesquisador do Núcleo de Pesquisa em Direitos Fundamentais, Relações Privadas e Políticas Públicas. E-mail: alexandre\_spf@hotmail.com. CV Lattes: <http://lattes.cnpq.br/8954789495709084>

<sup>iii</sup> Acadêmico do Curso de Direito do Instituto Três Rios/UFRRJ. Pesquisador do Núcleo de Pesquisa em Direitos Fundamentais, Relações Privadas e Políticas Públicas. E-mail: wemertonmonteiro@hotmail.com CV Lattes:

<sup>iv</sup> BRASIL. Decreto Federal nº 6.949. Ratifica o Tratado da ONU sobre os Direitos das Pessoas com Deficiência. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2007-2010/2009/decreto/d6949.htm](http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/decreto/d6949.htm)>. Acesso em: 13 abril de 2016.

<sup>v</sup> BRASIL. Lei n. 13.146, de 06 de julho de 2015. Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Disponível em <[http://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2015/Lei/L13146.htm](http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2015/Lei/L13146.htm)> Acesso em 15 de abril de 2016.

<sup>vi</sup> Organização Mundial de Propriedade Intelectual. Tratado de Marraqueche. Disponível em<<http://www.wipo.int/treaties/en/ip/marrakesh/>> Acesso em 15 de abril de 2016.

<sup>vii</sup> BRASIL. Constituição Federal da República do Brasil. Art. 5º § 3º Os tratados e convenções internacionais sobre direitos humanos que forem aprovados, em cada Casa do Congresso Nacional, em dois turnos, por três quintos dos votos dos respectivos membros, serão equivalentes às emendas constitucionais.

<sup>viii</sup> Ibid. p. 6.

<sup>ix</sup> BRASIL. Câmara dos Deputados. Parecer da Comissão de Cultura, de relatoria do Deputado Federal Leo de Brito, p.4. Disponível em: < <https://www2.camara.leg.br/>>. Acesso em: 20 de fev. de 2016.

<sup>x</sup> Ibid. p. 3.

<sup>xi</sup> Expressão também utilizada no parecer do Deputado Federal Aelton Freitas à Comissão de Defesa dos Direitos da Pessoa com Deficiência.

<sup>xii</sup> Ibid. p. 4.

<sup>xiii</sup> BRASIL. Senado Federal. Parecer da Comissão de Relações Exteriores de Defesa Nacional, de relatoria da Senadora Marta Suplicy, p.1. Disponível em: <<https://www.senado.gov.br/>>. Acesso em: 18 de fev. de 2016,

<sup>xiv</sup> Ibid. p. 2.

<sup>xv</sup>BRASIL. Câmara dos Deputados. Parecer da Comissão de Defesa dos Direitos da Pessoa com Deficiência, de relatoria do Deputado Federal Aelton Freitas, pp.5-6. Disponível em: < <https://www2.camara.leg.br/>>. Acesso em: 20 de maio de 2016.

---

## O retrato situacional das revistas científicas brasileiras

---

Danyelle Mayara Silva

Instituto Federal de Educação, Ciência e Tecnologia de Brasília (IFB)

[danyelle.silva@ifb.edu.br](mailto:danyelle.silva@ifb.edu.br)

Teila de Oliveira Carvalho

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

[teilacarvalho@ibict.br](mailto:teilacarvalho@ibict.br)

Bianca Amaro

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

[bianca@ibict.br](mailto:bianca@ibict.br)

Tainá Batista de Assis

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

[taina@ibict.br](mailto:taina@ibict.br)

### Resumo

O Brasil tem ocupado posições centrais em índices de acesso aberto e em indexadores de revistas científicas, desempenhando um importante protagonismo no desenvolvimento do acesso aberto na América Latina. O país se destaca em razão de ações bem-sucedidas na promoção de infraestruturas que permitem o desenvolvimento deste tipo de canal de comunicação científica livre de custos e barreiras de acesso aos seus leitores. Neste sentido, este trabalho buscou apresentar o quadro situacional das revistas científicas brasileiras de acesso aberto com base em sua presença em índices e indexadores de grande relevância. Com base nos resultados, observou-se que a presença das revistas científicas brasileiras em diferentes diretórios e indexadores, nacionais e internacionais, reflete de forma positiva as estratégias adotadas no país para a promoção de revistas científicas em acesso aberto. O trabalho que vem sendo desenvolvido dia-a-dia nesse contexto e a consequente maior visibilidade dessas publicações científicas demonstram que o Brasil mantém a caminhada para o alcance de maiores níveis e presença em diretórios e indexadores nacionais e internacionais no mundo do Acesso Aberto.

**Palavras-chave:** Periódicos Científicos, Produção Científica, Acesso Aberto, Indexadores, Internacionalização, América Latina

## **The situational frame of brazilian scientific journals**

### **Abstract**

Brazil has occupied central positions in open access indexes of scientific journals, playing an important role in the development of open access in Latin America. The country stands out due to successful actions to promote infrastructure that enable the development of this type of free scientific communication channel costs and barriers to access to your readers. Thus, this study aimed to present the situational context of Brazilian scientific journals open access based on their presence on indexes and indexes of great importance. Based on the results, it was observed that the presence of Brazilian scientific journals in different directories and indexes, national and international, reflects positively the strategies adopted in the country for the promotion of scientific journals in open access. The work that is being developed day by day in this context and the resulting increased visibility of these scientific publications show that Brazil has to walk to reach higher levels and presence in directories and national and international databases in the world of Open Access.

**Keywords:** Scientific Journals, Scientific Production, Open Access, Indexes, Internationalization, Latin America

## **Introdução**

A ciência está toda baseada na comunicação das pesquisas realizadas no âmbito da comunidade científica. Essa comunicação é necessária para que as pesquisas atinjam o objetivo de gerar novos conhecimentos e alcançar uma ampla gama de pesquisadores interessados no mesmo assunto. Além disso, segundo Oliveira (2005), o desenvolvimento científico está fundamentado na divulgação, leitura, assimilação e utilização de estudos anteriores, mas para que isso aconteça é preciso que as informações sejam divulgadas e estejam acessíveis aos pesquisadores. Por isso são utilizados meios de comunicação que viabilizam a disseminação e o acesso dos resultados das pesquisas em nível nacional e internacional.

Os periódicos assumem um papel fundamental nesse processo de propagação da ciência e podem ser considerados o principal canal de publicação científica. Para Kuramoto (2006), a informação científica é o fator primordial para o avanço científico e técnico de um país e essas informações, que decorrem dos resultados das pesquisas científicas, devem ser divulgadas por meio das revistas para a comunidade. Por isso, cada país busca uma forma de propagar a suas pesquisas, planejando se inserir em uma ciência desenvolvida e de grande repercussão para agregar valor e dar visibilidade a suas descobertas.

Porém, destacar-se em âmbito internacional não é uma realidade fácil de ser alcançada pelos países em desenvolvimento ou emergentes. A grande distribuição das revistas científicas dos países participantes do considerado “*mainstream*” da ciência costuma ultrapassar as publicações oriundas dos demais países, ficando em destaque no ambiente científico.

Alperin, Fischman e Willinsky (2008), no entanto, acreditam que a difusão das tecnologias de informação e comunicação (TIC), principalmente na América Latina, possibilitou que um número maior de cientistas tivesse acesso à informação de forma mais fácil e incentivou os pesquisadores a participarem e colaborarem mais efetivamente com as redes internacionais. Além disso, facilitou o desenvolvimento de revistas científicas permitindo a disseminação mais rápida das descobertas científicas no meio eletrônico. Fraga (2008) inclusive aponta que nos últimos anos a produção científica na América Latina teve um aumento considerável e que nesse avanço destacam-se principalmente os pesquisadores brasileiros em razão da quantidade de artigos publicados.

De forma geral, o Brasil tem tido um grande desenvolvimento não só na quantidade de artigos publicados por seus pesquisadores, mas também no número de revistas científicas criadas nos últimos anos. Mugnaini, Digiampietri, e Mena-Chalco (2014) e Fiorin (2007) consideram que as causas do crescimento da produção científica no país e seu melhor posicionamento no cenário mundial são os investimentos continuados no desenvolvimento da pós-graduação brasileira e em pesquisas, auxiliando na formação e aprimoramento de seus pesquisadores, a avaliação dos programas de pós-graduação realizada pela Capes, assim como as exigências para concessão de bolsas de Produtividade em Pesquisa do CNPq. Esses incentivos nacionais, ainda segundo os autores, têm estimulado o aumento da produtividade e a divulgação dos resultados das pesquisas por parte dos cientistas brasileiros, além de induzir a criação e o desenvolvimento de revistas científicas nacionais. Por essa razão, os periódicos científicos brasileiros têm se destacado no meio editorial pela grande quantidade que estão surgindo principalmente dentro das universidades.

Porém, sabe-se que apenas criar uma revista não garante que esta será de qualidade e alcançará o público desejado. Para desenvolver-se dentro da comunidade científica nacional e internacionalmente, as revistas procuram meios de divulgação que facilitem o acesso e promovam a visibilidade das suas publicações.

As bases de dados têm assumido esse papel de certificar que as revistas estarão acessíveis a um público maior e também de facilitar a localização dos artigos. Além disso, a indexação das revistas em bases de dados, especificamente nas de impacto internacional, possibilita a inclusão da produção científica em um sistema mundial, garantindo a sua internacionalização. Essa internacionalização é, segundo Fiorin (2007), um dos critérios mais importantes para averiguar a excelência da produção científica.

É por isso que, segundo Fraga (2008), o número de bases de dados em que os periódicos brasileiros estão sendo indexados tem crescido de maneira significativa, isso porque os pesquisadores têm procurando internacionalizar o conhecimento científico gerado no Brasil. Para o autor, os periódicos brasileiros estão buscando os melhores e mais atuais meios de divulgação e acesso para garantir um maior impacto na sociedade, serem conhecidos e terem seus conteúdos citados. Além disso, as bases podem ser consideradas fontes de informação de qualidade por, em sua grande maioria, possuir critérios que exigem que as publicações mantenham um padrão de qualidade para serem indexadas e permanecerem nelas.

Além do papel desempenhado pelas bases de dados no processo de indexação das revistas científicas, é indiscutível que o acesso aberto às publicações científicas também tem destaque fundamental no acesso, divulgação e visibilidade das pesquisas. O modelo de acesso aberto considera que as pesquisas científicas são bens de interesse público, especialmente se tiverem sido financiadas por meio de fundos públicos, e devem estar acessíveis sem nenhum tipo de barreira para acesso a todos aqueles que desejarem e se interessarem pelos conteúdos. O acesso aberto à informação científica facilita a comunicação na sociedade científica, além de promover uma igualdade de acesso entre os países desenvolvidos e em desenvolvimento e seus diversos públicos de pesquisadores e leitores.

De forma significativa, o Brasil tem desempenhado notável protagonismo no desenvolvimento do acesso aberto na região ibero-americana (Babini, 2011). Ocupando posições centrais em índices de acesso aberto e em indexadores de revistas científicas, o país se destaca em razão de ações bem-sucedidas na promoção de infraestruturas que permitem o desenvolvimento deste tipo de canal de comunicação científica livre de custos e barreiras de acesso aos seus leitores.

Importantes iniciativas como a criação da Scientific Electronic Library Online (SciELO), lançada em 1998, a ação adotada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), desde 2003, para promoção do uso do software Open Journal System (OJS) do Public Knowledge Project (PKP) e o apelo à adesão aos diretórios de acesso aberto como o Directory of Open Access Journals (DOAJ), Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (RedALyC), Diretório de Políticas Editoriais das Revistas Científicas Brasileiras (Diadorim) e o Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal (Latindex) fizeram e mostram que o acesso aberto é uma realidade na comunicação científica brasileira.

É notável que, no Brasil, os organismos ligados à ciência nacional têm agido tanto para estimular a indexação de suas revistas em bases de dados quanto para promover o acesso aberto. Tais ações foram e ainda são essenciais para impulsionar a publicação de revistas científicas em acesso aberto no país e também para o aumento da visibilidade e

impacto das pesquisas científicas brasileiras, especialmente em um cenário de ciência internacionalizada.

Observando todo esse contexto concernente às publicações, este trabalho tem por objetivo apresentar o quadro situacional das revistas científicas brasileiras de acesso aberto com base em sua presença em índices e indexadores.

## Metodologia

A metodologia adotada para o alcance de tal objetivo baseou-se uma pesquisa descritiva sob abordagem quantitativa tanto para a coleta como para a análise dos dados e foi realizada por meio do levantamento (*survey*) em índices e indexadores de grande relevância no âmbito da comunicação científica e em que as revistas científicas brasileiras de acesso aberto estão presentes. Entre as bases de dados analisadas figuram o DOAJ, o diretório de revistas do PKP, Scielo, Redalyc, Latindex (diretório e catálogo), Journal Citation Reports (JCR) e SCImago Journal & Country Ranking (SJR).

Cabe ressaltar que nem todos os indexadores são de registro exclusivo de revistas de acesso aberto, no entanto, naqueles em que são aceitos outros modelos de negócios (como periódicos baseados em assinaturas), foi usado o filtro “acesso aberto” a fim de identificar as revistas que diziam respeito ao objetivo deste trabalho (o Latindex apresentou uma nova versão em 2016 e os registros cadastrados anteriormente a esta data ainda estão em processo de atualização quanto a diversos dados, entre eles, o modelo de negócio). Pontua-se ainda que o diretório de revistas do PKP, apesar de ser relevante para este estudo em razão do fomento à criação de revistas em acesso aberto no Brasil, não pode ser tido como fonte segura quanto à natureza científica das revistas registradas (isto se deve ao fato de o software poder ser baixado livremente e usado pra criação de revistas sem o registro obrigatório no diretório. Assim, não há como inferir quantas revistas são científicas ou não). Por fim, para apresentar o quadro situacional das revistas científicas brasileiras de acesso aberto, foi avaliado o Diretório de Políticas Editoriais das Revistas Científicas Brasileiras com o propósito de observar o estado atual das declarações das políticas editoriais e identificar a proporção de revistas científicas de acesso aberto *versus* revistas baseadas em outros modelos de negócio.

A coleta de dados foi feita de maneira transversal e atualizada ocasionalmente entre o período de submissão e a versão final desta publicação, tendo sido adotado como o conjunto de dados concludente aqueles obtidos pela coleta realizada no dia 30 de setembro de 2016.

## Resultados

Como principais resultados do levantamento realizado tem-se que:



i) o Brasil ocupa o primeiro lugar no mundo em número de revistas científicas em acesso aberto registradas no DOAJ, com 875 registros, seguido pelo Reino Unido ( $n=766$ ) e Estados Unidos ( $n=673$ ), países com maior tradição científica e quantidade maior de periódicos científicos, mas que, em razão de aspectos históricos, têm alcançado êxito na conversão de revistas de outros modelos para acesso aberto e a criação de novas revistas sob a via dourada com menor celeridade do que o Brasil e outros países da América Latina (Aperín; Fischman, 2015);

ii) são aproximadamente 2.000 revistas científicas brasileiras que utilizam o software OJS/PKP ( $n=1.939$ ), traduzido no Brasil pelo Ibict como Sistema Eletrônico de Editoração de Revistas (SEER). Uma vez que o mapa do OJS apresenta um total de 10.049 revistas distribuídas em todo o mundo e 3.306 na América Latina, ressalta-se uma proporção de representatividade em relação ao total de cerca 20% do total mundial e 60% na região latino-americana;

iii) no sistema SciELO, o Brasil apresenta um conglomerado de 354 revistas científicas em acesso aberto, sendo desse total 286 títulos correntes e 68 títulos não-correntes;

iv) o RedALyC registra 224 revistas científicas;

v) no Sistema Latindex, o Brasil também se destaca quantitativamente, mas, neste caso, contabilizam-se revistas de acesso aberto e as revistas que apresentam restrições ao acesso. São 5.850 revistas científicas presentes no Diretório, tendo 2.490 em seu Catálogo. Conforme exposto na metodologia, estes números dizem respeito às revistas de acesso aberto, mas também de outros modelos de negócio. No entanto, devido aos percursos da publicação científica brasileira, bem como da América Latina, ser mais alinhada ao modelo de acesso aberto e, há algumas décadas, não ter atraído o interesse comercial para as grandes editoras científicas (fato que vêm mudando), avaliou-se como positiva a indicação da quantidade de registros apontados pela base de dados Latindex. Essa indicação torna-se ainda mais relevante diante dos números apresentados.

vi) o JCR apresenta 128 revistas científicas brasileiras, das quais 92 correspondem a publicações de acesso aberto. (TODOS OA: 977);

vii) na base SJR tem-se presente 338 registros brasileiros, sendo que destes 258 são revistas científicas de acesso aberto (TODOS OA: 3797);

viii) no Diretório Diadorim estão registradas 1.156 revistas, tendo declaradas 1.112 revistas sob o modelo de acesso aberto imediato à publicação, outras 34 permitem o acesso aberto após período de embargo e 10 declaram-se de acesso restrito. Cabe ressaltar que no Diadorim as informações referentes às políticas editoriais são registradas pelos próprios editores dessas revistas científicas;

Entre esses índices e indexadores, destaca-se a liderança do Brasil em relação à quantidade de revistas científicas cadastradas no DOAJ entre um total de 129 países

representados no diretório – o país corresponde à aproximadamente 10% de todos os registros de revistas científicas da base (875 revistas brasileiras em um total de 9.172 registros). O diretório, que passou por significativa mudança em sua forma de cadastro no ano de 2015 e atualmente exige dos editores o preenchimento de um formulário online com 58 campos, sendo 43 obrigatórios, deixou de ser apenas um índice de revistas científicas de acesso aberto e configura-se agora como um indexador. Isto porque indica dados como o tipo de avaliação feito pelo corpo editorial da revista, questões relativas à qualidade e transparência no processo editorial, o licenciamento dos conteúdos da revista e também exige a adesão a um conjunto de boas práticas editoriais para que a revista científica seja aceita no diretório. Assim, o elevado número de revistas científicas brasileiras cadastradas no DOAJ serve como uma importante fonte de informação sobre como estas revistas se apresentam não somente em relação à quantidade, mas também quanto a aspectos relativos a sua qualidade.

### **Considerações finais**

Os dados e análises realizados neste trabalho possibilitaram a traçar um quadro situacional das revistas científicas brasileiras de acesso aberto, tomando por base a sua presença em índices e indexadores.

Ao longo dos últimos quinze anos, o universo editorial científico brasileiro teve o que se chamaria de um crescimento exponencial em termos de criação de revistas científicas de acesso aberto. Vários foram os fatores que levaram à composição desta realidade onde se destacam: o incentivo financeiro concedido, regularmente, pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); uma maior pontuação na avaliação para aqueles Programas de Pós-graduação que possuem revistas, por parte da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); como também o surgimento do Sistema Eletrônico de Editoração de Revistas (SEER/OJS), que veio a ser traduzido para o português, disseminado com o oferecimento de capacitação por parte do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). Os fatores apresentados, entretanto, são apenas aqueles de relação evidente e significativas para os resultados observados. Há que se considerar também a hipótese de que, uma vez que no contexto de pesquisas científicas contemporâneas os meios eletrônicos fornecem grandes possibilidades em comunicação, compartilhamento e colaboração, os próprios pesquisadores, que se avultam cada vez mais em maiores quantidades, desejavam se comunicar mais e, assim, toda a ecologia da comunicação científica se viu afetada por este comportamento. Importante também ressaltar, como circunstância facilitadora da criação de revistas brasileiras, a questão de serem, maiormente, editadas por organismos públicos não comerciais. Trata-se de Universidades, Institutos de Pesquisa e Associações Científicas, mantidas com fundos públicos. A verdade é que a tradição do mundo editorial científico brasileiro, antes mesmo

do Movimento de Acesso Aberto, é a de fazer uso de modelos de negócio que se caracterizavam em conceder o acesso ao seu conteúdo.

O fato é que o Brasil alcançou posições de destaque na área editorial, em diretórios, bases de dados e indexadores, notadamente aqueles relacionados com o mundo do Acesso Aberto à Informação Científica. Faz-se igualmente necessário apontar, que os periódicos científicos se fazem mais presentes nas bases de dados nacionais, como o Scielo, e nas regionais, como o RedALyC e o Latindex. Há que se perseguir um caminho de desenvolvimento a fim de alcançar maior participação nos canais de disseminação dos conteúdos científicos de relevância internacional e não se limitar, assim, a expressiva participação nos contextos de alcance reduzidos.

Neste ponto, é possível refletir também sobre a necessidade da busca do equilíbrio entre o binômio: quantidade X qualidade. Grande parte das revistas brasileiras, e neste caso tanto revistas científicas de acesso aberto como de outros modelos de negócios, ainda não alcançou posições de destaque nas chamadas “grandes bases de dados indexadoras internacionais”, tais como a JCR e SJR. Aumentar a presença brasileira nesses indexadores internacionais é muito importante já que, ao estarem neles presentes, os efeitos esperados será crescimento da visibilidade da ciência nacional, aumentando o seu número de citações e assim possibilitando a criação de novas redes de colaboração e produção de conhecimento científico. Entretanto, para que se possa alcançar níveis satisfatórios de internacionalização faz-se necessário conhecer, primeiramente e mais profundamente, as revistas científicas brasileiras para que se possa encontrar formas que aumentem a sua qualidade, garantindo assim percentuais maiores de aceitação para indexação em bases de dados.

A respeito deste trabalho de pesquisa, uma vez tendo alcançado o objetivo proposto, sugere-se, a título de investigações futuras, que sejam realizados estudos que adotem em sua metodologia uma análise longitudinal no tocante à distribuição de tempo, de forma que acompanhe a presença das revistas científicas brasileiras de acesso aberto em um amplo período de tempo para que assim se possam observar os movimentos de crescimento ou retração da presença deste tipo de publicação nas bases de dados, o seu comportamento, suas características e tendências.

### Referências bibliográficas

ALPERIN, Juan P.; FISCHMAN, Gustavo E. (2015) – Revistas científicas hechas en Latinoamérica. In LÓPEZ, Eduardo A. [et al.] – *Hecho en Latinoamérica: acceso abierto, revistas académicas e innovaciones regionales*. Ciudad Autónoma de Buenos Aires : CLACSO, 2015. [Consult. 22 Set. 2016]. Disponível na Internet: <URL: <http://biblioteca.clacso.edu.ar/clacso/se/20150722110704/HechoEnLatinoamerica.pdf>>. ISBN 978-987-722-067-4.

ALPERIN, Juan P.; FISCHMAN, Gustavo; WILLINSKY, John (2008) – Open Access and Scholarly Publishing in Latin America: Ten flavours and a few reflections. *Liinc em Revista* [Em linha]. Vol. 4, N° 2. [Consult. 16 Set. 2016]. Disponível na Internet: <URL: <http://liinc.revista.ibict.br/index.php/liinc/article/view/269>>. ISSN 1808-3536.

BABINI, Dominique (2011) – Acceso abierto a la producción científica de América Latina y el Caribe: identificación de principales instituciones para estrategias de integración regional. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad* [Em linha]. Vol. 6, N° 17 [Consult. 10 Abr. 2016]. Disponível na Internet: <<http://www.redalyc.org/html/924/92422634002/>>. ISSN 1668-0030.

FIORIN, José L. (2007) – Internacionalização da produção científica: a publicação de trabalhos de Ciências Humanas e Sociais em periódicos internacionais. *Revista Brasileira de Pós-Graduação* [Em linha]. Vol. 4, N° 8. [Consult. 22 Set. 2016]. Disponível na Internet: <URL: <http://ojs.rbpg.capes.gov.br/index.php/rbpg/article/view/133>>. ISSN 2358-2332.

FRAGA, Maria N. O. (2008) – Crescimento do periódico científico e sua relação com o trabalho dos pareceristas. *Revista da Rede de Enfermagem do Nordeste* [Em linha]. Vol. 9, N° 1. [Consult. 16 Set. 2016]. Disponível na Internet: <URL: <http://www.redalyc.org/articulo.oa?id=324027961001>>. ISSN 1517-3852.

KURAMOTO, Hélio (2006) – Informação científica: proposta de um novo modelo para o Brasil. *Ciência da Informação* [Em linha]. Vol. 35, N° 2. [Consult. 20 Set. 2016]. Disponível na Internet: <URL: <http://dx.doi.org/10.1590/S0100-19652006000200010>>. ISSN 0100-1965.

MUGNAINI, Rogério; DIGIAMPIETRI, Luciano A.; MENA-CHALCO, Jesús P. (2014) – Comunicação científica no Brasil (1998-2012): infraestrutura nacional e internacionalização. In *XIII Congreso Internacional de Información*. La Habana, Cuba: 2014. [Consult. 19 Set. 2016]. Disponível na Internet: <URL: [https://www.researchgate.net/publication/262048529\\_Comunicacao\\_cientifica\\_no\\_Brasil\\_1998-2012\\_infraestrutura\\_nacional\\_e\\_internacionalizacao](https://www.researchgate.net/publication/262048529_Comunicacao_cientifica_no_Brasil_1998-2012_infraestrutura_nacional_e_internacionalizacao)>.

OLIVEIRA, Érica B. (2005) – Produção científica nacional na área de geociências: análise de critérios de editoração, difusão e indexação em bases de dados. *Ciência da Informação* [Em linha]. Vol. 34, N° 2. [Consult. 15 Set. 2016]. Disponível na Internet: <URL: <http://www.scielo.br/pdf/ci/v34n2/28553.pdf>>. ISSN 0100-1965.

---

## Integrar uma rede de repositórios no ecossistema de gestão de ciência nacional

---

José Carvalho

Universidade do Minho

[jcarvalho@sdum.uminho.pt](mailto:jcarvalho@sdum.uminho.pt)

João Mendes Moreira

FCT/FCCN

[jmm@fccn.pt](mailto:jmm@fccn.pt)

Cátia Laranjeira

FCT/FCCN

[catia.laranjeira@fccn.pt](mailto:catia.laranjeira@fccn.pt)

Fernando Ribeiro

FCT/FCCN

[fernando.ribeiro@fccn.pt](mailto:fernando.ribeiro@fccn.pt)

Paulo Graça

FCT/FCCN

[paulo.graca@fccn.pt](mailto:paulo.graca@fccn.pt)

Paulo Lopes

FCT/FCCN

[plopes@fccn.pt](mailto:plopes@fccn.pt)

### Resumo

PTCRIS (Portugal Current Research Information System) é um programa que visa a criação e desenvolvimento sustentado de um ecossistema integrado de informação nacional para apoiar a gestão da investigação, de acordo com os melhores padrões e práticas internacionais.

Para alcançar essa visão, o PTCRIS tem dois objetivos principais. O primeiro é definir um quadro regulamentar com base nas melhores normas e práticas internacionais. O segundo é promover a adoção de tal quadro nos diversos sistemas de informação, tanto a nível nacional (incluindo RCAAP) e local (ao nível institucional).

Este artigo relata o contexto, estratégia e trabalho desenvolvido até agora para tornar a rede repositórios portugueses (RCAAP), com 50 repositórios, compatíveis com o ecossistema de investigação nacional (PTCRIS), mediante a adoção do seu quadro normativo.

**Palavras-chave:** Repositórios; Rede Nacional; CRIS; Gestão

## **Integration of a repository network on a national science management ecosystem**

### **Abstract**

PTCRIS (Portugal Current Research Information System) is a program that aims at the creation and sustained development of a national integrated information ecosystem to support research management, according to the best international standards and practices.

In order to achieve such vision, PTCRIS has two main goals. The first one is to define a regulatory framework based on the best international standards and practices. The second is to foster the adoption of such framework in the various information systems, both national (including RCAAP) and local (at the institution level).

This paper reports the context, strategy and work developed thus far to make the Portuguese repositories network (RCAAP), with more than 50 repositories, compliant with the national research ecosystem (PTCRIS) by adopting its regulatory framework.

**Keywords:** Repositories; National Network; CRIS; Management

## **Introdução**

Este trabalho descreve o papel e a importância das normas e diretrizes na integração de repositórios institucionais dentro dos ecossistemas mais amplos para melhor gerir a investigação, bem como para melhor apoiar o acesso aberto, dados abertos e uma ciência aberta.

Este documento relata o contexto, estratégia e trabalho desenvolvido até agora, para tornar a rede portuguesa de repositórios (RCAAP), com 50 repositórios, compatíveis com o ecossistema de investigação nacional (PTCRIS), adotando o seu quadro normativo.

A primeira seção descreve o ecossistema nacional de gestão da investigação (PTCRIS), a rede nacional de repositórios (RCAAP) assim como o quadro normativo PTCRIS.

A segunda secção descreve os desafios, necessidades, cenários, metodologia e as ferramentas utilizadas para implementar rapidamente o quadro normativo PTCRIS na rede de repositórios.

A terceira secção concentra-se no trabalho e nos resultados alcançados no projeto piloto.

## 1-PTCRIS

O PTCRIS é um programa iniciado em maio de 2014 pela Fundação para a Computação Científica Nacional (FCCN), uma plataforma de alto desempenho para desenvolvimento e teste de aplicações e serviços avançados de comunicações da Fundação para a Ciência e Tecnologia (FCT), a agência pública nacional para a ciência, tecnologia e inovação. O objetivo principal do PTCRIS é assegurar a criação e o desenvolvimento sustentado de um ecossistema nacional integrado de informação de suporte à atividade científica desenvolvida em Portugal, de acordo com as melhores normas e práticas internacionais.

Para este fim, PTCRIS visa:

1. Definir o quadro normativo a adotar pelos diversos sistemas de informação
2. Coordenar a integração de sistemas da FCT de acordo com o quadro normativo
3. Coordenar os sistemas externos de integração (nacional e internacional) com a FCT de acordo com o quadro normativo
4. Apoiar e promover a utilização de sistemas PTCRIS dentro da comunidade.

O cumprimento destes objetivos traduz-se em inúmeros benefícios para a comunidade de Ciência e Tecnologia, nomeadamente o acesso facilitado a informação fidedigna, completa e atualizada sobre a atividade científica; a simplificação dos processos de gestão, reporte e tomada de decisão; a otimização do processo de financiamento e o aumento da visibilidade e acessibilidade dos resultados da atividade científica portuguesa, potenciando o seu uso e impacto na comunidade científica nacional e internacional.

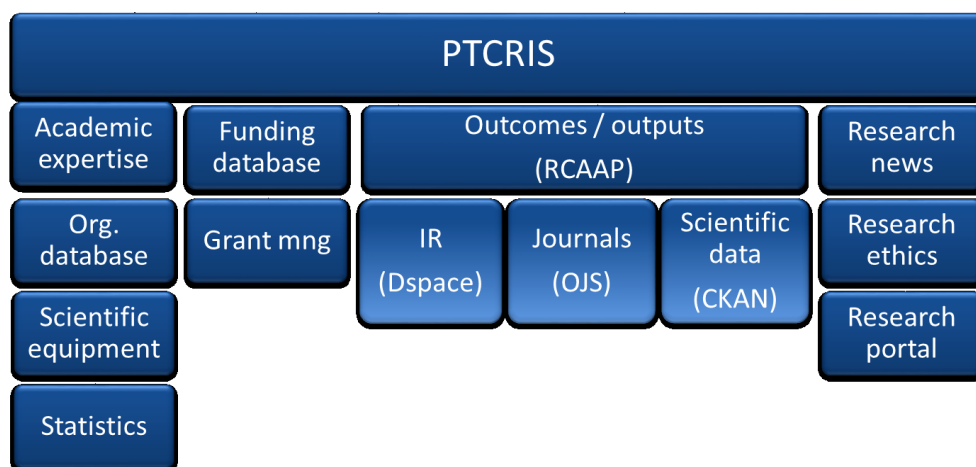


Figura 1: Componentes do ecossistema PT-CRIS

## RCAAP

Como se pode observar na Figura 1, os módulos chave do ecossistema PTCRIS são os produtos e os resultados. Em Portugal, estes componentes são coordenados dentro da atividade do projeto RCAAP (Repositório Científico de Acesso Aberto de Portugal).

O RCAAP é o principal instrumento para a implementação do Open Access (OA) em Portugal. Tem por missão promover, apoiar e facilitar a adoção do acesso aberto ao conhecimento científico em Portugal. Um dos principais serviços consiste em armazenar, disponibilizar e preservar a produção científica dos diversos repositórios institucionais a nível nacional. O RCAAP oferece ainda uma ampla gama de serviços para cumprir a sua missão OA. O serviço eletrónico mais conhecido é o serviço de alojamento de repositórios institucionais (SARI), mas há outros serviços prestados pelo RCAAP, tais como: alojamento de revistas científicas (SARC), repositórios de dados científicos (SARDC) e um repositório partilhado (Comum). Os resultados da investigação destes recursos são agregados no Portal RCAAP.

O modelo de serviço pelo qual se optou garante a realização do trabalho comum pela equipa de especialistas do projeto, libertando as instituições para a realização das atividades core. Como resultado da aplicação deste modelo, regista-se um custo de operação reduzido (em resultado das economias de escala e do trabalho especializado), um prazo de implementação de novos serviços também reduzido e permite que os gestores dos repositórios e respetivas equipas se foquem nas atividades nucleares. O projeto RCAAP procura ainda consolidar uma visão holística e integrada do movimento de acesso aberto promovendo ações de divulgação, promoção e formação e desenvolvendo parcerias nacionais e internacionais para a expansão dos serviços disponibilizados.

## Quadro Normativo PTCRIS

Para garantir que todos os componentes do PTCRIS funcionem sem problemas, um quadro regulamentar comum teve de ser definido. Com base em estudos e análises de *state-of-the-art* foram adotadas as seguintes normas e diretrizes:

1. Interoperabilidade de Dados – um dos principais resultados da PTCRIS é conseguir a integração efetiva entre elementos do ecossistema. Para este fim, tanto o perfil de aplicação (com base em CERIF-XML), como o formato de dados, a semântica e classificação de dados são essenciais;
2. Identificadores únicos – usando identificações exclusivas, persistentes e internacionais para investigadores / docentes (ORCID), organizações (ISNI / Ringgold), etc.



3. Sincronização de dados – Os dados do ecossistema residem atualmente em vários sistemas nacionais distribuídos. Para garantir a consistência dos dados, os protocolos de sincronização devem ser utilizados e neste contexto é sugerido o PTCRISync. O PTCRISync é definido com um protocolo para apoio à sincronização de publicações entre sistemas por intermédio do sistema ORCID. O seu principal benefício é o de evitar muito do trabalho repetitivo necessário para manter o perfil do autor atualizado em diferentes plataformas. O modo de funcionamento consiste num mecanismo que exporta para o ORCID todas as publicações selecionadas para sincronização pelo utilizador num sistema CRIS. Posteriormente todas as alterações efetuadas a estas publicações serão automaticamente exportadas para o ORCID, mantendo-se desta forma a informação no ORCID permanentemente atualizada. Quando novos identificadores únicos (DOIs, Handles, etc) são adicionados a uma publicação no ORCID que esteja em modo de sincronização, estes serão automaticamente adicionados à publicação respetiva no sistema CRIS;

4. Privacidade e Proteção de Dados – Para construir indicadores de confiança depende da existência de dados abrangentes, completos, consistentes e confiáveis. No entanto, a necessidade de coletar, gerir e utilizar esta informação pode estar em desacordo com os direitos de proteção de dados.

## 2– O Desafio

O desafio era integrar os serviços RCAAP, nomeadamente repositórios institucionais, no ecossistema de gestão da investigação português (PTCRIS). Isto teve de ser feito respeitando alguns princípios: código aberto, não disruptivo e com uma abordagem sustentável.

Para atingir este fim foram considerados dois cenários: S1) adaptar soluções existentes; S2) mudança da rede de repositórios institucionais (IR) para um software, tanto quanto possível, compatível com o quadro normativo PTCRIS.

Depois de algumas análises tornou-se claro que a solução mais rentável e sustentável era a utilização do DSpaceCRIS<sup>i</sup> como solução que está em conformidade com muitos requisitos do quadro normativo PTCRIS regulamentar, ou seja, integra com o sistema ORCID e inclui o formato CERIF-XML.

A rede de repositórios RCAAP tem 50 repositórios, sendo 28 deles geridos em regime de SaaS. Para testar esta abordagem, decidiu-se executar um piloto com três instituições. Desta forma proporciona-se um ambiente seguro para testar a abordagem proposta.

Atualmente, está prevista uma abordagem metodológica que irá guiar o projeto piloto e preparar as ferramentas, integrações e métodos que irão apoiar as três instituições piloto.

Um dos desafios deste projeto é manter o "trabalho em série" e a "cadeia produtiva" que nos permitem hoje gerir todos os serviços RCAAP (incluindo os 28 repositórios) com 2 homens-mês.

Além dos aspetos metodológicos, deverão ainda ser mantidos os níveis de compatibilidade com o depósito legal de teses e dissertações, a monitorização da política da FCT, assim como integrações com outros serviços e sistemas integrados na rede RCAAP.

## **Resultados**

O projeto piloto em desenvolvimento no projeto RCAAP relativamente à implementação de três instâncias do DSpaceCRIS foi iniciado através de uma candidatura aberta para a qual foram selecionadas as instituições: Universidade do Algarve, Universidade Aberta e a Egas Moniz – Cooperativa de Ensino Superior CRL.

Enquadrada na metodologia de trabalho, foi desenvolvida uma instância master que agrega todos os desenvolvimentos e alterações ao código base da plataforma e que depois é instanciado nas 3 instalações ou mais no futuro.

A metodologia adotada incluiu uma fase de preparação, com instalações e configurações técnicas mas também de contacto inicial com as instituições para definição das tarefas, expectativas e informação necessária.

Posteriormente avançou-se para a importação da informação dos repositórios existentes para o DSpaceCRIS no sentido de reaproveitar toda a informação existente.

Foram consideradas duas fases de organização da informação, uma primeira fase com foco nas publicações, pessoas, organizações e projetos, e uma segunda focada nos eventos, revistas, e outro tipo de entidades que possam vir a ser criadas.

A grande alteração face à plataforma DSpace é a possibilidade de descrever diversas entidades, implicando por si só um esforço e uma curadoria de informação mais cuidada e pormenorizada. No sentido de otimizar o trabalho, foi considerada a utilização de identificadores externos e internacionais para criar os índices das entidades em cada instância do DSpaceCRIS. Deste modo, essa informação poderá ser interpretada fora do contexto do repositório.

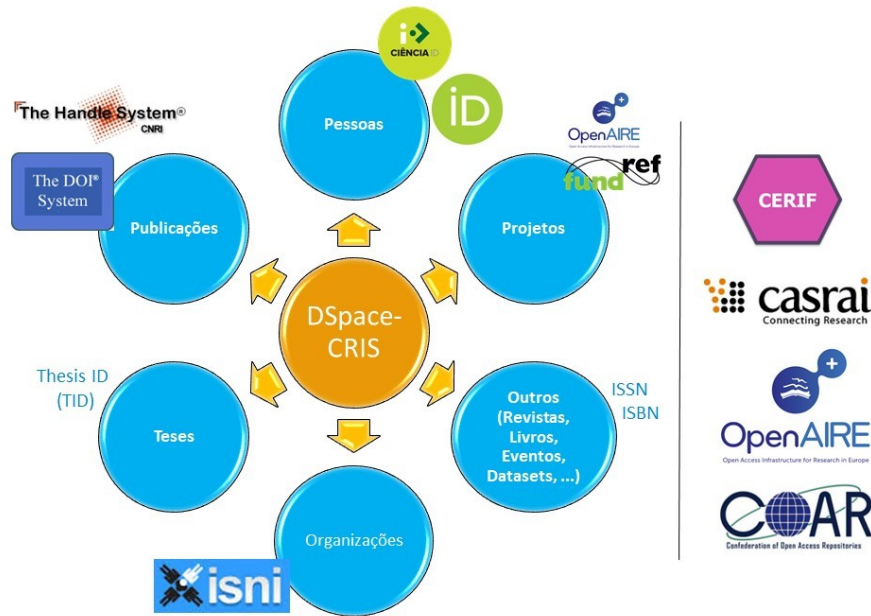


FIGURA 1 – IDENTIFICADORES DAS INSTÂNCIAS NO DSPACECRIS

Aliado aos identificadores, foram considerados vocabulários e diretrizes internacionais para a organização e partilha de informação do DSpaceCRIS.

No decorrer do processo foram identificados alguns aspetos negativos e positivos relevantes para a implementação de um repositório com DSpaceCRIS:

- **Curadoria de Informação** – É necessária a correção e uniformização de vários campos como os autores, projetos, nomes de revistas, editoras assim como lidar com situações em que um campo pode ter diversos usos conforme o tipo de trabalho descrito.
- **Bugs do Sistema** – Considerado o sistema nas versões testadas, existem ainda diversos bugs que foram resolvidos e outros por resolver, denotando ainda alguma falta de maturidade do projeto. De realçar que a equipa gestora do projeto DSpaceCRIS colaborou na resolução dos problemas.
- **Funcionalidades não implementadas** – Algumas funcionalidades não estão ainda implementadas ou não existem na sua forma final e otimizada como o interface CERIF–XML como interface máquina–máquina.
- **Versatilidade** – Considerado normalmente um aspeto positivo, o facto de a ferramenta ser altamente configurável torna também difícil a sua estabilização, principalmente com algumas lacunas de documentação.

Contudo, e tal como a primeira análise à ferramenta revelou, existem bastantes aspetos positivos a considerar:

- **Qualidade dos Metadados** – A alteração da informação descritiva dos recursos para o formato CERIF–XML em vez de Dublin Core permite e obriga a um maior cuidado na curadoria dos metadados pois neste caso a informação está ligada entre as várias instâncias.

Implica também mais tempo para que toda essa informação seja registada, mas a termo compensa pois o sistema implementa mecanismos internos de índices que ajudam no preenchimento e ligação da informação.

– **Informação pormenorizada** – O alargamento do esquema de metadados leva a que cada instância seja descrita com mais pormenor, permitindo análises mais finas e relatórios sobre a informação mais ricos, quer para o gestor, quer para os diferentes níveis de gestão na instituição (departamentos, centros de investigação ou instituição como um todo).

– **Perfil de Investigador** – O perfil do investigador é um conceito muito importante neste contexto pois existe uma relação direta entre as publicações do autor e o seu perfil, permitindo-lhe ter um papel mais ativo na construção do seu próprio perfil.

– **Mais métricas** – Como cada entidade possui indicadores e métricas próprias, permite uma monitorização mais alargada, quer das publicações, mas também da página do investigador, da página do projeto ou da página do departamento.

– **Metodologia de importação de DSpace** –> DSpaceCRIS – Este projeto piloto permite ainda definir uma metodologia, otimizada e testada para a importação da informação existente num repositório institucional com DSpace para uma instância do DSpaceCRIS. Assim que totalmente terminada, permitirá a sua aplicação a 28 instâncias do projeto SARI assim como será partilhada pela comunidade para que outras instituições possam usar essa mesma metodologia para a importação e gestão da informação na nova instância DSpaceCRIS.

– **Integrar o quadro normativo PTCRIS** – Com esta iniciativa, é mais um passo dado na consolidação da integração dos repositórios portugueses no quadro normativo PTCRIS. Sendo os repositórios sistemas base onde é organizada, gerida e validada a informação científica das instituições, será sempre o ponto inicial para enriquecer os dados que serão depois usados noutros contextos como os portais agregadores, sistemas de currículos ou o próprio ORCID.

Face ao decorrer das atividades do piloto DSpaceCRIS, será necessário melhorar alguns aspetos na plataforma no sentido de estar totalmente compatível com o quadro normativo. Contudo, para tal é necessário o contributo da comunidade para que seja melhorada a documentação e sejam afinadas configurações e comportamentos que se enquadrem nos requisitos da comunidade.

## Conclusão

Os repositórios institucionais são parte importante do ecossistema de investigação. Para integrar sem problemas os repositórios institucionais num ecossistema mais amplo, estes devem cumprir com um quadro normativo. Por outro lado, nos últimos anos, os sistemas de gestão de publicações como os repositórios tornaram-se insuficientes para atender às necessidades dos gestores de informação. Além de publicações, as informações sobre autores, afiliações, financiamento e infraestruturas de investigação são cada vez mais importantes.

Nesse sentido, o desenvolvimento de sistemas integrados como o DSpaceCRIS, alinhados com orientações, diretrizes, e identificadores internacionais são fundamentais para definir uma base de informação validada, perto da fonte autoritativa que permitirá alimentar o conceito de CRIS nacional como o PTCRIS.

Com este trabalho, iremos contribuir para a integração com os sistemas externos da web e, conseqüentemente, proporcionar um melhor apoio do conceito de Ciência Aberta.

---

<sup>i</sup> <https://wiki.duraspace.org/display/DSPACECRIS/DSpace-CRIS+Home>

---

## Perfil dos usuários do repositório institucional da Universidade Federal da Bahia

---

Flávia Rosa

Universidade Federal da Bahia

[flaviagr@ufba.br](mailto:flaviagr@ufba.br)

Susane Barros

Universidade Federal da Bahia

[cooreditorial@ufba.br](mailto:cooreditorial@ufba.br)

Rodrigo França Meirelles

Universidade Federal da Bahia

[rodrigomei@ufba.br](mailto:rodrigomei@ufba.br)

### Resumo

O objetivo principal deste trabalho é avaliar os resultados da implantação do Repositório Institucional da UFBA para a Universidade e sua comunidade, sendo que os dados aqui apresentados correspondem a etapa de avaliação que busca conhecer o perfil, as expectativas e interesses dos usuários. Esta pesquisa utilizou como método de investigação o survey e como técnica de coleta de dados a aplicação de questionário. Contribuições para a melhoria desse serviço foram apresentadas, algumas possíveis de serem implementadas outras que independem da atuação da equipe para efetiva solução. Destacou-se nas respostas, a necessidade de maior divulgação do repositório na instituição e padronização dos termos de busca já que um mesmo termo é encontrado com grafias variadas.

**Palavras-chave:** Usuário. Serviço de informação. Repositório Institucional. Universidade Federal da Bahia. Comunicação científica.

### **Abstract**

The main objective of this study is to evaluate the results of UFBA's Institutional Repository implementation for the use of University itself as an institution, as well as for the benefit of its community. The data presented in here are those regarded to the evaluation phase, which seeks to establish the users profile, expectations and interests. An inquiry was carried out using the survey method and a questionnaire as data collection technique. A number of contributions to the service improvement were presented, some of which are feasible to be implemented, other, however, go beyond implementation team powers for their effective resolution. The inquiry responses highlighted the need for greater disclosure on the repository within the institution and the search terms standardization, once the same term is found with different spellings.

**Keywords:** User. Service information. Institutional Repository. Universidade Federal da Bahia. Scientific communication.

### **Introdução**

Na sociedade contemporânea as redes eletrônicas de comunicação, em particular a internet, alteraram o fluxo informacional da comunicação e possibilitaram uma ampla circulação dos resultados das pesquisas, aproximaram pesquisadores e contribuíram para a recepção mais ágil da informação também para o usuário. Em 2002, cientistas do Open Society Institute (OSI) reunidos em Budapeste e descontentes com o até então fluxo da comunicação científica formal, redigiram e publicaram um documento chamado Budapest Open Access Initiative (BOAI, 2002), que indicavam alterações decisivas nesse fluxo possibilitando que o acesso a conteúdo científico digital, *on-line*, fosse livre de barreiras de custo e de permissão. Nesse contexto e nesse mesmo ano (2002), foram criados os primeiros Repositórios Institucionais (RIs) disponibilizando arquivos especialmente de conteúdo acadêmico e científico. Atualmente, segundo dados do Registry of Open Access Repositories (ROAR, 2016),<sup>1</sup> há 2938 RIs no mundo, registrados no ROAR, sendo 155 deles registrados como de instituições do Brasil.

Este trabalho apresenta parte dos resultados de um projeto de pesquisa sobre a adoção de uma política de acesso aberto pela Universidade Federal da Bahia (UFBA). Seu objetivo principal é avaliar os resultados da implantação do Repositório para a Universidade e sua comunidade, assim como a percepção de suas instâncias superiores para torná-lo uma ferramenta de gestão da produção científica observando sua importância para a disseminação da pesquisa realizada na instituição como atendimento a uma política institucionalizada de acesso aberto. Os resultados aqui expostos correspondem a etapa do

projeto que busca conhecer o perfil, as expectativas e interesses dos usuários do Repositório Institucional da UFBA (RIUFBA). Em 2010, pesquisa semelhante foi realizada no Repositório da Universidade do Minho,<sup>ii</sup> da qual adaptou-se o instrumento de coleta utilizado para pesquisa desenvolvida no RIUFBA.

A motivação principal para a realização dessa etapa foi o entendimento do papel central do usuário para os serviços de informação. É o *feedback* do usuário que orientará o aprimoramento dos serviços oferecidos e garantirá sua visibilidade. Para cumprir com sua finalidade um serviço de informação deve estar integrado à instituição e ser percebido como importante ao seu funcionamento. Segundo Cunha, Amaral e Dantas (2015) compreender a relação entre pessoas e grupos de pessoas bem como as formas como interatuam com a informação e na busca pela informação são importantes contribuições dos estudos de usuários. Passados cinco anos da implantação do RIUFBA é necessário avaliar se há essa integração, se os usuários percebem a importância dos serviços oferecidos, como interagem com esse mecanismo na busca por informação e o que pode ser melhorado para potencializar sua utilização.

O repositório representa a concretização de uma política institucionalizada de acesso aberto à produção científica, cultural e artística da UFBA, pois até recentemente não existia uma iniciativa de um mecanismo eletrônico com uma política para a disseminação da produção científica. Tradicionalmente, as políticas e práticas estabelecidas foram tratadas de forma fragmentada na Instituição, o que certamente afetou a preservação da produção intelectual. Percebe-se com isso que não apenas a disseminação e visibilidade da produção irão justificar a manutenção e investimentos em tal mecanismo, mas também seu uso e capacidade de preservação do conhecimento produzido na Universidade.

A qualidade do serviço, no entanto, está vinculada ao investimento que é feito pela Instituição para que o RIUFBA cumpra com sua função. O RI possui um Grupo Gestor com representantes da Pró-Reitoria de Pesquisa, Criação e Inovação, Pró-Reitoria de Ensino de Pós-Graduação, Superintendência de Tecnologia e Informação, Sistema de Universitário de Bibliotecas (Sibi), Instituto de Ciência da Informação (ICI) e Editora da UFBA (Edufba). Atualmente, a estrutura física do Repositório está sediada no Sibi. Dois técnicos de nível médio e dois bibliotecários fazem a gestão técnica e suporte e um grupo de bibliotecários são responsáveis pela validação dos depósitos. A Edufba, comunidade piloto do RIUFBA, oferece apoio com estagiários na divulgação do conteúdo.

### **Repositório Institucional como serviço de informação**

A década de 1960 trouxe muitas inovações no que se refere aos serviços de informação especializada. De acordo com Figueiredo (1996) houve nesse período um crescimento no número de publicações de revisão, de publicações de síntese e avaliação, criação de bases de dados, diretórios e disseminação seletiva da informação. E foi também



quando surgiram os estudos de usuários. Entre 1970 e 1980 as tecnologias na área de informação começam a ser utilizadas. Surgem nesse período as primeiras bases de dados de texto integral, começam a ser implantados sistemas de bases de dados, assim como novos serviços extraídos das bases começam a ser criados. A disseminação seletiva da informação passa a ser automatizada e a avaliação de serviços de referência torna-se central. Surge também o CD-ROM, mas só passa a ser amplamente utilizado na década de 1990. A partir desse momento outros serviços mediados pelas tecnologias de informação e comunicação são criados e com a internet surgem também ambientes científicos digitais como bibliotecas, repositórios e periódicos científicos.

Um repositório institucional é, ao mesmo tempo, um serviço e uma fonte que fornece informações especializadas de diversas áreas do conhecimento. Serviço porque como todo serviço caracteriza-se como um produto intangível que visa à satisfação de necessidades e possui propriedades que os distinguem de outras atividades: fluxo, variedade e uso intensivo de recursos humanos. Um serviço só acontece no fluxo do trabalho contínuo, no momento que o usuário solicita e por isso, não pode ser estocado e é difícil de ser mensurado. Um serviço é diverso porque exige mão de obra especializada em relação às técnicas produtivas. Um serviço é interativo porque no momento da prestação há um nível de relacionamento entre usuário e prestador de serviço. (MEIRELLES, 2006)

Uma fonte de informação tem como função principal responder a uma demanda de informação, produto ou serviço de informação, uma pessoa, grupo de pessoas, uma organização, etc. (CAMPELLO; CENDÓN; KREMER, 2000). Para cumprir com sua função o usuário tem de ter atribuído valor a ela para então utilizá-la, respondendo assim a necessidade de conhecer. Essa necessidade de conhecer é inerente à espécie humana que busca informação para mudar o estado atual de conhecimento sobre algo. No contexto acadêmico-científico é o aporte para os avanços científicos e tecnológicos. Ou seja, é por meio de serviços de informação, ao acessar as fontes, que o usuário tem sua necessidade de informação respondida.

Na visão de Gomes (1975, p. 64) “adequada e rápida transferência de conhecimento entre as comunidades científicas” é a principal função de um serviço de informação. O serviço de informação tem como finalidade fornecer a informação solicitada com a maior agilidade e precisão pelo menor custo possível. De acordo com Whitaker (apud GROGAN, 2001, p. 8) “a finalidade de um serviço de referência e informação é permitir que as informações fluam eficientemente entre as fontes de informação e quem precisa de informações.” Sendo assim, todo serviço de informação deve ser centrado no usuário porque tem nele sua razão de existir.

Para atender às necessidades dos usuários um serviço de informação num ambiente científico digital, especialmente os repositórios, Camargo e Vidotti (2009) consideram que alguns elementos essenciais – ferramenta de busca, metadados, política, interoperabilidade,

preservação, acessibilidade e usabilidade – têm de ser abordados. A política é uma das características mais importantes porque determina, dentre outras questões, as que envolvem avaliação e ajustes. Tendo a UFBA adotado uma política de acesso aberto por intermédio do repositório faz-se necessário avaliar os resultados de sua implementação. Por entender que implementar envolve, além da implantação, um período de adaptação, acolhimento e incorporação do que foi implantado, é que se mostrou necessário o acompanhamento deste processo do RIUFBA. Nessa perspectiva foi que se buscou conhecer o perfil dos usuários, suas expectativas e interesses e que uso estão fazendo do RIUFBA.

Nos resultados obtidos por Rosa (2011) com a pesquisa de usuários do repositório institucional da Universidade do Minho (RepositoriUM) verificou-se que a maioria dos usuários era vinculado a Universidade do Minho (UMinho), principalmente como estudante, sendo reduzido, no entanto, o número de professores e/ou pesquisadores. Entre os usuários finais, havia um número significativo daqueles que não tinham qualquer tipo de vínculo com a Instituição. Ressalta-se que a maioria tinha como área de interesse as Ciências Sociais e a representação dos estudantes de pós-graduação, principalmente ao nível do mestrado, como usuários do RepositóriUM, explica o fato do principal tipo de trabalho para o qual os usuários recorrem ao RepositóriUM serem “trabalhos acadêmicos ao nível da pós-graduação (mestrado)”.

Conhecida pelo pioneirismo, a UMinho foi uma das primeiras instituições europeias a implementar um repositório institucional. Durante o processo de implementação, em 2005, a Instituição lançou um programa de incentivo financeiro para que as unidades povoassem o repositório, pois este processo estava sendo mais lento que o esperado. A política produziu efeitos importantes tanto no aumento do número de comunidades quanto no número de documentos, sobretudo porque divulgou o repositório por intermédio de uma política de recompensa, mas também porque apontou para a necessidade de se exigir a obrigatoriedade no autoarquivamento. Cinco anos depois, a política de obrigatoriedade foi implantada, o que demonstra o compromisso e o entendimento das instâncias superiores de que um repositório amplia o impacto da pesquisa e a visibilidade da instituição, além de ser uma ferramenta de gestão da produção científica.

Observando outros aspectos, Marcelino e Siebra (2015), em trabalho sobre a interação de usuários com repositórios institucionais de instituições federais de ensino superior brasileiras, levantaram os tipos de dispositivos mais utilizados pelos usuários dos repositórios para acessar a internet, o tempo “gasto” com acesso a internet e serviços web mais utilizados. Aplicaram testes de usabilidade e acessibilidade e buscaram identificar dificuldades apresentadas pelos usuários no acesso aos sites e principalmente na execução de tarefas. Na avaliação de acessibilidade as autoras perceberam que todos os repositórios da amostra – um repositório ativo por capital brasileira – apresentam erros que podem dificultar o uso dos repositórios por pessoas com deficiência. Os resultados da pesquisa

apontaram também um grande percentual (70%) de usuários que não conhecia os repositórios, o que indica que são pouco divulgados em suas instituições. No que se refere à usabilidade as autoras concluem pela dificuldade que os usuários enfrentam para acessar as informações que desejam, sobretudo usuários com deficiência, e recomendam que os modelos de interação dos repositórios sejam repensados.

Em outra perspectiva, Bell e Sarr (2010) relatam a experiência da University of Rochester, que, com base em estudos de usuários, desenvolveu um novo *software* para seu repositório que tem a função de disseminar e preservar informação, mas também está integrado às atividades de pesquisa desde a fase de criação para envolver a comunidade criando uma solução atraente para o autoarquivamento. Para isso se colocaram o desafio de projetar um repositório que fosse suficientemente interessante para a comunidade induzindo mudança de hábitos e incorporando-o às suas tarefas rotineiras. O repositório da University of Rochester foi implantado em 2002 e inicialmente o *software* utilizado foi o DSpace. Foi povoado de forma bem lenta e a comunidade rapidamente se mostrou desinteressada. Após os estudos de usuários identificaram uma série de questões que os ajudariam a desenvolver um novo *software*, o IR+, que pudesse atender às necessidades da comunidade. Eles não queriam apenas um “receptáculo” para o trabalho final, queriam uma ferramenta com um espaço privado onde pudessem coletar material, compartilhar textos em várias fases, desenvolver trabalhos em colaboração com outros pesquisadores em diversas plataformas entre outras coisas.

## Metodologia

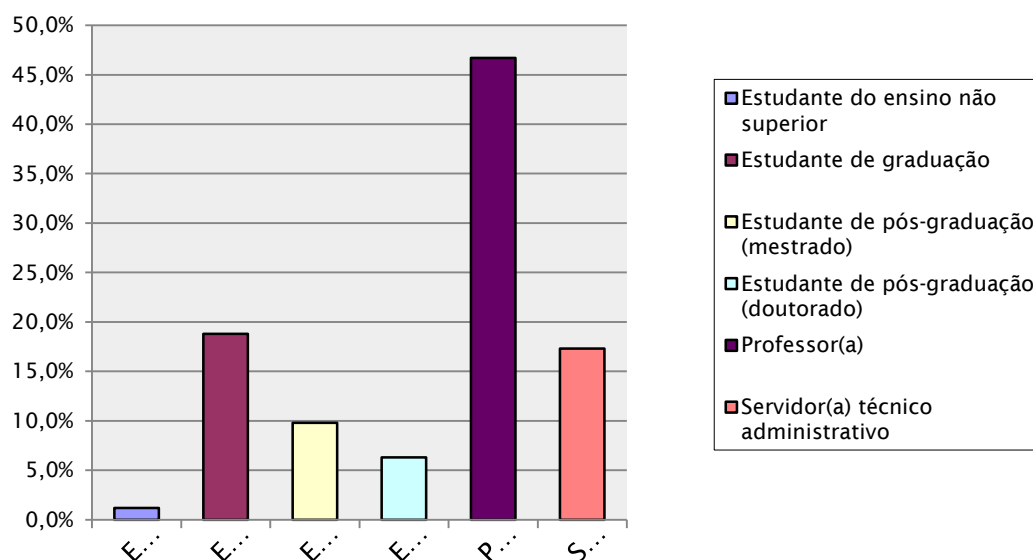
Esta pesquisa utilizou como método de investigação o survey e como técnica de coleta de dados a aplicação de questionário e foi realizada no período de 22 de abril de 2015 a 14 de agosto de 2015 e obteve 255 respondentes. O instrumento de coleta foi construído no software Survey Monkey e disponibilizado na página do RIUFBA para livre acesso. O instrumento foi formulado com 16 questões, sendo quatro fechadas de resposta única, 11 fechadas com a opção outros e uma com a possibilidade de marcar mais de uma alternativa. Foram disparados e-mails de divulgação em listas e realizadas postagens no facebook do RI. Além de informações sociodemográficas o questionário contemplou perguntas sobre uso, frequência de uso, percepção sobre a importância dos serviços, interesses e expectativas com relação a oferta de serviços do repositório.

## Perfil dos Usuários do RIUFBA

Quanto às informações sociodemográficas a faixa etária varia entre 20 e 50 anos. A prevalência é de usuários brasileiros, não houve registro de respondentes de outros países de língua oficial portuguesa. Mas, se observarmos os dados estatísticos dos documentos depositados no RIUFBA existem acessos por usuários de todos os continentes. O questionário, no entanto, foi respondido apenas por brasileiros.

Conforme apresentado no Gráfico 1 quase metade dos respondentes (46,7%) são docentes diferentemente dos resultados apresentados na pesquisa do repositório da UMinho, onde a maioria dos usuários são estudantes do mestrado. Considerando o universo de alunos matriculados regularmente – a Instituição tem cerca de 30 mil alunos de graduação e 5.500 de pós-graduação para cerca de 2500 professores –, a quantidade de respondentes confirma que o repositório não tem alcançado os discentes. Dentre os respondentes 9,8% são estudantes de mestrado, 6,3% são estudantes de doutorado e 18,8% são estudantes de graduação.

Gráfico 1 - Ocupação

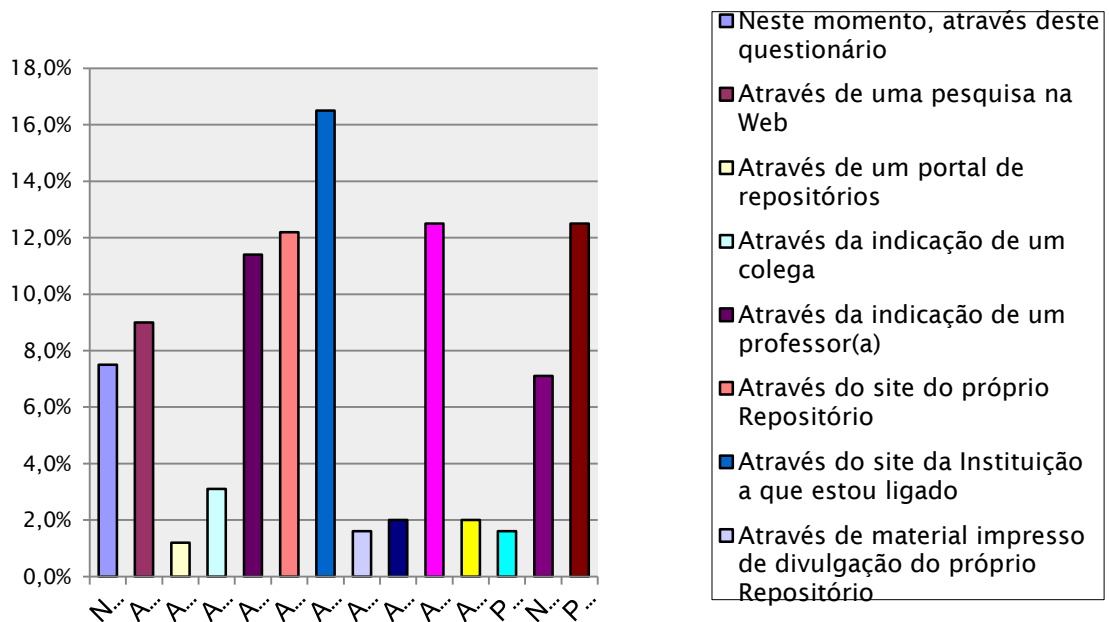


Fonte: elaborado pelo autor

Quando perguntados sobre a vinculação institucional cerca de 90% dos usuários são vinculadas a uma instituição de ensino superior, destes 75% são vinculados à própria UFBA que apresenta os usuários distribuídos de maneira bastante pulverizada entre suas unidades, sendo a Faculdade de Odontologia a única unidade que não apresentou nenhuma participação. Dentre as categorias de usuários (usuário, depositante, validador e membro de grupo gestor) 52,4% são apenas usuários e 38,4% usuários/depositantes. Esses dados revelam um baixo número de usuários depositantes em relação ao número de potenciais depositantes, docentes e alunos de mestrado e doutorado, que somam 62,8% dos respondentes que são, em sua maioria, usuários que possuem vinculação com a UFBA. Os resultados reforçam o que Rosa, Barros e Meirelles (2015) afirmam sobre a necessidade dos programas estimularem seu corpo docente e discente a autoarquivarem suas produções no RI; da UFBA de criar políticas de estímulo e/ou obrigatoriedade de autoarquivamento; assim como de ampliar as estratégias de divulgação e conscientização da comunidade da UFBA sobre a importância do RI.

O Gráfico 2 demonstra que a maioria dos usuários tomou conhecimento do RIUFBA por meio do *site* da Instituição (16,5%) à qual estão vinculados, mas o Alerta<sup>iii</sup> (12,5%) e o *site* (12,2%) do repositório também foram citados como o primeiro contato dos usuários com o RIUFBA. Questionou-se aos respondentes, há quanto tempo são usuários do RI e 29,7% responderam que desde 2010, ou seja, no ano de criação do RI, enquanto 22,4% informaram que a partir de 2015. Sobre as áreas de interesse dos usuários são: Educação (32%) e Interdisciplinar (24,3%). A área das Engenharia foi a que apresentou o menor número de interessados.

Gráfico 2 - Como tomou conhecimento da existência do Repositório Institucional da UFBA?

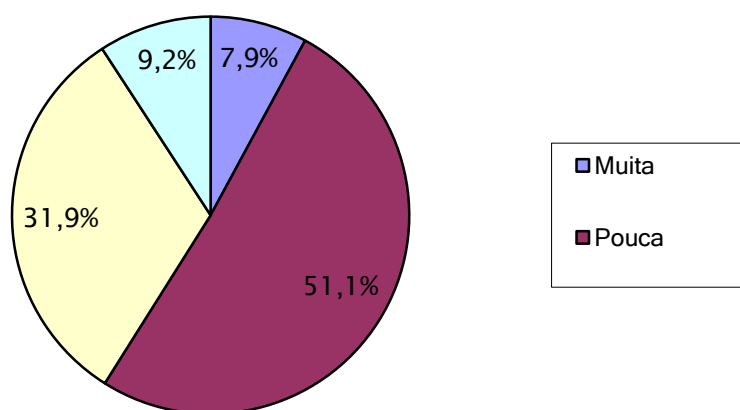


Fonte: elaborado pelo autor

No que se refere à usabilidade, outro elemento essencial apontado por Camargo e Vidotti (2009) – adequação da interface do sistema à necessidade do usuário – o Gráfico 3 apresenta que apenas 7,9% dos usuários têm muita dificuldade em acessar o RI, enquanto 51,1% afirmam que têm pouca dificuldade e 31,9% nenhuma dificuldade. Quanto às opções de pesquisa (busca por autor, título, assunto, tipo de documento, data) disponíveis no RIUFBA são consideradas adequadas por 73,4% dos respondentes e inadequadas por 2,2%. Esses dados convergem quando perguntados sobre a localização de documentos relevantes “na maior parte das vezes” é indicada por mais de 50% dos respondentes e somente 2,6% informaram que nunca localizam documentos relevantes, demonstrando que usuários têm pouca dificuldade em localizar os documentos e na relevância nos resultados da busca. No

entanto, os usuários apontam que existem inconsistências nas entradas para autores, quantidades os filtros insuficientes também considerados elementos essenciais para atender as necessidades dos usuários, ferramentas de busca e metadados. Essas questões devem ser aprimoradas para um melhor desempenho do RIUFBA.

**Gráfico 3** - Sente dificuldades na pesquisa de informação no Repositório?



Fonte: elaborado pelo autor

Questionados sobre a finalidade para a qual acessam o repositório a indicação de uso para elaboração de “Trabalhos técnicos (Nível médio)” confirma a presença de usuários com idade inferior a 20 anos, conforme dado apresentado anteriormente. Do mesmo modo que a maior prevalência para “Escrita de artigos e *papers* de pesquisa” vincula-se diretamente a ocupação de 46,7% dos respondentes que são professores. No que diz respeito à frequência de acesso, 59,8% responderam que acessam ocasionalmente em função de necessidades concretas, o que confirma a prevalência de uso para “Escrita de artigos e *papers* de pesquisa” para justificar as situações em que acessam o repositório.

## Conclusão

A ênfase de um serviço de informação deve ser o usuário e a identificação da sua necessidade de informação. Vários tipos de documentos podem dar suporte às respostas num serviço de informação. Esses documentos vão subsidiar a criação de novo conhecimento e de novas fontes.

Para o bom funcionamento do serviço de informação e satisfação do usuário é imprescindível o aperfeiçoamento de processos informacionais e o desenvolvimento de competências pelos bibliotecários para lidar tanto com técnicas e suportes tradicionais quanto com objetos digitais. Para isso é necessário também uma boa gestão, que se faz com recursos humanos, financeiros e materiais.

Os resultados obtidos com a pesquisa que definiu o perfil do usuário do RI da UFBA, serão norteadores para a melhoria do serviço prestado e desenvolvimento de ações que contribuam para ampliar sua divulgação. A revisão dos meta dados para padronizá-los é uma ação já em desenvolvimento. Espera-se contribuir cada vez mais para ampliar a visibilidade, o número de acessos, o autoarquivamento e que de fato o RI da UFBA cumpra com o seu papel de disponibilizar a produção científica da Instituição.

### Referências bibliográficas

BARROS, Susane S.; ROSA, Flavia G. M. G.; MEIRELLES, Rodrigo F. (2015). Repositório Institucional da Universidade Federal da Bahia como mecanismo de visibilidade para os programas de pós-graduação. *PontodeAcesso* [Em linha]. Vol 9, Nº 3. [Consult. 25 set. 2016]. Disponível em:

<<http://www.portalseer.ufba.br/index.php/revistaici/article/view/15086>> ISSN 1981-6766

BELL, Suzanne; SARR, Natha. (2010). Case Study: Re-Engineering an Institutional Repository to Engage Users. *New Review of Academic Librarianship*, [Em linha]. Vol 16, Nº Supl 1. [Consult. 25 set. 2016]. Disponível em: <<http://www.tandfonline.com/doi/full/10.1080/13614533.2010.509517?scroll=top&needAccess=true>>

CAMARGO, Liriane S. de A. de; VIDOTTI, Silvana A. B. G. (2009). Arquitetura da informação para repositórios científicos digitais. In: SAYÃO, Luis et al. *Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação*. Salvador: EDUFBA.

CUNHA, Murilo; AMARAL, Sueli Angelica de; DANTAS, Edmundo Brandão. (2015). *Manual de estudo de usuários da informação*. São Paulo: Atlas.

FIGUEIREDO, Nice Menezes de. (1996). *Textos avançados em referência & informação*. São Paulo: Polis: Associação Paulista de Bibliotecários.

GOMES, Hagar E. (1975) – *O pensamento de Paul Otlet e os princípios do UNISIST*. Niterói: Universidade Federal Fluminense. Tese de Livre-Docência.

GROGAN, Denis. (2001). *A prática do serviço de referência*. Tradução de Antonio Agenor Briquet de Lemos. Brasília, DF: Briquet de Lemos/Livros.

MARCELINO, Críssia de S.; SIEBRA, Sandra de A. (2015). Análise da interação de usuários com repositórios institucionais de instituições federais de ensino superior brasileiras. *Revista Gestão.Org*, [Em linha]. Vol. 13, Nº esp. [Consult. 25 set. 2016]. Disponível em:

<<http://www.revista.ufpe.br/gestaoorg/index.php/gestao/article/view/788>>. ISSN 1679-1827.

MEIRELLES, D. S. e. (2006). O conceito de serviço. *Revista de Economia Política*, [Em linha]. Vol. 26, Nº 1, [Consult. 27 ago. 2015]. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0101-31572006000100007&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-31572006000100007&lng=en&nrm=iso)>.

ROSA, Flavia G. M. G. (2011) – *A disseminação da produção científica da Universidade Federal da Bahia através da implantação do seu repositório institucional*. Uma política de acesso aberto. Salvador: Instituto de Humanidades, Artes e Ciências Professor Milton Santos, Universidade Federal da Bahia. 242 f. Tese de doutoramento. Disponível em: <http://www.repositorio.ufba.br/ri/handle/ri/3031> Acesso em: 14 set. 2015.

---

<sup>i</sup> <<http://roar.eprints.org/>>.

<sup>ii</sup> Rosa (2010)

<sup>iii</sup> É uma publicação do Núcleo de Disseminação do Conhecimento (NDC) e destina-se a divulgar a produção acadêmica da UFBA registrada no seu Repositório Institucional. O Núcleo foi criado e é mantido pelo Grupo Gestor do Repositório Institucional da Universidade Federal da Bahia (RI/UFBA). Para mais informações e para acessar todas as edições do Alerta, visite: [www.ndc.ufba.br](http://www.ndc.ufba.br).



---

## Repositórios temáticos e memória: a constituição da educação em saúde no Brasil por meio da trajetória de Virgínia Schall

---

Denise Nacif Pimenta

Instituto Centro de Pesquisas René Rachou–Fiocruz Minas, Fundação Oswaldo Cruz

[denise.pimenta@cpqrr.fiocruz.br](mailto:denise.pimenta@cpqrr.fiocruz.br)

Leandro da Conceição Borges

Instituto Brasileiro de Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz (Fiocruz)

[leandrocb@bol.com.br](mailto:leandrocb@bol.com.br)

Kizi Mendonça de Araújo

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz /Ficocruz

[crisrina.guimaraes@icict.fiocruz.br](mailto:crisrina.guimaraes@icict.fiocruz.br)

Maria Cristina Soares Guimarães

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz /Ficocruz

[crisrina.guimaraes@icict.fiocruz.br](mailto:crisrina.guimaraes@icict.fiocruz.br)

Cícera Henrique da Silva

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz /Fiocruz

[cicera.silva@globo.com](mailto:cicera.silva@globo.com)

## **Resumo**

Relata-se a trajetória da pesquisadora brasileira Virgínia Torres Schall (1954–2015) e sua contribuição para o campo da Educação em Saúde. Um protótipo de Repositório Temático (RT) na área de Educação em Saúde foi estruturado em parceria entre o Centro de Pesquisas René Rachou e o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, ambos da Fundação Oswaldo Cruz. Realizou-se uma coleta da produção da pesquisadora na plataforma Currículo Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil. Tal iniciativa almeja sistematizar e divulgar em acesso aberto sua diversificada produção acadêmica e de produtos informativos/educacionais, pois como pioneira no campo, sua produção constitui enorme riqueza e memória do campo de Educação em Saúde no Brasil.

**Palavras-chave:** Repositório temático, Memória, Acesso Livre, Educação em Saúde, Virgínia Torres Schall

## **Subject repositories and memory: the establishment of health education in Brazil through the trajectory of Virgínia Schall**

### **Abstract**

We report the trajectory of a Brazilian researcher Virginia Torres Schall (1954–2015) and her contribution to the field of Health Education. A prototype of a Subject Repository in the field of Health Education was organized in partnership with the René Rachou Research Center and the Institute of Scientific and Technological Communication and Information in Health, both from the Oswaldo Cruz Foundation. A collection of the production of the researcher was organized from the platform Curriculum Lattes from the Brazilian National Council for Scientific and Technological Development. This initiative aims to systematize and disseminate in open access the diverse production of academic and informative/educational products. As a pioneer in the field, her production is of great wealth for the memory of the field of Health Education in Brazil.

**Keywords:** Subject repositories, Memory, Open access, Health education, Virgínia Torres Schall.

## **Introdução**

A preservação e disseminação da memória e informação são consideradas atualmente como fatores importantes dentro do processo informacional (Borges, 2015). Neste cenário, as Tecnologias da Informação e Comunicação (TIC) e o Movimento de Acesso Livre (Open Access) surgem nas últimas décadas como ferramentas potencializadoras desse processo. Assim, os Repositórios Digitais (temáticos ou institucionais) são utilizados como uma das

principais ferramentas para a implementação do Acesso Livre no mundo e como mecanismo de preservação da memória e gestão da informação.

A trajetória de uma pesquisadora brasileira e sua contribuição para a área do conhecimento Educação em Saúde é objeto de estudo do projeto em desenvolvimento no Centro de Pesquisas René Rachou (CPqRR), em parceria com o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT), ambas unidades da Fundação Oswaldo Cruz (Fiocruz).

Virginia Schall teve trajetória fundamental na constituição de inúmeras linhas de pesquisa no Brasil, tal como a Educação em Saúde, Saúde Coletiva, Ensino de Ciências e Divulgação Científica. Psicóloga, mestre em neurofisiologia e comportamento; doutora em educação, pesquisadora titular Fiocruz, a cientista tem mais de uma centena de artigos completos publicados em periódicos nacionais e internacionais e anais de congressos. Suas linhas de pesquisa foram diversas e de cunho essencialmente interdisciplinar, tais como Ciências Humanas e Sociais em Saúde, Informação, Educação e Comunicação (IEC) em Saúde, Divulgação científica sobre saúde para escolares, professores e profissionais de saúde e Popularização da Ciência e da Tecnologia. Virginia Schall teve participação fundamental na construção e no desenvolvimento da área de divulgação científica na Fiocruz. Foi peça chave na criação do Museu da Vida e de vários outros projetos de grande importância para a instituição.

Foi docente do Programa de Pós-Graduação em Ciências da Saúde e do Programa de Saúde Coletiva na Fiocruz Minas. Publicou onze livros infanto-juvenis, alguns dos quais compõem as coleções: Ciranda da Saúde (Antares, RJ, 1986), Ciranda do Meio Ambiente (Memórias Futuras, 1989) e Ciranda da Vida (Memórias Futuras, 1990).

Recebeu também prêmios pela atividade de divulgação científica, um nacional, em 1990, Prêmio José Reis de Divulgação Científica do CNPq, e um estadual, em 2002, o Prêmio Francisco de Assis Magalhães Gomes, da Secretaria de Ciência e Tecnologia do Estado de Minas Gerais. Como pesquisadora foi consultora do Ministério da Saúde, Ministério da Educação, Coordenação de Aperfeiçoamento de Ensino Superior (Capes), CNPq, Fundação de Apoio ao Estado de Minas Gerais (Fapemig), tem proferido palestras e escreveu para o Canal Futura de Televisão (Programa Viva Legal) e colaborou em outros programas da emissora. Foi também uma das pesquisadoras percutoras da área de ensino em ciências no país.

Virgínia Schall tem grande e significativa produção intelectual, não só em textos acadêmicos, como em jogos e vídeos para a área da Educação em Ciências. Orientou mais de 100 estudantes do ensino básico ao pós-doutorado.

Além da atividade científica, foi poetisa e integrante da Academia Feminina Mineira de Letras. Como poeta, recebeu vários prêmios, e têm dois livros de poesias publicados, um para crianças (Cedo ou Tarde, 2007, Franco Editora, Juiz de Fora) e outro para adultos (In Minas Memória, 2009, Editora O Lutador, Belo Horizonte). Dentre os prêmios recebidos para a sua poesia, estão: em 1994, Concurso de Poesias Vinicius de Moraes, da Prefeitura do Rio

de Janeiro; Concurso Poesia na Vale da CVRD; em 1995, Prêmio Raul de Leoni, de poesia, da União Brasileira de Escritores (RJ); em 1998 e 2000, prêmios de Poesia da Academia Feminina Mineira de Letras e Medalha de Prata no Concurso de Poesias “Brasil – 500 Anos” (Juiz de Fora, MG). Faleceu em abril de 2015 e deixou um enorme legado que merece ser organizado, analisado e divulgado para futuros projetos de pesquisa em diversas áreas do conhecimento.

Assim, devido à diversidade e riqueza do arquivo pessoal da pesquisadora, este se configurou em objeto de análise deste projeto. Como arquivo pessoal, toma-se aqui a definição proposta por Bellotto (2004, p. 266):

«conjunto de papéis e material audiovisual ou iconográfico resultante da vida e da obra/atividades de estadistas, políticos, administradores, líderes de categorias profissionais, cientistas, escritores, artistas etc. Enfim, pessoas cuja maneira de atuar, agir, pensar e viver possa ter algum interesse para as pesquisas nas respectivas áreas onde desenvolveram suas atividades; ou ainda pessoas detentoras de informações inéditas em seus documentos que, se divulgadas na comunidade científica e na sociedade civil, trarão fatos novos para as ciências, a arte e a sociedade».

Já com relação ao campo da Educação em Saúde no qual Virginia auxiliou a consolidar, este é definido como «[...] um campo multifacetado, para o qual convergem diversas concepções, das áreas tanto da educação, quanto da saúde, as quais espelham diferentes compreensões do mundo, demarcadas por distintas posições político-filosóficas sobre o homem e a sociedade» (Schall; Struchiner, 1999).

Assim, como mulher na ciência, pesquisadora das artes e da educação, contribuiu em muitos campos dentro e fora da saúde, além do desenvolvimento de inúmeros produtos e materiais educativos. Toda esta contribuição e conhecimento delinearão uma melhor compreensão da história da saúde Coletiva e Educação em Saúde no Brasil republicano. Destacar a importância de Virgínia Schall para a Educação em Saúde é recuperar a memória da trajetória e construção desta área no país, à qual ela se dedicou durante quatro décadas. Dessa forma, o trabalho biográfico é capaz de mobilizar atividades de difusão e de construção do conhecimento histórico para amplas audiências. Busca-se, portanto, com esse projeto, delinear pela primeira vez no Brasil por meio de uma biografia, a história e memória do campo da Educação em Saúde e suas inter-relações com a Saúde Coletiva.

### **Memória e Repositórios Temáticos**

Segundo Thompson (1998) toda história depende, basicamente, de sua finalidade social. Há uma série de conformações e conjuntos de documentos de tipo biográfico, ao lado de memórias e autobiografias, que permitem compreender como indivíduos experimentaram e interpretam acontecimentos, situações e modos de vida de um grupo ou da sociedade em geral. Isso torna o estudo da história mais concreto e próximo, facilitando a apreensão do passado pelas gerações futuras e a compreensão das experiências vividas por outros (Thompson, 1998). A partir da relevância da produção acadêmica e de materiais

informativos/educativos, constata-se a importância de buscar respostas fundamentadas em uma metodologia que possibilite a identificação da memória das ações de educação em saúde realizadas no Brasil, além da construção da história de vida e obra de Virgínia Schall.

É neste sentido que um RT foi pensado como forma de sistematização e divulgação do importante trabalho da pesquisadora. Para Márdero Arellano (2008) os repositórios institucionais (RIs) podem ser considerados como sistemas de informação que armazenam, preservam, gerem e disponibilizam o acesso à produção científica de uma instituição e/ou comunidades científicas, por meio de repositórios centrais que são provedores de serviços nacionais e internacionais. Os RIs têm por objetivo preservar a memória institucional e tornar disponível a produção científica da instituição, alinhando-se ao Movimento Internacional de Acesso Aberto ao Conhecimento Científico.

Seguindo o movimento internacional, o Repositório Institucional Arca (<http://www.arca.fiocruz.br/>) foi criado pelo ICICT, unidade da Fiocruz que tem como um dos seus compromissos a disseminação da produção intelectual da instituição. Lançado em 07 de abril de 2011, tem como objetivo acolher e disponibilizar a produção intelectual da Fiocruz buscando uma melhoria contínua deste modelo de preservação da memória institucional, bem como uma maior visibilidade da produção de conhecimento gerada nessa instituição.

A ideia deste projeto é desenvolver um protótipo de RT que integrará o repositório ARCA com a produção acadêmica de Virgínia Schall, bem como com os materiais informativos/educativos produzidos pela pesquisadora.

## **Objetivo**

O objetivo central deste trabalho é apresentar a experiência de construção de um repositório Temático (RT), que procurou organizar e sistematizar a produção acadêmica e de materiais informativos/educativos desenvolvidos por Virgínia Schall. Além disso, almeja-se discutir o conceito de memória e sua relação com os Repositórios Temáticos (RT) no campo da Saúde.

## **Metodologia**

A produção da pesquisadora Virginia Schall foi compilada a partir dos dados contidos no Currículo Lattes da autora. Uma síntese de sua biografia foi resgatada por meio da análise de sua produção registrada no Currículo Lattes, uma fonte declaratória, criada pelo CNPq ([www.lattes.cnpq.br](http://www.lattes.cnpq.br)), que integra as bases de dados de Currículos, de grupos de pesquisa e de instituições em um único sistema de informação do Brasil. Esta fonte se tornou um padrão nacional no registro do percurso acadêmico de estudantes e pesquisadores do Brasil e tem sido adotado pela maioria das instituições de fomento, universidades e institutos de pesquisa do país como instrumento de coleta de dados da produção intelectual do país.

Os dados presentes na fonte foram extraídos por meio da plataforma Stela Experta, aplicativo desenvolvido pelo Instituto Stela contratado pela Fiocruz, o qual realiza uma espécie de harvesting no Currículo Lattes.

Após a extração de toda a produção, estas foram categorizadas e excluídas as duplicidades e produções não autorais (entendidas como as orientações de trabalhos de final de curso, dissertações e teses). Após esta etapa, foi realizada uma padronização das entradas de autoria.

PRODUÇÃO	QUANTIDADE
Apresentação de trabalho	36
Artigo publicado em periódico	175
Capítulo de livro	40
Demais trabalhos relevantes	4
Desenvolvimento de material didático	24
Livro publicado	27
Mídias sociais, blogs, websites	1
Produção artística/ cultural	12
Patentes e registros	2
Produto tecnológico	5
Programa de rádio/tv	5
Texto em jornal e revista	21
Trabalhos publicados em anais de eventos	194
Trabalhos técnicos	18
<b>Total</b>	<b>564</b>

**Quadro 1:** Perfil de tipologias da produção acadêmica de Virgínia Schall extraída do CV Lattes via script

Com o objetivo de triar as produções da pesquisadora na temática específica de Educação em Saúde para posterior inserção do material no ARCA, foi realizada análise das temáticas a partir da leitura dos títulos e palavras-chave dos documentos.

Em seguida à busca e localização dos textos completos dos artigos de periódicos, foi feita a exportação dos metadados de forma automática por meio de script desenvolvido pela equipe do Arca. Nesta etapa, se priorizou a exportação dos metadados das tipologias documentais que fossem de acesso imediato, estando sendo feito em paralelo, a localização das demais, conforme sua acessibilidade: se já digitais serão disponibilizadas a seguir. Caso não estejam disponíveis em versão digital, será feita digitalização, observando-se a Lei de Direitos Autorais vigente no país.

Com relação à disponibilização dos dados, criou-se um espelho que dará acesso a sua produção depositada originalmente no Arca, sem que tenha sido necessário criar outra instância para o RT de Educação em Saúde.

## Resultados

Após a eliminação de duplicatas e de itens descartados da análise, como descrito no item anterior, foram encontrados 564 registros distribuídos em diferentes tipologias de

produção (trabalhos de congresso, artigos de periódicos, capítulos de livros e outras tipologias documentais) e escopo temático (Educação em Saúde, Divulgação Científica).

Para o enquadramento ao RT e buscando delinear a contribuição desta importante pesquisadora para o campo no Brasil, deste total foram separadas todas as produções na temática Educação em Saúde, o que totalizou 313 itens (56% do total de produções), conforme discriminado no Quadro 2. Desta forma, observa-se sua forte presença no campo, seja pela discussão da temática de Educação em Saúde em eventos científicos e outros veículos formais ou pela diversidade de produtos que podem ser considerados de divulgação científica.

Tipologia da produção de Virgínia Schall sobre educação e saúde	Nº	%
Trabalhos publicados em anais	140	45
Artigos de periódicos	84	27
Capítulos de livros	28	9
Material didático	23	7
Livros	13	4
Textos em jornais e revistas de divulgação	10	3
Outras produções (artísticas e bibliográficas)	6	2
Produtos tecnológicos	5	2
Programas rádio, tv	4	1
<b>Total</b>	<b>313</b>	<b>100</b>

**Quadro 2:** Distribuição da produção de Virgínia Schall sobre Educação em Saúde, segundo tipologia documental

Grande parte do material está em processo de digitalização para composição do RT. Até o momento já foram localizados e incorporados ao ARCA 81 (25%) documentos da pesquisadora na área de Educação em Saúde.

### Considerações Finais

O trabalho aqui apresentado representa um primeiro esforço de organização, sistematização e representação da produção intelectual da Pesquisadora e poetisa Virgínia Torres Schall na área de Educação em Saúde.

Esperamos que a evolução deste trabalho possa auxiliar no processo de acesso e compreensão da história e memória do campo de Educação em Saúde e que a construção deste RT possa conferir maior visibilidade e acesso ao patrimônio digital de Virgínia, bem como garantir a preservação e conservação da memória de seu valoroso trabalho no campo de Educação em Saúde no Brasil.

## Referências bibliográficas

ARCA – Repositório Institucional da Fiocruz [Em linha]. Rio de Janeiro. [Consult. 12 jun. 2016]. Disponível na Internet: <URL: <http://www.arca.fiocruz.br/>>.

BELLOTO, Heloísa Liberalli (2004) – Arquivos permanentes: tratamento documental. 2ª ed. rev. e aum. Rio de Janeiro : FGV. 320 p. ISBN: 85-225-0474-1

BORGES, Leandro da Conceição (2015) – Repositório educacional em saúde: homenagem à Virgínia Torres Schall. Rio de Janeiro : Instituto de Comunicação e Informação Científica e Tecnológica em Saúde. 37 f. Monografia de especialização

CNPq – Sobre a plataforma Lattes [Em linha]. Brasília, DF: CNPq. [Consult. 10 Out. 2016]. Disponível na Internet: < URL: <http://lattes.cnpq.br/>>

MÁRDERO ARELLANO, Miguel Àngel (2008) – Critérios para a preservação digital da informação científica. Brasília, DF: Universidade de Brasília. 354 f. Tese de doutorado. [Consult. 22 Out. 2016]. Disponível na Internet: <URL: [http://repositorio.unb.br/bitstream/10482/1518/1/2008\\_MiguelAngelMarderoArellano.pdf](http://repositorio.unb.br/bitstream/10482/1518/1/2008_MiguelAngelMarderoArellano.pdf)>

SCHALL, Virgínia Torres (2015) – Currículo Lattes. [Em linha]. Brasília, DF: CNPq. [Consult. 14 nov. 2015]. Disponível na Internet: < URL: <http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4787216D3>>

SCHALL, Virgínia Torres; STRUCHINER, Miriam (1999) – Educação em saúde: novas perspectivas. Cadernos de Saúde Pública [Em linha]. Vol 15, supl 2. [Consult. 24 Jun. 2016]. Disponível na Internet: <URL: <http://www.scielo.br/pdf/csp/v15s2/1282.pdf>>. ISSN: 1678-4464

THOMPSON, Paul (1998) – A voz do passado: história oral. Rio de Janeiro : Paz e Terra, 385 p. ISBN: 85-219-0309-X



---

# Avaliação de repositórios institucionais brasileiros: Uma proposta de método de avaliação

---

Ana Kelly Alves Lameira

Instituto Brasileiro de Informação em Ciência e Tecnologia

[anakellylameira@gmail.com](mailto:anakellylameira@gmail.com)

## Resumo

Avaliação de repositórios é uma etapa fundamental na construção de um sistema de informação, pois permite quantificar a eficiência do sistema, de modo a garantir que ele acompanhe o dinamismo do objeto de trabalho e dos usuários. Diversos modelos teóricos de avaliação foram propostos na literatura, porém a falta de padronização e clareza na definição das características de avaliação são fatores que dificultam a comparação dos modelos e tornam os resultados das avaliações muitas vezes subjetivos. Desta forma, propõe-se neste trabalho um método padronizado de avaliação de repositórios institucionais, baseado em um questionário a ser aplicado por meio de observação do repositório e entrevista com o gestor. Tanto o questionário quanto a entrevista são formulados a partir de perguntas objetivas, de modo a minimizar a subjetividade da avaliação. A título de validação, este método foi utilizado para realizar a avaliação de sete repositórios institucionais não-universitários. Por meio desta avaliação, identificou-se pontos fracos, como falta de política institucional e de indexação; e, também, pontos fortes, como participação em diretórios internacionais, padronização das páginas e mais de 75% do acervo em acesso aberto.

**Palavras-chave:** Avaliação de Repositórios Institucionais, Indicadores, Comunicação Científica

## Evaluation of Brazilians Institutional Repositories: proposal of an evaluation method

### Abstract

Evaluation of institutional repositories is an important step in the design of an information system, because it allows to measuring the efficiency of the system in order to ensure that it follows the dynamism of the subject work and the users.

Several theoretical models of evaluation have been proposed in the literature, but the lack of standardization and of a clear definition of the evaluation characteristics are factors that make it difficult to compare the models and render the results of the evaluations subjective. Therefore, this paper proposes a standardized method of evaluation of institutional repositories, based on a questionnaire to be applied through examination of the repository and interview with the manager. Both the questionnaire and the interview consist of objective questions, in order to make the evaluation less subjective. As a form of validation, this method was used for evaluating seven non-university institutional repositories. Through this evaluation, we identified weaknesses such as lack of institutional policy and indexing; and also strengths, such as participation in international directories, and more than 75% of documents in open access.

## Introdução

Repositórios Institucionais são bases de dados compostas por coleções digitais de textos completos disponibilizados em acesso aberto, que são desenvolvidas a partir de serviços de gestão relacionados com a coleta, organização, disseminação e preservação da produção acadêmica dos membros de uma instituição (Costa, 2014).

Avaliar repositórios institucionais (RIs) é importante para desenvolver, ou aprimorar estes sistemas (Café, 2016). Assim como a informação e a sociedade da informação são dinâmicas, os sistemas de informação também devem ser dinâmicos para acompanhar o seu objeto de trabalho e os seus usuários. Por isso é importante saber como avaliar estes sistemas de forma a identificar pontos fracos e fortes buscando sempre atualiza-los tanto em informação, quanto em serviços.

Arouck explica que «diversos modelos teóricos para avaliação de sistemas de informação foram propostos na literatura. Entretanto, a maioria dos trabalhos indica os atributos a avaliar, sem definir claramente o modo que estas características devem ser avaliadas» (Myers, Kappelman e Prybutok, 1997, Cit. Por 2001, p. 5). Este trabalho se propõe a resolver este problema apontado por Arouck.

O Ibict realiza avaliações anuais de RIs coletados pelo buscador oasisbr, já dispondo, portanto, de um dispositivo de avaliação de repositórios que vem sendo aprimorado conforme o surgimento de tópicos na literatura. Por isso, o objetivo dessa pesquisa é aprimorar o método de avaliação de repositórios institucionais desenvolvido pelo Ibict definindo claramente o modo de avaliação de cada característica e aplica-lo aos sete repositórios institucionais não universitários coletados pelo oasisbr como forma de validar esse método. A motivação é apresentar um método de avaliação que possa caminhar para um modelo padronizado de avaliação de RIs.

O problema da avaliação de RIs será abordado com base em pesquisa bibliográfica, onde se levantarão características consideradas importantes para o sucesso de um repositório institucional. Em seguida, será elaborado um questionário de avaliação com

essas características. O questionário será aplicado por meio de observação dos RIs e por meio de entrevista com os gestores.

Assim, espera-se iniciar um debate que leve a uma consolidação de critérios de avaliação para RIs e da forma de avaliação. Havendo padronização neste sistema de avaliação, tanto o trabalho do gestor, quanto a busca dos usuários serão facilitados. Isso porque, os gestores saberão como melhorar os seus sistemas, dispondo das características mais importantes dentro dos repositórios. Quanto aos usuários, todo o modelo de avaliação será estruturado visando obter sistemas mais usuais, claros e completos.

## Referencial teórico

Existem algumas formas de avaliar repositórios digitais, uma delas é a avaliação formativa, ou seja, realizada durante o processo de construção do sistema, para detectar problemas com antecedência e ter baixo custo na correção. Outra forma de avaliação é a somativa, ou seja, realizada no final do processo de construção do sistema, para julgar a qualidade de uso e identificar se as metas foram alcançadas (Barbosa; Silva, 2010 Cit. Por Café, 2016). Este estudo é centrado apenas na avaliação somativa, considerando repositórios institucionais já existentes.

O primeiro recorte de tempo realizado para este estudo foi de 1978 a 2000 e foi baseado na revisão de literatura do Arouck (2001). Baseado nos estudos de DeLone e McLean, 1992; Pitt, Watson e Kavan, 1995; e Myers, Kappelman e Prybutok, 1997; Arouck (2001) considera três dimensões de avaliação, são elas: qualidade do sistema, qualidade da informação e qualidade do serviço. Dentro de cada dimensão, existem diferentes características a serem avaliadas. Baseado nos estudos já apresentados, Arouck (2001) considera as seguintes características para a dimensão 'Qualidade do sistema': utilidade, facilidade de uso, confiabilidade, tempo de resposta, integração, flexibilidade e acessibilidade.

A dimensão 'Qualidade da informação' possui várias características citadas na literatura. O estudo do Arouck (2001, p. 15) possui um quadro com todas elas, entretanto, considerando que estas características não são consolidadas dentro da comunidade acadêmica, para este estudo, selecionamos apenas as mais citadas. São elas: legibilidade, oportunidade, relevância, exatidão, completude, utilidade, clareza, confiabilidade e atualização. Finalmente, na dimensão 'Qualidade do Serviço', as características são: confiabilidade, acessibilidade, comunicação, credibilidade, segurança e compreensão.

As definições de cada característica são dadas na Tabela 1.

Característica	Definição	Autor
Acessibilidade	Redução do tempo necessário ao acesso à informação. Também pode ser definido como característica que	Valente & Fujino, 2016, p. 148 ; Cybis; Bettiol;

	afeta especificamente usuários com deficiência.	Faust, 2010 Cit. Por Café 2016, p. 52
Atualização	Trata de critérios de atualização ou revisão dos itens depositados no RI (errata, nova versão) e da definição do que é feito com os itens retirados (descartados) do RI.	Prestes, 2010, p. 78
Clareza	Divulgação de suas políticas de funcionamento, que devem informar sobre: a cobertura de assunto do mesmo; suas comunidades e coleções; como o repositório procede em relação à preservação dos documentos digitais etc.	Prestes, 2010, p. 55
Compreensão	Característica que melhora a utilidade da informação por possibilitar que a informação seja classificada, caracterizada e apresentada de forma a se tornar compreensível. Ligada à clareza e concisão.	Valente & Fujino, 2016, p. 156
Confiabilidade	Credibilidade no conteúdo e na fonte de informação.	Valente & Fujino, 2016, p. 149
Facilidade de uso	Sinônimo de usabilidade: a interface apresenta uso fácil e agradável.	Nielsen, 2012 Cit. Por Café, 2016, p. 49
Flexibilidade	Apresenta possibilidade de atalhos para usuários experientes.	Nielsen, 1995 Cit. Por Café, 2016, p. 51
Segurança	Segurança aos documentos depositados nos repositórios através de técnicas que primam pela preservação de arquivos digitais.	Prestes, 2010, p. 13
Utilidade	Valor de uso da informação na perspectiva de quem ou para que será útil. Apresentação de características necessárias aos seus usuários.	Valente & Fujino, 2016, p. 149 ; Nielsen, 2012 Cit. Por Café, 2016, p. 49
Completude	Definição não encontrada.	

Tempo de resposta	Definição não encontrada.	
Integração	Definição não encontrada.	
Legibilidade	Definição não encontrada.	
Oportunidade	Definição não encontrada.	
Relevância	Definição não encontrada.	
Credibilidade	Definição não encontrada.	
Exatidão	Definição não encontrada.	
Comunicação	Definição não encontrada.	

**Tabela 1** – Características sobre análise de sistemas de informação

O segundo recorte de tempo realizado para este estudo foi baseado nos estudos de Costa (2014; 2015). Nestes trabalhos, a autora realiza uma pesquisa bibliográfica e documental, nas bases de dados Google Scholar e Web of Knowledge, para descrever os elementos que definem o acesso aberto e as características dos repositórios institucionais. Como resultado desta etapa foram definidos cinco critérios de seleção e oito dimensões de análise para os repositórios institucionais. Para esta pesquisa, os cinco critérios de seleção e as oito dimensões de análise serão considerados 13 características. Estas características estão representadas na Tabela 2.

Característica	Definição	Autor
Ser institucionalmente definida	Relaciona-se explicitamente com uma instituição e armazena exclusivamente a produção acadêmica dos membros desta instituição	Costa (2014)
Tratar da produção acadêmica	Tipos de documentos considerados como produção acadêmica: – Artigos científicos; – Pré-prints de artigos científicos; – Trabalhos apresentados em eventos científicos; – Teses e dissertações; – Livros acadêmicos; – Dados primários; – Periódicos eletrônicos; – Material audiovisual acadêmico; – Recursos educacionais acadêmicos.	Costa (2014)

Ser coleções digitais de texto completo	São bases de dados digitais que disponibilizam o texto completo dos documentos.	Costa (2014)
Ser interoperável	Adequação ao protocolo OAIPMH.	Costa (2014)
Ser uma iniciativa de acesso aberto	Disponibilização livremente e sem custos, por meio da Internet pública, publicações acadêmicas, acompanhada de licenças de uso.	Costa (2014)
Tamanho dos RIs	Autora não define.	Lynch e Lippincott (2005); Westrienen e Lynch (2005) Cit. Por Costa (2014)
Tipos de documentos depositados	Autora não define.	Lynch e Lippincott (2005); Westrienen e Lynch (2005); Rieh et al (2007) Cit. Por Costa (2014)
Softwares utilizados	Autora não define.	Lynch e Lippincott (2005) Cit. Por Costa (2014)
Características do sistema	Autora não define.	Rieh et al (2007) Cit. Por Costa (2014)
Responsabilidades administrativas e políticas	Lideranças, Fontes de financiamento, Colaboradores	Lynch e Lippincott (2005); Rieh et al (2007) Cit. Por Costa (2014)
Existência de políticas	Autora não define.	Lynch e Lippincott (2005) Cit. Por Costa (2014)
Serviços oferecidos	Autora não define.	Lynch e Lippincott (2005) Cit. Por Costa (2014)
Áreas do conhecimento	Autora não define.	Lynch e Lippincott (2005); Westrienen e Lynch (2005) Cit. Por Costa (2014)

**Tabela 2** – Características sobre análise de repositórios institucionais

A literatura mostra que existe uma discussão bem desenvolvida sobre critérios de avaliação para sistemas de informação e, especificamente, para repositórios institucionais. Mostrou ainda que o campo de avaliação da usabilidade de RIs caminha para uma consolidação de critérios e métodos. Entretanto, não foram encontrados trabalhos que abordem métodos específicos de avaliação de sistemas de informação, especialmente, repositórios institucionais.

## Metodologia

Esta pesquisa tem como objetivo estruturar um modelo de avaliação e verificar a sua aplicabilidade. O levantamento de requisitos para a avaliação se deu por meio de revisão de literatura sobre os temas «avaliação de repositórios» e «avaliação de sistemas de informação». A busca, pela literatura utilizada nesta pesquisa, foi realizada no buscador *Google Scholar* com finalidade de complementar mecanismo de avaliação já existente no Ibict. A busca foi realizada em fevereiro de 2016.

A literatura sobre avaliação de sistemas de informação mostrou que ainda não existem padrões de métodos de avaliação que possam consolidar dados para estudos comparativos (Myers, Kappelman e Prybutok, 1997; Brancheau, Janz e Wetherbe, 1996; Galletta e Lederer, 1989; Dickson e Wetherbe, 1983; Hamilton e Chervany, 1981 Cit. Por Arouck, 2001). Por isso, neste estudo, optou-se por fazer uma extensa busca de autores que falam sobre o tema e selecionar, de acordo com os estudos destes autores, as características de avaliação que mais se reparam, focando o método de avaliação apenas nessas características.

O método de avaliação de repositórios utilizado pelo Ibict nos últimos anos se baseia em um questionário com características que se dividem em quatro categorias: metadados, visibilidade, políticas e usuários. Mantendo essa categorização, as características encontradas foram distribuídas dentro de cada categoria de acordo com a forma que são utilizadas no processo de avaliação, atribuindo-se assim perguntas objetivas (perguntas de sim-não) para cada característica.

Uma das maiores dificuldades observadas na categorização e detalhamento das características foi justamente a definição das perguntas objetivas que compõem cada característica. Uma vez que a literatura não define estes conceitos, a definição destas perguntas fica a cargo da interpretação pessoal do avaliador, sendo necessária experiência em gerência de RI e um amplo conhecimento deste tipo de sistema para que as melhores perguntas sejam definidas.

A aplicação do questionário se deu por meio de observação dos portais dos RIs e por entrevista com os gestores realizada ao telefone. As avaliações foram realizadas em março de 2016. As respostas obtidas foram utilizadas para calcular a média de cumprimento de cada tópico. A média foi organizada em escala de 0 a 10, onde o desempenho é representado da seguinte maneira:

- 0 a 2,09 corresponde a um desempenho Péssimo;
- 2,1 a 4,09 corresponde a um desempenho Ruim;
- 4,1 a 6,09 corresponde a um desempenho Regular;
- 6,1 a 8,09 corresponde a um desempenho Bom; e
- 8,1 a 10 corresponde a um desempenho Ótimo.

Desta forma, além de levantar o resultado de cada repositório institucional, percebendo os pontos fortes e fracos de cada sistema, o Ibict ainda pôde ter uma visão geral e perceber quais orientações devem ser passadas aos gestores de RIs.

Como modo de verificar a aplicabilidade deste questionário, realizou-se aplicação dele a sete RIs não universitários coletados pelo oasisbr. A escolha de RIs não universitários se deu em virtude do volume desta categoria, que apesar de conter apenas sete RIs, já é suficiente para verificar a utilização do método proposto.

Com base em avaliações de RI já realizadas pelo IBICT e na revisão de literatura, formulou-se um novo questionário, contendo 40 requisitos para avaliação. Os trabalhos utilizados foram duas dissertações, Café (2016) e Costa (2014), uma revisão de literatura, Arouck (2001), e um artigo de revista, Valente e Fujino (2016). O resultado final do questionário, assim como a correspondência entre o questionário e as características encontradas na literatura são apresentados na Tabela 3.

<b>Característica observável (pergunta objetiva)</b>	<b>Característica na Literatura</b>	<b>Como avaliar</b>
<b>Metadados</b>		
<b>Os metadados estão estruturados com Dublin Core?</b>	Ser interoperável; Integração	Observação do portal
<b>O campo dc.type está preenchido com informação sobre o tipo de documento?</b>	Ser interoperável; Integração; Compreensão; Utilidade	Observação do portal
<b>O campo dc.rights está preenchido com informação sobre o direito de acesso ao documento?</b>	Ser interoperável; Integração; Compreensão; Utilidade	Observação do portal
<b>Cada item possui um identificador único que ligue os metadados e os documentos? Esse identificador não deve apontar para ambientes externos ao RI.</b>	Segurança	Observação do portal
<b>Preenchimento dos metadados dc.type e dc.rights é feito conforme as diretrizes especificadas?</b>	Ser interoperável; Integração	Observação do portal
<b>Visibilidade</b>		



Possui estatísticas de acesso?	Utilidade; Oportunidade	Observação do portal
O nome repositório aparece na URL com o termo «repositório»?	Utilidade	Observação do portal
O RI participa de diretórios internacionais como ROAR e OpenDOAR	Legibilidade	Pesquisar nos portais dos diretórios em questão
O nome ou logotipo da instituição está sempre visível no RI?	Ser institucionalmente definida	Observação do portal
As páginas do RI estão padronizadas?	Facilidade de uso	Observação do portal
O RI aparece no ranking webometrics?	Credibilidade	Pesquisar no portal do ranking webometrics
Pelo menos 75% dos recursos textuais que o repositório oferece estão em Acesso Aberto?	Ser uma iniciativa de acesso aberto	Pesquisar por tipo de acesso, quando houver a possibilidade. Ou perguntar em entrevista com o gestor
O repositório está sendo coletado por pelo menos dois portais nacionais e/ou internacionais?	Acessibilidade	Pesquisar nos buscadores Google Acadêmico e Oasisbr por títulos depositados no RI
<b>Políticas</b>		
As políticas utilizadas no repositório estão disponíveis no próprio repositório?	Comunicação; Utilidade; Existência de políticas	Observação do portal
Na política do RI estabelece-se pelo menos quem e o que pode depositar?	Comunicação; Utilidade; Clareza	Observação do portal, quando a política estiver disponível. Ou perguntar em entrevista com o gestor
Há alguma política sobre preservação dos conteúdos?	Segurança; Existência de políticas	Observação do portal, quando a política estiver disponível. Ou perguntar em entrevista com o gestor

Há alguma declaração sobre a missão e os objetivos do RI?	Comunicação; Utilidade	Observação do portal
Identifica-se o uso de licenças, tais como Creative Commons, no RI?	Confiabilidade; Segurança; Legibilidade	Observação do portal
<b>Usuários</b>		
Há uma seção na homepage para informações sobre o RI?	Acessibilidade; Utilidade; Ser institucionalmente definida	Observação do portal da instituição
O RI disponibiliza textos de ajuda para o depósito?	Compreensão; Utilidade	Observação do portal
O RI disponibiliza e-mail e telefone da administração em sua página principal?	Utilidade; Responsabilidades administrativas e políticas	Observação do portal
O RI disponibiliza o mapa do site para navegação?	Compreensão; Utilidade	Observação do portal
O RI permite que o usuário faça login?	Flexibilidade	Observação do portal
O e-mail de contato esta na forma impessoal?	Credibilidade; Legibilidade; Ser institucionalmente definida	Observação do portal
Há algum texto no RI orientando os usuários como citar documentos?	Utilidade	Observação do portal
O RI possibilita o uso de feeds RSS?	Utilidade; Serviços oferecidos	Observação do portal
Há redes sociais ligadas ao RI?	Oportunidade	Observação do portal e pesquisa em redes sociais
Oferece algum serviço (por exemplo, disseminação seletiva)?	Utilidade; Serviços oferecidos	Entrevista com o gestor
O RI permite interfaces em outras línguas?	Compreensão; Utilidade; Oportunidade	Observação do portal
Esclarece quem pode depositar, o que e em que formato?	Clareza; Comunicação; Ser institucionalmente definida	Observação do portal
Possui, por exemplo, possibilidade de aumentar a	Acessibilidade	Observação do portal

<b>fonte?</b>		
<b>Institucionalização</b>		
Há algum esclarecimento sobre direito autoral?	Legibilidade; Confiabilidade; Comunicação; Clareza	Observação do portal
É produção acadêmica / intelectual?	Relevância; Credibilidade; Tipos de documentos depositados	Entrevista com o gestor
Os documentos depositados são produção de membros da instituição?	Ser institucionalmente definida	Entrevista com o gestor
A produção depositada é igual à produção da instituição?	Compleitude; Tamanho dos RIs	Entrevista com o gestor
Participa de iniciativas nacionais (coletado pelo oasisbr, por exemplo)?	Legibilidade; Credibilidade	Pesquisar no portal do oasisbr
É realizado autoarquivo?	Ser institucionalmente definida	Entrevista com o gestor
Há iniciativas de incentivo ao depósito?	Atualização	Entrevista com o gestor
Há uma política de indexação de conteúdos?	Exatidão; Existência de políticas	Entrevista com o gestor
Há vocabulário controlado aplicado ao RI?	Exatidão	Entrevista com o gestor

Tabela 3 – Tradução entre a literatura e o observável

### Análise e interpretação dos dados

Realizadas as avaliações dos sete repositórios institucionais brasileiros não universitários coletados pelo oasisbr, os desempenhos gerais de cada categoria estão dispostos na Tabela 4 e os desempenhos individuais dos RIs estão apresentados na Tabela 5.

<b>Categoria</b>	<b>Média numérica</b>	<b>Média na escala</b>	<b>Nº de requisitos</b>
<b>Metadados</b>	7,71	Bom	5
<b>Visibilidade</b>	6,07	Regular	8

<b>Políticas</b>	4,57	Regular	5
<b>Usuário</b>	3,17	Ruim	14
<b>Institucionalização</b>	4,64	Regular	8
<b>Média ponderada total</b>	4,78	Regular	40

Tabela 4 – Resultado geral da avaliação

Categoria	Repositórios						
	A	B	C	D	E	F	G
<b>Metadados</b>	4	6	10	10	10	10	4
<b>Visibilidade</b>	5	7,5	6,25	3,75	8,75	5	6,25
<b>Políticas</b>	4	4	10	4	4	4	2
<b>Usuário</b>	2,14	2,14	7,92	2,14	2,14	3,57	2,14
<b>Institucionalização</b>	3,75	8,75	5	5	3,75	2,5	3,75
<b>Média ponderada total</b>	3,49	5,24	7,52	4,24	4,99	4,49	3,49
<b>Média na escala</b>	Ruim	Regular	Bom	Regular	Regular	Regular	Ruim

Tabela 5 – Resultado individual da avaliação

Outros pontos de avaliação propostos no trabalho de Costa (2014) podem ser interessantes para levantar características gerais de determinadas áreas (estaduais, nacionais, continentais). Os resultados dos sete RIs avaliados são apresentados a seguir:

- Tipos de documentos predominantes: trabalhos apresentados em congressos; artigos em periódicos; e teses e dissertações.
- Áreas do conhecimento predominantes: não foi possível identificar, pois 6 repositórios classificam as áreas do conhecimento por indexadores livres (não estruturados em linguagem documental) e não pela área geral. Logo, o tempo dessa pesquisa se estenderia caso esse trabalho se ocupasse em classificar cada indexador dentro de suas áreas de conhecimento.
- Quantidade dos documentos: o maior acervo contém 71772 documentos e o menor contém três documentos. O total de documentos contidos nos 7 repositórios é de 115619. A média de documentos por repositório, portanto, é 16517.
- Serviços oferecidos: três repositórios oferecem disseminação seletiva. Os outros não oferecem nenhum serviço.
- Características operacionais do sistema: os sete repositórios operam em DSpace configurados em diferentes versões, a predominante é a versão 4.2.

A aplicação do questionário elaborado neste estudo levantou pontos fracos e fortes. Pode-se perceber, enquanto pontos fracos, que apenas um RI disponibiliza sua política em seu portal. A entrevista com os gestores mostrou que apenas três RIs possuem políticas

próprias institucionalmente definidas. Vale ressaltar que a categoria «Políticas» teve um dos piores desempenhos na avaliação.

Outro ponto fraco identificado na avaliação foi a falta de política de indexação e vocabulário controlado. A literatura mostra que tem surgido uma preocupação com a recuperação da informação nestes sistemas. Em outros tipos de sistemas de informação, tais como bibliotecas físicas e digitais e base de dados, existe uma discussão consolidada sobre indexação e vocabulário controlado como ferramenta de organização e recuperação da informação. Entretanto, nas discussões sobre repositórios institucionais esse tema ainda não é muito abordado. Dentre os sete repositórios abordados, apenas um possui política de indexação e vocabulário controlado.

O contato com os gestores também se mostrou um ponto fraco, nenhum RI possui telefone e e-mail para contato. Três RIs possuem ou e-mail, ou o telefone. Os outros quatro não possuem nem e-mail, nem telefone para contato. Outra fraqueza são os serviços de acessibilidade, nenhum RI avaliado possui qualquer serviço de acessibilidade para pessoas com deficiência. Por fim, destaca-se ainda a falta de clareza e comunicação destes sistemas avaliados sobre quem pode depositar documentos e em que formato.

A categoria «Usuários», que teve o pior desempenho na avaliação, mostrou que há problemas na comunicação e apresentação dos RIs para seus usuários. Os resultados mais baixos nessa categoria tratavam de textos de ajuda para o depósito, disponibilização de e-mail e telefone da administração, explicação sobre quem pode depositar, o que e em que formato e, por fim, serviços de acessibilidade.

Sobre os pontos fortes, destaca-se a participação em diretórios internacionais, seis dentre os sete RIs participam do ROAR ou do OpenDOAR; a padronização das páginas, todos os RIs avaliados possuem suas páginas padronizadas; todos os avaliados possuem um identificador único em cada item que ligue os metadados e os documentos; todos os avaliados possuem apenas documentos de produção científica; e, por fim, mais de 75% do acervo em acesso aberto, cinco RIs possuem esta estatística em seus acervos.

### **Considerações finais**

Este trabalho apresentou uma proposta de avaliação para repositórios institucionais estruturada por meio de pesquisa bibliográfica sobre o tema. A proposta de avaliação se baseia em um questionário formulado a partir das características encontradas na literatura. Este questionário deve ser aplicado em duas vias: observação do RI a ser avaliado e entrevista com o gestor do RI.

A estrutura do questionário apresentou 40 questões que avaliam cinco dimensões dos RIs, são elas: metadados, visibilidade, políticas, usuários e institucionalização. A

aplicação do questionário aos sete repositórios institucionais não universitários e coletados pelo oasisbr teve a intenção de testar o modelo de avaliação estruturado neste trabalho.

Os resultados gerais do teste foram apresentados na Tabela 4. Também foram discutidos os pontos fracos – falta de políticas institucional e de indexação, difícil contato com os gestores e falta de serviços de acessibilidade – e fortes – participação em diretórios internacionais, padronização das páginas e mais de 75% do acervo em acesso aberto. Os resultados individuais foram apresentados na Tabela 5 e mostraram as dimensões fracas e fortes de cada instituição. Esses resultados são de grande interesse a gestores que desejam melhorar seus sistemas.

O questionário resultado deste estudo foi apresentado de forma inteiramente aberta, com explicação sobre como aplicá-lo, para que instituições interessadas possam realizar avaliação de seus repositórios institucionais. O resultado do teste de aplicação mostrou que este instrumento de avaliação atingiu seu objetivo ao mostrar fraquezas e forças de seus objetos de avaliação.

Em trabalhos futuros, propõe-se a integração do método desenvolvido com dispositivos de avaliação de usabilidade, de modo a aprimorar a avaliação dos repositórios institucionais, levando-se em conta o usuário final. O desempenho da categoria «Usuários» sugere que avaliar a usabilidade destes sistemas é necessário para que as equipes dos repositórios institucionais conheçam as necessidades de seus usuários.

## Referências bibliográficas

- AROUCK, Osmar. (2001) – Avaliação de sistemas de informação: revisão da literatura. *Transinformação* [Em linha]. Vol. 13, N° 1. [Consult. 13 Out. 2016]. Disponível na Internet: <<http://bogliolo.eci.ufmg.br/downloads/AROUK%20Avaliacao%20de%20sistemas%20de%20informacao.pdf>>. ISSN 2318-0889
- BAILEY, James E.; PEARSON, Sammy W. (1983) – Development of a tool for measuring and analyzing computer user satisfaction. *Management Science* [Em linha]. Vol. 29, N° 5. [Consult. 13 Out. 2016]. Disponível na Internet: <<http://pubsonline.informs.org/doi/abs/10.1287/mnsc.29.5.530>>. ISSN 1526-5501
- CAFÉ, Luísa C. – Avaliação da usabilidade na interação e recuperação da informação dos usuários pós-graduandos no Repositório Institucional da Universidade de Brasília. Brasília: Universidade de Brasília, 2016. Dissertação de mestrado
- COSTA, Michelli P. da – Características e contribuições da via verde para o acesso aberto à informação científica na América Latina. Brasília: Universidade de Brasília, 2014. Dissertação de mestrado

COSTA, Michelli P. da; LEITE, Fernando C. L. (2015) – Repositórios institucionais de acesso aberto à informação científica: proposta de modelo de avaliação. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde* [Em linha]. Vol. 9, Nº 3. [Consult. 13 Out. 2016]. Disponível na Internet:

<<http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/996/pdf996>>. ISSN 1981-6278

PRESTES, Catarina de Q. (2010) – Construção de políticas para repositórios institucionais: análise da ferramenta OPENDOAR. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2010. Trabalho de Conclusão de Curso

VALENTE, Nelma T. Z.; FUJINO, Asa (2016) – Atributos e dimensões de qualidade da informação nas Ciências Contábeis e na Ciência da Informação: um estudo comparativo. *Perspectivas em Ciência da Informação* [Em linha]. Vol. 21, Nº 2. [Consult. 13 Out. 2016]. Disponível na Internet:

<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2530>>. ISSN 1981-5344

---

## Implementação de métricas alternativas no portal oasisbr

---

Danielly dos Santos Ribeiro

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

[daniellyribeiro@ibict.br](mailto:daniellyribeiro@ibict.br)

Luiza Moreira Camargo

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

[luizacamargo@ibict.br](mailto:luizacamargo@ibict.br)

### Resumo

Objetiva identificar a viabilidade de implementação de ferramenta *altmetrics* para medição de compartilhamento de publicações científicas do Portal oasisbr em mídias sociais. Foi realizado levantamento bibliográfico sobre métodos tradicionais de aferição de impacto e o surgimento de métricas alternativas, assim como os tipos de ferramentas disponíveis no mercado, pagas ou gratuitas. A ferramenta escolhida para teste foi *Altmetric API*. Após sua implementação no oasisbr, verificou-se que não foi calculado em tempo real o compartilhamento de artigos em redes sociais *online*, sendo que o compartilhamento no Facebook não foi contabilizado, enquanto que no Twitter o registro foi efetuado com sucesso. Considerou-se que, embora a ferramenta não tenha tido total aproveitamento no Portal, o uso de métricas alternativas é um caminho favorável tanto para o pesquisador, quanto para ambientes de pesquisa *online*. Dessa forma, torna-se necessário realizar estudos mais aprofundados para que, futuramente, a ferramenta de *altmetrics* seja implantada de forma efetiva no oasisbr.

**Palavras-chave:** Métricas tradicionais. Métricas alternativas. Altmetrics. Ferramentas altmetrics. Portal oasisbr. Mídias sociais.

### Implementation of alternative metrics in Portal oasisbr

#### Abstract:

Aims to identify the feasibility of implementing altmetrics tool for measuring share of scientific publications of Portal oasisbr in social media. It was conducted literature on traditional impact assessment methods and the emergence of alternative metrics, as well as the types of tools available in the market, paid or free. The tool chosen for testing was Altmetric API. After its implementation in oasisbr, It



was found that the tool did not calculate in real time the sharing of articles in *online* social networks, being that the share on Facebook was not recorded, while on Twitter the record was successfully made. It was considered that although the tool have not had full use in the Portal, the use of alternative metrics is a favorable way for both the researcher and for *online* research environments. Thus, it becomes necessary to conduct depth studies so that in future the altmetrics tool can be deployed effectively in oasisbr.

**Keywords:** Tradicional metrics. Alternative metrics. Altmetrics. Altmetrics tools. Portal oasisbr. Social media.

## Introdução

A chegada da *Web* e as consequentes novas tecnologias modificaram o modo de interação entre as pessoas. Assim como aconteceu na sociedade, o meio acadêmico também sofreu modificações, sendo que os cientistas migraram para o ambiente *online*, de forma a expandir os métodos de disseminação do conhecimento.

Os índices tradicionais de aferição do impacto das publicações científicas (NASCIMENTO; ODDONE, 2015, p. 3) para avaliar a extensão de uma publicação são frequentemente utilizados no meio acadêmico. A partir de seu uso, é possível medir o impacto de determinado artigo científico fazendo uso dos principais filtros da ciência (PRIEM et al., 2010 *apud* NASCIMENTO; ODDONE, 2015, p. 4): Contagem de Citação, Avaliação por Pares, Fator de Impacto.

- A Contagem de Citação refere-se a quantas vezes um determinado artigo foi citado. Nascimento (2016, não paginado) subentende que “quanto maior o número de citações de um artigo, maior é o seu alcance dentro do campo da ciência”. Porém, deve-se considerar que alguns autores possuem afinidade com outros, e, assim, as citações de artigos podem sofrer influência devido a essa proximidade, de forma que, por mais que determinado artigo seja citado, ele poderá estar restrito dentro de um grupo limitado de autores (NASCIMENTO, 2016, não paginado);
- Avaliação por Pares tem como objetivo “ajudar a controlar o fluxo de ideias acadêmicas e são recompensados com um entendimento de como o conhecimento científico é transmitido e adotado” (MORENO, 2015, p. 20). Contudo, para Moreno (2015, p. 20), o rastreamento da citação nunca foi capaz de acompanhar os tópicos que muitas vezes são mais importantes para colégios invisíveis, mas possuem menor visibilidade;
- O Fator de Impacto, por sua vez, foi criado em 1955 por Eugene Garfield. Sua função é avaliar a frequência com que um artigo é citado, contabilizando as citações efetuadas 2 anos anteriores a um dado ano (COSTA et al., 2012).

Esses principais filtros da ciência, entretanto, passaram por uma crise, sendo criticados pela comunidade acadêmica. Os questionamentos têm como base a possibilidade de uso impróprio desses indicadores, assim como a lentidão para obtenção de resultados, em oposição ao “atual cenário tecnológico da comunicação científica” (GOUVEIA, 2013 apud NASCIMENTO; ODDONE, 2015, p. 6). A partir desse contexto, surge, em 2010, de autoria de Jason Priem, as métricas alternativas – ou *altmetrics* –, como alternativa a esses questionamentos. Priem *et al.* (2010 apud NASCIMENTO; ODDONE, 2015, p. 4) sugerem que as métricas alternativas poderiam ser uma solução mais apropriada para mensurar e avaliar o impacto da produção científica.

*Altmetrics* é o termo utilizado para designar o “estudo e uso de medidas de impacto acadêmico com base na atividade de ferramentas e ambientes online” (BARROS, 2015, p. 21), o que possibilita medir a popularidade de uma publicação científica não apenas no âmbito acadêmico, mas também no meio social. Por meio delas, os pesquisadores podem identificar onde e por quem seu trabalho foi visualizado, compartilhado e discutido nas redes sociais *online* (NASCIMENTO; ODDONE, 2015, p. 4).

Embora comumente seja utilizado para designar um ambiente de interação e relacionamento *online*, o início do termo “rede social”, na realidade, data do século XVIII (RECUERO, 2009 apud CALIL JUNIOR; CORRÊA; SPUDEIT, 2013), momento em que ocorreu a primeira menção em um estudo científico. De acordo com as relações interpessoais na sociedade, “redes sociais seria o conjunto de relações sociais entre um conjunto de atores e também entre os próprios atores” (CALIL JUNIOR; CORRÊA; SPUDEIT, 2013). Dessa forma, o termo “rede social” não é uma exclusividade do surgimento das mídias sociais do início da década de 2000. Sendo assim, é errôneo utilizar este termo para tratar de ferramentas de interação social *online*. Logo, a designação correta é “mídias sociais”, “redes sociais *online*”, “sites de redes sociais”, “redes sociais na internet”, “ferramentas de redes sociais”, “ferramentas colaborativas” (CALIL JUNIOR; CORRÊA; SPUDEIT, 2013) e também “ferramentas da *Web 2.0*” (JESUS; CUNHA, 2012, p. 114).

A chegada da *Web 2.0* permitiu fazer uso de ferramentas que “possibilitam a criação de canais rápidos, democráticos e segmentados, permitindo que as pessoas compartilhem projetos, gostos pessoais, ideologias e diversos outros tipos de informações utilizando fotos, vídeos e mensagens através de redes sociais, blogs, *bookmarkings* entre outros” (COSTA, 2013, p. 2–3).

Este estudo objetiva apresentar uma análise da viabilidade do uso de métricas alternativas no Portal Brasileiro de Publicações Científicas em Acesso Aberto (oasisbr) para contabilizar compartilhamento de trabalhos científicos em redes sociais *online*, tendo em vista os requisitos do sistema.

## Métricas alternativas

O crescente número de interações dos usuários em redes sociais *online*, além de promover conexões entre pesquisadores e leitores, corrobora para o compartilhamento e a disseminação de informações científicas. As interações como compartilhar, comentar, “curtir”, menções em blogs e registros em gerenciamentos de referências, quando são analisadas acabam produzindo indicadores não tradicionais, surgindo, assim, estudos voltados às métricas alternativas.

As métricas alternativas, ou *altmetrics*, visam complementar as métricas baseadas em citações, apontando sob uma nova perspectiva o modo de medir e avaliar a produção científica. O termo *altmetric* foi cunhado, em 2010, por Jason Priem e posteriormente consolidado em seu artigo “*altmetrics: a manifesto*”, no qual surgiu da necessidade de desenvolver um mecanismo que ajudasse os pesquisadores a filtrar as informações mais relevantes e significantes (NASCIMENTO, 2016). No manifesto, a altmetria é apresentada “[...] como uma alternativa à crise dos três principais filtros da ciência – a revisão por pares – a contagem de citações e o fator de impacto – frente ao movimento de migração dos cientistas para ambiente *online*” (NASCIMENTO, 2016, não paginado).

*Altmetrics* é definida como o estudo e a utilização de medidas de impacto acadêmico baseado em atividade de ferramentas e ambientes *online*, sendo considerado um subconjunto que engloba a Cienciometria e a Webometria (PRIEM; GROTH; TARABORELLI, 2012). Ou seja, “[...] são indicadores da comunicação científica baseados na *Web* social. Não se trata de uma medida única, mas de um conjunto de métricas muito diversos [...]” (SOUZA, 2015, p. 58).

Destacam-se atualmente quatro argumentos favoráveis às métricas alternativas: diversidade de fontes, velocidade, transparência e a menção de novos indicadores (WOUTERS; COSTA, 2012). A primeira refere-se à diversidade das publicações *online*, sendo necessário um conjunto de ferramentas que atendam essa nova demanda. Quanto à velocidade, as métricas alternativas se acumulam de modo mais rápido que as medidas tradicionais de citação (WOUTERS; COSTA, 2012). O terceiro argumento aponta que a coleta de dados por meio de *Application Programming Interface* (APIs) pode ser tão transparente quanto o cálculo utilizado para medir o impacto das publicações (WOUTERS; COSTA, 2012). Ou seja, “*altmetrics* podem apoiar mapas da ciência mais transparentes avaliações mais amplas e equitativas, e melhorias para o sistema de revisão por pares” (BARROS, 2015). E por fim, a altmetria mensura e apresenta indicadores além dos tradicionais (WOUTERS; COSTA, 2012).

As métricas alternativas são utilizadas para calcular o número de compartilhamentos, visualizações e downloads de artigos científicos em revistas, repositórios institucionais, mídias sociais, como *Twitter* e *Facebook*, e em ferramentas de gestão de referências bibliográficas *online*, por exemplo, *Mendeley* e *CiteUlike* (NASCIMENTO, 2016). Desse modo, os dados altmétricos são gerados por meio de sites de mídia social, redes acadêmicas e ferramentas utilizadas para criar, coletar, compartilhar, organizar e gerenciar as informações

(REMER; BORCHARDT, 2015). Os dados coletados pelas ferramentas *altmetrics* são excelentes fontes sobre os hábitos e necessidades informacionais dos pesquisadores, tanto que, atualmente, as métricas alternativas estão sendo consideradas como um novo método de revisão por pares pós-publicação (ELÉSPURU BRICEÑO; HUAROTO, 2016).

Ademais, são caracterizadas por agregarem diversos indicadores que apontam a interação do usuário com o resultado da pesquisa, tais como o número de visualização, compartilhamento, comentários, ou a junção de outros documentos de variadas fontes (NASCIMENTO, 2016). Os indicadores são obtidos de forma imediata, tendo como resultado a implementação das métricas existentes por meio das análises estatísticas das novas mídias e ferramentas, possibilitando o uso de dados quantitativos e qualitativos, para avaliar o artigo, o impacto do periódico e do autor (HAUSTEIN; PETERS, 2012).

A partir do contexto supracitado, é possível vislumbrar a altmetria como um novo método, cujo objetivo é complementar a mensuração da produção científica e monitorar o modo pelo qual os artigos são disseminados e discutidos tanto por pesquisadores quanto pelo público em geral (MARQUES, 2014).

### **Ferramentas altmetrics**

As ferramentas voltadas para as métricas alternativas desempenham a função de agregar indicadores tradicionais e alternativos, possibilitando o monitoramento da interação do usuário com a pesquisa. Contudo, para coletar e armazenar os dados de citações é necessário que cada artigo contenha ao menos um identificador único, por exemplo: Digital Object Identifier (DOI), Pubmed ID, Handle e Uniform Resource Identifier (URI). A ausência desses identificadores dificulta o rastreamento e a menção da produção de um autor ou o desempenho de um artigo (NASCIMENTO; ODDONE, 2015).

Destacam-se, atualmente, quatro empresas cujas ferramentas são as mais utilizadas pela comunidade acadêmica: *Altmetric* da Altmetric.com, *ImpacStory* da *National Science Foundation* and Alfred P. Sloan Foundation, PLOS ALM da *Public Library of Science* (PLOS) e a *Plum Analytics* da empresa EBSCO. Desse conjunto de ferramentas, a *Altmetric* tem sido a que mais se destaca, pois além de atender um público diverso como bibliotecários editores e pesquisadores, oferece também a instalação gratuita, embora limitada.

A *Altmetric* é uma empresa britânica que foi fundada em 2011, sendo considerada uma das primeiras a oferecer ferramentas de métricas alternativas e soluções voltadas para bibliotecários, editores e pesquisadores. Para gerar o indicador, *altmetric attention score*, são rastreadas as menções de publicações acadêmicas em sites de redes sociais *online*, revistas, jornais, blogs, Wikipedia e gestores de referências (NASCIMENTO, 2016). Após a coleta, os dados de interação são calculados por um algoritmo automatizado, que representa uma contagem de acordo com os fatores volume, fonte e autoridade. (NASCIMENTO, 2016).

Os produtos oferecidos pela *Altmetric* são: *Altmetric Bookmarklet*, *Altmetric Explorer* e *Altmetric API*. A *Altmetric Bookmarklet* é um aplicativo que pode ser instalado em diversos

navegadores e possibilita a visualização das métricas dos artigos que contenham um identificador único, como DOI, *Handle*, Pubmed ID ou arXiv ID (ALTMETRIC, 2016).

A *Altmetric Explorer* foi desenvolvida para editores, bibliotecários, autores e instituições. Essa ferramenta coleta e reúne os dados de citações em uma interface, utilizando os filtros palavra-chave, data de publicação, periódico, instituição e autor, para posteriormente recuperá-los.

O *Application Programming Interface* (API) da *Altmetric* permite que os metadados e menções da base de dados *Altmetric* sejam consultados e apresentados na página de cada artigo por meio do ícone “*altmetric donut*” (ALTMETRIC, 2016).

O *Impactstory* possibilita aos pesquisadores criar um perfil, adicionar documentos da sua produção acadêmica e monitorar as menções recebidas. Foi desenvolvido em 2011 por Jason Priem e Heather Piwowar. Em 2012, passou a ser patrocinado *pela National Science Foundation e Sloan Foundation*. No início de 2016, a ferramenta foi remodelada e começou a atuar como um agregador de serviços de outros provedores altmétricos. Com o uso do sistema de identificação de autores ORCID, as informações do perfil e os dados da produção acadêmica são sincronizados de tal modo que, a cada atualização, o *Impactstory* passa a coletar as menções recebidas. Essas métricas são disponibilizadas pela *Altmetric*, e combinadas aos números de leitores no *Mendeley* e referências na *CrossRef*. Além disso, disponibiliza, no perfil do pesquisador, as menções junto com a lista das publicações acadêmicas (NASCIMENTO, 2016).

O periódico *Public Library of Science* (PLOS) foi o primeiro a desenvolver métricas de artigo com a ferramenta *PLOS ALM*. Essa monitora a influência dos seus artigos a curto, médio e longo prazo, utilizando diversos indicadores de impacto como estatísticas de uso, compartilhamento em redes sociais *online*, citações acadêmicas e bookmarks (SOUZA, 2014).

A *Plum Analytics* oferece produtos voltados para universidades e instituições de pesquisas. Sua principal ferramenta é o *PlumX*, no qual coleta e organiza métricas alternativas, agrupando-as em cinco categorias: uso (quantidade de downloads e visualizações), capturas (quantas vezes foi seguido e marcado nos favoritos), menções (quantidade de posts, comentários e revisões), mídia social (tweets, curtidas e compartilhamentos) e citações (SOUZA, 2014).

## Portal oasisbr

O Portal de Publicações Científicas em Acesso Aberto (oasisbr), desenvolvido e mantido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), reúne a produção científica brasileira em um único ambiente online de consulta, sem custos de acesso e de download dos textos completos (OASISBR, [201-]). Através dele, é possível auxiliar pesquisadores na divulgação de seus trabalhos e garantir ao usuário “o acesso a livros, capítulos de livros, teses, dissertações, artigos e publicações apresentadas em

eventos, abrangendo a produção brasileira e também a portuguesa” (SILVA; ALENCAR; SOUZA, 2014, p. 7), por meio do Repositório Científico de Acesso Aberto de Portugal (RCAAP). Para que uma instituição brasileira de ensino e pesquisa participe do oasisbr, ela necessita atender aos critérios de coleta automática que estão relacionados na página do Portal (OASISBR, [201-]).

O oasisbr faz a coleta automática dos metadados a partir de três fontes brasileiras: revistas científicas eletrônicas nacionais, Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) e repositórios institucionais. Ademais, também faz parte da coleta o Repositório Científico de Acesso Aberto de Portugal (RCAAP). Essa parceria portuguesa entre o oasisbr e RCAAP é fruto de acordo firmado entre os Ministros da Ciência e Tecnologia de Portugal e do Brasil (SILVA; ALENCAR; SOUZA, 2014, p. 7).

Para realização da coleta, como critérios gerais para todos os provedores de conteúdo (ou seja, as instituições de ensino e pesquisa participantes), além de manter conexão permanente com a internet, é preciso gerenciar ou publicar informação de natureza científica e/ou tecnológica, sendo que os recursos de informação devem estar disponíveis em texto completo e armazenados no próprio provedor de conteúdos. Além disso, os provedores de conteúdo devem fazer uso do *Open Archives Initiative - Protocol for Metadata Harvesting* (OAI-PMH). Para que ocorra também a interoperabilidade e a padronização das informações entre os sistemas locais e o Portal, exige-se que todos os provedores façam uso de um conjunto mínimo de metadados, tais como: título, autor, resumo, palavras-chave, ano da publicação, tipo de documento e idioma. Como último requisito, é necessária a adoção de metadados que especifiquem o tipo de documento (por exemplo: *dc:type*) no qual o preenchimento deve atender às orientações do vocabulário controlado das Diretrizes DRIVER (OASISBR, [201-]).

Em complemento a esses critérios, existem exigências específicas para cada fonte. A revista científica eletrônica interessada em participar do oasisbr necessita fazer parte da SciELO e registrar suas políticas no Diadorim. Caso ela não faça parte da SciELO, esta deve possuir ISSN; ser de natureza científica e/ou tecnológica; adotar arbitragem por pares; ter indicação explícita do conselho editorial ou similar e que este possua caráter exógeno; manter a periodicidade; possuir, no mínimo, três números publicados e garantir que o acesso aos artigos seja isento de custos e irrestrito. De toda forma, é imprescindível que a revista registre as políticas no Diadorim (OASISBR, [201-]).

No caso das teses e dissertações, o único critério para integrar ao Portal é que elas já estejam presentes na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Já para repositórios institucionais, é necessário que seja de caráter científico e/ou tecnológico, possua acesso aos trabalhos gratuito, acesso aos artigos também seja gratuito e irrestrito, o repositório faça adoção de arbitragem científica e, assim como as revistas científicas eletrônicas, possua indicação explícita do conselho editorial ou similar e que este seja de natureza exógena/multinstitucional (OASISBR, [201-]).

O *software* utilizado pelo Portal é o *VuFind*. Gerenciado pela biblioteca *Falvey Memorial Library, VillaNova University* (Pensilvânia, Estados Unidos), o *VuFind* é uma ferramenta de descoberta e entrega em *software* livre de código aberto (WIKI IBICT, 2016). Com o objetivo de colaborar com a criação de portais integrados de informação geridos por bibliotecas, o *VuFind* permite a integração de ferramentas variadas, criando, assim, um portal de busca consolidado (WIKI IBICT, 2016).

Embora a plataforma tecnológica utilizada pelo oasisbr possibilite o compartilhamento de publicações científicas em mídias sociais, como *Twitter* e *Facebook*, o sistema ainda não apresenta o uso de métricas alternativas para mensurar as citações dos documentos depositados no Portal.

## Metodologia

Trata-se de uma pesquisa aplicada, cuja abordagem de coleta de dados é qualitativa. Foi realizada a leitura de artigos sobre ferramentas voltadas às métricas alternativas, além do estudo e posterior teste para a implementação da ferramenta *Altmetric API* no Portal oasisbr. Desta maneira, a pesquisa segue estruturada em duas etapas: bibliográfica e empírica (teste de implementação).

O levantamento bibliográfico seguiu o seguinte processo: (i) no buscador Google Acadêmico foram utilizados os seguintes termos de busca: "*altmetrics*", "*altmetrics tools and library*", "métricas alternativas and *altmetrics*" "*los indicadores altmetrics*" "*Web 2.0*" "ferramentas da *Web 2.0* e biblioteca"; (ii) a partir dos resultados obtidos foram considerados os artigos que citam sistemas para cálculo de métricas alternativas, sendo recuperados os artigos, em português, inglês e espanhol, publicados entre 2012 e 2016, possibilitando, assim, a seleção e análise teórica das ferramentas *altmetrics*, de forma a localizar aquela que se enquadrasse nos seguintes critérios previamente estabelecidos pelas autoras:

- Ferramenta de *altmetric* desenvolvida para instituições e editores;
- Possibilidade de rastreamento dos metadados mediante identificadores persistentes, a exemplo DOI, *Handle* ou URI;
- Permitir a instalação gratuita, mesmo que limitada.

(iii) seguindo os critérios supracitados, foram identificadas três ferramentas voltadas para instituições e editores; além de possibilitar o uso do identificador persistente: *Plum X metrics*, *Altmetric API* e *Altmetric Explorer*.

Optou-se pela ferramenta *Altmetric API*, pois, segundo a pesquisa realizada em artigos e nos sites das empresas que ofertam os sistemas, somente essa ferramenta permite a instalação de modo gratuito, embora limitado.

## Resultados



A pesquisa resulta no teste da ferramenta *Altmetric API* no Portal oasisbr. A *Application Programming Interface* (API) da *Altmetric* permite que os metadados e as menções da base de dados *Altmetric* sejam consultados e apresentados na página de cada artigo por meio do ícone “*altmetric donut*” (ALTMETRIC, 2016).

Iniciou-se o teste da ferramenta com a inserção do código *Java Script* `<script type='text/javascript' src='https://d1bxh8uas1mnw7.cloudfront.net/assets/embed.js'></script>` e o elemento *Div* `<div class='altmetric-embed' data-badge-type='donut' data-handle='''></div>` no *VuFind*, sendo especificado o identificador persistente “*Handle*” no campo “*data-handle*” da própria ferramenta, para a medição dos dados de compartilhamento. Após a implementação, a ferramenta consulta os metadados dos artigos na base de dados *Altmetric*, considerando os seguintes critérios: DOI, *Handle*, URI, assunto e título do periódico.

As informações coletadas são representadas por três elementos: (I) o *altmetric donut* do artigo, cujas cores representam as fontes de citação; (II) uma lista com as fontes das citações e a quantidade de menções recebidas; (III) e, por fim, calcula quantas vezes o artigo foi salvo nos gerenciadores de referências bibliográficas *CiteULike* e *Mendeley*.

No teste realizado no Portal oasisbr, a ferramenta *Altmetric API* não calculou em tempo real o compartilhamento dos artigos em redes sociais *online*, tendo a espera de 1h para aparecer a menção. Ainda durante o teste, observou-se que o compartilhamento pelo *Facebook* não foi contabilizado pela ferramenta, entretanto, no *Twitter* foi contabilizado com sucesso.

## Conclusão

Entende-se que ferramentas de *altmetrics* no contexto atual de interação no ambiente *online* devem ser vistas como uma forma de aumentar a visibilidade dos portais de pesquisa e dos pesquisadores, sendo que aqueles ganham visibilidade através dos acessos e esses podem avaliar a popularidade de seus trabalhos científicos.

Embora no teste realizado não tenha sido possível contabilizar em tempo real o compartilhamento das publicações do Portal oasisbr em mídias sociais, acredita-se que a experiência da implementação tenha sido válida por apontar as funcionalidades na prática de uma ferramenta de métricas alternativas gratuita. Assim, com base no experimento realizado, torna-se necessária elaboração de estudos aprofundados para que a ferramenta *Altmetric API* seja implementada de forma efetiva no Portal.

## Referências bibliográficas

- ALTMETRICS. Disponível em: [<https://www.altmetric.com/>](https://www.altmetric.com/). Acesso em: 29 ago. 2016.
- BARROS, Moreno. Altmetrics: métricas alternativas de impacto científico com base em redes sociais. *Perspectivas em Ciência da Informação*, v. 20, n. 2, 2015.



CALIL JUNIOR, Alberto; CORRÊA, Elisa Cristina Delfini; SPUDEIT, Daniela. O uso das mídias sociais nas bibliotecas brasileiras: análise dos trabalhos apresentados no SNBU e CBB. XXV Congresso Brasileiro de Biblioteconomia, Documento e Ciência da Informação – Florianópolis, SC, Brasil, 07 a 10 de julho de 2013.

COSTA, Clara Bessa da. Curtiu? : a experiência da biblioteca do Senado Federal no uso das mídias sociais. In: Seminário Nacional de Documentação e Informação Jurídica, n. 4, [2013], Florianópolis, SC.

COSTA, T.; SÍLVIA LOPES, S.; FERNÁNDEZ-LLIMÓS, F.; AMANTE, M. J.; LOPES, P. F. A Bibliometria e a Avaliação da Produção Científica: indicadores e ferramentas. In: Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas. Disponível em: <<http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/429/pdf>>. Acesso em: 02 set. 2016.

ELÉSPURU BRICEÑO, Liliana; HUAROTO, Libio. Los repositorios institucionales como herramientas para medir los indicadores Altmetrics: experiencia de la Universidad Peruana de Ciencias Aplicadas (UPC). In: CONGRESO INTERNACIONAL DE BIBLIOTECAS UNIVERSITARIAS, 5., 2016, Lima. Disponível em: <[http://repositorioacademico.upc.edu.pe/upc/bitstream/10757/617434/1/repositorios\\_institucionales\\_altmetric\\_elsespuru\\_huaroto.pdf](http://repositorioacademico.upc.edu.pe/upc/bitstream/10757/617434/1/repositorios_institucionales_altmetric_elsespuru_huaroto.pdf)>. Acesso em: 25 set. 2016.

HAUSTEIN, S.; PETERS, I. Using social bookmarks and tags as alternative indicators of journal content description. *First Monday*, v. 17, n. 11. 5 nov. 2012. Disponível em: <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/4110>>. Acesso em: 01 set. 2016.

JESUS, Deise Lourenço de; CUNHA, Murilo Bastos. Produtos e serviços da web 2.0 no setor de referência das bibliotecas. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 17, n. 1, p. 110–133, jan./mar. 2012.

MARQUES, Fabrício. Retuíte ou pereça. 2014. Disponível em: <<http://revistapesquisa.fapesp.br/2014/07/15/retuite-ou-pereca/>>. Acesso em: 16 set. 2016.

NASCIMENTO, Andrea Gonçalves do; ODDONE, Nanci Elisabeth. Uso de altmetrics para avaliação de periódicos científicos brasileiros em ciência da informação. *Ciência da Informação Revista*, Maceió, v. 2, n. 1, p. 3–12, jan./abr. 2015.

\_\_\_\_\_. *Altméria para bibliotecários: Guia prático de métricas alternativas para avaliação da produção científica*. Revolução eBook, 2016.

OASISBR. Portal brasileiro de publicações científicas em acesso aberto. Acesso em: 30 set. 2016. Disponível em: <<http://oasisbr.ibict.br/vufind/>>.

OASISBR. Critérios para coleta. Acesso em: 30 ago. 2016. Disponível em: <<http://oasisbr.ibict.br/vufind/Contents/Home?section=how>>

PRIEM, J.; GROTH, P.; TARABORELLI, D. The altmetrics collection. *PloS one*, v. 7, n. 11, p. e48753, 2012.

ROEMER, Robin Chin; BORCHARDT, Rachel. Major Altmetrics Tools. *Library Technology Reports*, v. 51, n. 5, p. 11, 2015.

SILVA, Ana Paula Lopes da; ALENCAR, Lucido Lopes de; SOUZA, Jaqueline Silva de. Iniciativas de acesso aberto à informação científica: uma abordagem à informação científica – oasisbr. In: *Seminário Nacional de Bibliotecas Univesitárias*, 18, 2014, Belo Horizonte. *Anais... Belo Horizonte*, [s.n.], 2014. P. 1–17.

SOUZA, Iara Vidal Pereira de. Almetria: estado da arte. *Sistema Eletrônico de Editoração de Revistas, Minas Gerais*, v. 7, n. 2, p.251–268, 2014. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/viewFile/164/205>>. Acesso em: 20 ago 2016.

WIKI IBICT. Vufind. Acesso em: 30 set. 2016. Disponível em: <<http://wiki.ibict.br/index.php/Vufind>>.

---

## Gestão de dados: Sobreposições ou convergências entre infraestruturas?

---

Pedro Moura Ferreira

Instituto de Ciências Sociais, Universidade de Lisboa, APIS

[pedro.ferreira@ics.ulisboa.pt](mailto:pedro.ferreira@ics.ulisboa.pt)

Bárbara Rodrigues

Instituto de Ciências Sociais, Universidade de Lisboa, APIS

[barbara.rodrigues@ics.ulisboa.pt](mailto:barbara.rodrigues@ics.ulisboa.pt)

### Resumo

No domínio da gestão de dados, o debate sobre a estratégia mais adequada para a curadoria e preservação dos dados de investigação em Portugal tem ganho uma pertinência crescente. A experiência acumulada nesta área aponta para as vantagens de um modelo em rede que articule *soluções de âmbito genérico*, lideradas pelas estruturas de coordenação central, e *soluções de âmbito disciplinar*, dinamizadas pelas infraestruturas de investigação de acordo com as necessidades e desenvolvimentos das diferentes áreas científicas. Nesta comunicação, a atividade dos arquivos de dados em ciências sociais é usada como exemplo para demonstrar as vantagens de conciliar estas duas abordagens.

**Palavras-chave:** Curadoria de Dados; Dados de Investigação; Gestão de Dados; Infraestruturas de Investigação; Preservação Digital.

### Data management: overlapping or convergence between infrastructures?

#### Abstract

In the field of data management, the debate concerning a suitable strategy for the curation and preservation of research data in Portugal has become increasingly pertinent. Previous experience points to the advantages of a network-based model that combines a general approach, headed by the central coordination structures, with a disciplinary approach, driven by the research infrastructures in line with the

needs and changes of the different scientific areas. In this paper, the activities of social sciences data archives will be used to demonstrate the advantages of combining these two approaches.

**Keywords:** Data Curation, Research Data, Data Management, Research Infrastructures, Digital Preservation.

## Introdução

Em Portugal, a curadoria dos dados de investigação constitui um dos campos insuficientemente desenvolvidos do sistema científico e tecnológico. Com efeito, são ainda limitadas as iniciativas neste âmbito, nomeadamente ao nível de infraestruturas que suportem a preservação e a disseminação dos dados. Por sua vez, as iniciativas existentes são bastante diversas, apresentando ora um âmbito *nacional*, se considerarmos o RCAAP; *disciplinar*, no caso das infraestruturas de investigação; e, ainda, *institucional*, no caso de ações como as desenvolvidas pelos Serviços de Documentação da Universidade do Minho e Universidade do Porto.

Tendo em vista responder às atuais e, sobretudo, futuras necessidades de depósito, arquivo, preservação e disseminação de dados de investigação, a articulação entre os diversos atores constitui um ponto que merece ser discutido e aprofundado, tanto mais quando um conjunto diversificado de fatores, como os requisitos de acesso aberto, a exigência de transparência dos processos de investigação, a reutilização dados e os receios de perdas de informação, têm vindo a contribuir para uma nova atitude dos investigadores face aos dados de investigação, ao mesmo tempo que lhes é exigido mais competências em termos de gestão de dados (Corti et al, 2014).<sup>1</sup> Esta mudança no perfil das responsabilidades e competências dos investigadores tem de ser acompanhada pelo desenvolvimento de infraestruturas de dados de investigação, porque, conforme a experiência de outros comprova, a existência dessas mesmas infraestruturas estimula e aprofunda o princípio da reutilização dos dados, contribuindo para maximizar o benefício público da investigação.

## Objetivo

Neste sentido, esta comunicação tem por propósito identificar e explorar as possíveis convergências, complementaridades e eventuais sobreposições entre os repositórios institucionais de dados de investigação e o papel que as infraestruturas de investigação, designadamente as que integram o no Roteiro Nacional de Infraestruturas de Investigação de Interesse Estratégico 2014–2020 da Fundação para a Ciência e Tecnologia, poderão desempenhar ou vir a desempenhar relativamente à gestão dos dados de investigação. O objetivo é esboçar possíveis articulações entre os diferentes tipos de infraestruturas, chamando a atenção para a necessidade de uma moldura institucional coerente e, tanto quanto possível, sob uma orientação unificada e integrada, decorrente de uma política científica relativa aos dados de investigação.

## Metodologia

Do ponto de vista metodológico, a relação entre os repositórios de dados de investigação e as infraestruturas de investigação é analisada e discutida no âmbito das ciências sociais, em particular a partir do caso da infraestrutura *Production and Archive of Social Science Data* (PASSDA), que reúne o Arquivo Português de Informação Social e o Programa das Atitudes Sociais e Políticas dos Portugueses. A análise realizada teve naturalmente em consideração a experiência e a documentação produzida quer no âmbito das infraestruturas, quer no âmbito dos repositórios dos dados, em especial a que foi realizada pelo RCAAP, sem ignorar as contribuições de importantes eventos como as Jornadas da Fundação para a Computação Científica. Considerando, no entanto, o propósito de clarificação do papel dos repositórios disciplinares associados às infraestruturas de investigação torna-se imprescindível recorrer à experiência internacional, especialmente relevante no que se refere aos arquivos de dados de ciências sociais reunidos no consórcio CESSDA e ao qual se espera que Portugal venha a aderir.

## Resultados e discussão

O Repositório Científico de Acesso Aberto de Portugal constitui um projeto-chave do sistema científico e tecnológico nacional e o principal dinamizador do movimento de acesso aberto em Portugal. A missão do projeto é «promover, apoiar e facilitar a adoção do acesso aberto ao conhecimento científico em Portugal e armazenar, disponibilizar e preservar a produção científica» (RCAAP, 2016).

Inicialmente comprometido com a disponibilização de literatura científica, o projeto alargou a sua atividade aos dados de investigação em 2010. A importância do RCAAP como um protagonista da gestão de dados em Portugal foi revigorada em eventos recentes como as Jornadas para a Computação Científica Nacional de 2016 e a Conferência de Dados de Investigação e Ciência Aberta – Rumo a uma Estratégia Nacional no mesmo ano.

Apesar do forte movimento político atual em direção à ciência aberta, é ainda pouco clara a distribuição de papéis e responsabilidades ao nível da gestão de dados em Portugal. Num estudo elaborado no âmbito do projeto RCAAP (Rodrigues, 2010), considerava-se que pelo menos a curadoria dos dados deveria ficar a cargo de profissionais, ao invés dos investigadores. O perfil destes profissionais era considerado ainda pouco definido. Em termos de infraestruturas de suporte, os repositórios institucionais das universidades eram vistos como uma das possíveis respostas para o alojamento de dados, disponibilizando, por exemplo, acesso aos dados produzidos pela «pequena ciência» (Rodrigues, 2010, p.17), constituindo deste modo um *upgrade* dos serviços já existentes. Ou seja, estes repositórios estariam mais vocacionados para dados de alcance mais limitado junto da respetiva comunidade científica, enquanto os dados mais relevantes ou mesmo de maior qualidade estariam a cargo de centros de dados.

Independentemente das soluções referidas, o estudo desaconselhava fortemente a criação de mega soluções para o arquivo de dados, propondo antes uma solução de

equilíbrio entre um *modelo de base genérica* e um *modelo de base disciplinar* (Rodrigues, 2010). Segundo este documento, «a curadoria dos dados científicos, para ser verdadeiramente efetiva e sustentável, exige a participação de todas as partes (os investigadores, as instituições onde trabalham e os organismos de financiamento) envolvidas na produção dos dados e no processo de investigação» (Rodrigues et al., 2010, p.42).

No contexto de um modelo de base disciplinar, os arquivos europeus de dados em ciências sociais representam uma das soluções com mais tradição na gestão de dados de investigação. Os primeiros arquivos foram criados na década de 60 e 70 do século passado – primeiro na Alemanha, seguida dos EUA, Reino Unido e Noruega – pela mão de investigadores empenhados na criação de estruturas que disponibilizassem recursos para a análise comparada e para a análise secundária de dados em ciências sociais. Foi assim que surgiu também, em 1976, o atual *Consortium of European Social Science Data Archive* (CESSDA), (Kaase, 2013), que em 2016 assumirá o estatuto de infraestrutura europeia de investigação (ERIC). O CESSDA tem 24 arquivos associados, dos quais 15 são membros (CESSDA, 2016). A existência desta rede internacional tem permitido uma ação concertada na gestão de dados de investigação, através da partilha de boas práticas, normas e conhecimentos técnicos, mas também na diminuição das barreiras entre países no acesso a recursos de investigação de qualidade.

Orientados para a o apoio à investigação, os arquivos oferecem um conjunto diversificado de serviços além, naturalmente, da função de pesquisa e de disponibilização dos dados, como o apoio à elaboração de planos de gestão de dados, consentimentos informados ou aconselhamento ético, banco de questões, abarcando por isso diferentes fases do ciclo de vida dos dados, desde a sua criação até à sua reutilização. Em muitos casos, os arquivos trabalham em estreita colaboração com as infraestruturas de investigação, como será o caso do Arquivo Português de Informação Social já que faz parte da infraestrutura *Production and Archive of Social Science Data Archive* (PASSDA).

As infraestruturas de investigação em ciências sociais têm ganho um protagonismo crescente nos sistemas científicos e tecnológicos, podendo ser definidas como “instituições duráveis, ferramentas técnicas e plataformas, e/ou serviços implementados para apoiar e promover a investigação enquanto recurso público para a comunidade de ciências sociais” (Renschler, Kleiner e Wernli, 2013, p.1). O investimento em infraestruturas constitui uma matéria consensual ao nível europeu patente na criação de instrumentos como o *European Strategy Forum on Research Infrastructures Roadmap* (ESFRI) e o *European Research Infrastructure Consortium* (ERIC). Tais instrumentos conferem às infraestruturas as condições necessárias para a estabilidade na prestação de serviços ao longo do tempo. A produção e a localização de recursos para a promoção da investigação de excelência, a inovação das práticas e a forte internacionalização das suas redes, constituem alguma das características destas infraestruturas.

Os arquivos asseguram a continuidade do acesso aos dados produzidos pelos projetos de investigação através de um conjunto de atividades de curadoria, preservação e disseminação. Ocupam-se assim de coleções, que incluem, para além das bases de dados, documentação como questionários e relatórios metodológicos.

Na medida em que o objetivo do arquivo de dados é permitir o uso futuro dos mesmos e não somente a sua preservação ou replicação, os arquivos desempenham um conjunto de tarefas diversificadas em comparação com o repositório institucional. Entre estas tarefas, destaca-se a documentação detalhada da informação depositada pelos investigadores com vista a facilitar a sua reutilização.

«Ao pensar na integração de bases de dados abertas, é inevitável abordar a questão dos *standards* e dos metadados. Apenas com esta preocupação com a disponibilização de dados será possível criar uma rede de fontes que permitam a interoperabilidade de dados e que não constitua um “caos informacional”, no qual seja impossível aos investigadores guiarem-se nas buscas pela informação pretendida» (Cardoso, 2012, p. 32).

Os arquivos europeus de dados em ciências sociais baseiam a sua documentação no *Data Documentation Initiative* (DDI). O DDI é um *standard* internacional para a descrição estruturada de dados em ciências sociais, criado em 1995. De acordo com Wackerow e Vardigan (2013), existe um conjunto diversificado de motivos para a utilização de metadados estruturados como seja (1) a preparação de dados para a sua reutilização por terceiros para análise secundária; (2) uma descrição compreensiva da coleção de dados; (3) relacionar estudos longitudinais; (4) o suporte à preservação, já que se baseia na linguagem XML. O DDI é também interoperável, não proprietário e garante a preservação a longo prazo pelo que está mais imune a perdas de informação.

A documentação dos dados encontra-se fortemente relacionada com a função de preservação digital dos arquivos de dados em ciências sociais, permitindo-lhes garantir o acesso às suas coleções no longo prazo.<sup>ii</sup> Porém, além desta documentação, existem ainda outras estratégias de preservação da informação que incluem, por exemplo, a migração de formatos, *backups*, verificação de autenticidade de ficheiros (ex.: *fixity checks*), melhorias nas infraestruturas técnicas, gestão de segurança e avaliação de riscos.

Neste contexto, a utilização de modelos de referência, como o *Open Archival Information System* (OAIS), é comum entre os arquivos do CESSDA. O OAIS é um modelo de referência, passível de ser utilizado por qualquer organização que tenha dados a seu cargo, que define as relações entre as principais atividades e funções relacionadas com a preservação digital – *ingest function*, *archival storage*, *data management function*, *administration function*, *access function* (Lavoie, 2000).

Outro aspeto fundamental no desenvolvimento dos arquivos tem sido o fortalecimento da confiança junto dos investigadores, financiadores e utilizadores de dados, através da certificação dos arquivos. Atualmente, seis dos 15 arquivos do consórcio CESSDA estão certificados pelo *Data Seal of Approval* (DSA). Este «selo» tem por objetivo garantir a

segurança dos dados e uma gestão de dados fiável e de qualidade, baseando-se numa autoavaliação, posteriormente revista por pares (DSA, 2016).

### **Conclusões: convergências entre infraestruturas**

A breve incursão realizada aos arquivos europeus permitiu chamar a atenção para algumas das especificidades disciplinares relacionadas com a gestão de dados de investigação. No caso das ciências sociais, um modelo disciplinar de gestão de dados apresenta vantagens do ponto de vista da (1) proximidade com os investigadores e respetivas infraestruturas de investigação, (2) adequação dos *standards* aos dados disponibilizados, (3) documentação compreensiva, (4) conhecimento de boas práticas, (5) coordenação internacional das inovações e das boas práticas. E, ainda, no acompanhamento de necessidades relacionadas com a investigação que se estendem para além dos dados em acesso aberto; por exemplo a disponibilização de facilidades para análise de dados sensíveis recolhidos por organizações como os institutos nacionais de estatísticas.

Importa, no entanto, referir que uma estratégia nacional – como aquela que está em curso para a Ciência Aberta em Portugal – deve ter também em consideração uma ação concertada entre as diferentes áreas disciplinares. A formação de competências na gestão de dados, a articulação com a política científica nacional e a garantia de persistência dos objetos através da atribuição de DOI aos dados de investigação são exemplos de assuntos de interesse para qualquer infraestrutura de investigação. O ponto, porém, mais crítico para qualquer infraestruturas de investigação, relaciona-se com a preservação digital dos seus objetos no longo prazo. Os dados de investigação e a sua efetiva reutilização requerem que se dê particular atenção à preservação digital de modo a neutralizar os riscos de perda de informação que decorrem da obsolescência tecnológica ou da gestão inadequada dos dados. Só garantindo o acesso continuado aos dados é possível criar condições que satisfaçam a reutilização dos mesmos, contribuindo ao mesmo tempo para uma maior eficiência do sistema de Ciência & Tecnologia, no sentido em que se evita a duplicação de estudos e se promove a análise secundária.

Tendo em conta a necessidade de conciliar interesses comuns e disciplinares, faz sentido referir propostas avançadas no passado no âmbito da atividade do RCAAP e que poderão ser reconsideradas num futuro próximo, se esse for o entendimento prevalecente, agora considerando explicitamente os dados de investigação. Estas propostas sugeriam a criação de um grupo para a preservação e curadoria digital que, entre outros, concretizasse «um projeto piloto, no domínio da preservação digital, com a participação de vários repositórios portugueses, com o recurso a arquitetura(s) que possa(m) dotar os repositórios participantes de ferramentas abrangentes em termos de preservação digital»; avaliasse a «exequibilidade e os termos de uma possível cooperação entre o projeto RCAAP e/ou os repositórios individualmente com o RODA – Repositório de Objetos Digitais Autêntico»; desenvolvesse ou disseminasse «documentos de divulgação, formação e suporte, como



*Briefing papers*, modelos de políticas e procedimentos, boas práticas e casos exemplares de preservação digital» (Ferreira et al., 2012, p. 45–46).

A concretização desta iniciativa pressupõe não só aprofundar um diálogo já iniciado entre os vários protagonistas mas também a elaboração de orientações claras no que respeita à construção das infraestruturas e à gestão dos dados, as quais reclamam e exigem definições mais claras e previsíveis por parte das instâncias responsáveis pela condução da política científica.

## Referências bibliográficas

CARDOSO, Gustavo; JACOBETTY, Pedro; DUARTE, Alexandra (2012) – *Para Uma Ciência Aberta*. 1ª ed. Lisboa : Mundos Sociais. 126 p. ISBN 978–989–8536–07–5.

CESSDA. *National Data Services*. [Em linha]. Bergen: CESSDA AS. [Consult. Set. 2016]. Disponível na Internet: <URL:<http://cessda.net/National-Data-Services>>.

CORTI, Louise [et al.] (2014) – *Managing and Sharing Research Data: A Guide to Good Practice*. 1ª ed. Londres : Sage. 222 p. ISBN 978–1–4462–6725–7.

DSA. Information. *About*. [Em linha]. [Consult. Out. 2016]. Disponível na Internet: <URL:<http://www.datasealofapproval.org/en/information/about/>>.

FCT (2014) – *Portuguese Roadmap for Research Infrastructures*. [Em linha]. [Consult. Abr. 2016]. Disponível na Internet: <URL:[https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese\\_Roadmap\\_of\\_Research\\_Infrastructures.pdf](https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese_Roadmap_of_Research_Infrastructures.pdf)>.

FERREIRA, Miguel; SARAIVA, Ricardo; RODRIGUES, Eloy (2012) – *Estado da Arte em Preservação Digital*. RCAAP. [Em linha]. [Consult. Set. 2016]. Disponível na Internet: <URL:<http://hdl.handle.net/1822/17049>>.

KAASE, Max (2013) – Research infrastructures in the social sciences: The long and winding road. In Kleiner, Brian [et al.] – *Understanding Research Infrastructures in the Social Sciences*. 1ª ed. Zurique : Seismo. ISBN 978–3–03777–133–4. Pt. 1, p. 19–27.

LAVOIE, Brian (2014) – *Meeting the challenges of digital preservation: The OAI reference model*. [Em linha]. OCLC. [Consult. Out. 2016]. Disponível na Internet: <<http://www.oclc.org/research/publications/library/2000/lavoie-oais.html>>.

RCAAP. Sobre o RCAAP. *Missão e Objetivos*. [Em linha] Lisboa: FCT–FCCN. [Consult. Set. 2016] Disponível na Internet: <URL:<http://projeto.rcaap.pt/index.php/lang-pt/sobre-o-rcaap/missao-e-objectivos>>.

RENSCHLER, Isabelle; KLEINER, Brian; WERNLI, Boris (2013) – Concepts and key features for understanding social science research infrastructure. In Kleiner, Brian [et al.] – *Understanding Research Infrastructures in the Social Sciences*. 1ª ed. Zurique : Seismo. ISBN 978–3–03777–133–4. Pt. 1, p. 11–18.

RODRIGUES, Eloy [et al.] (2010) – *Os Repositórios de Dados Científicos: Estado da Arte*. [Em linha]. [Consult. Set. 2016]. Disponível na Internet: <URL:<http://hdl.handle.net/1822/10830>>.

VAN DEN EYNDEN, Veerle; BISHOP, Libby (2014) – *Incentives and motivations for sharing research data: researcher's perspectives* [Em linha]. 1ª ed. Essex : UK Data Archive, University of Essex. [Consult. 14 Jun. 2016]. Disponível na Internet: <URL:[http://repository.jisc.ac.uk/5662/1/KE\\_report-incentives-for-sharing-researchdata.pdf](http://repository.jisc.ac.uk/5662/1/KE_report-incentives-for-sharing-researchdata.pdf)>.

WACKEROW, Joachim; VARDIGAN, Mary (2013) – An established international metadata standard: The Data Documentation Initiative (DDI). In Kleiner, Brian [et al.] – *Understanding Research Infrastructures in the Social Sciences*. 1ª ed. Zurique : Seismo. ISBN 978-3-03777-133-4. Pt. 2, p. 158-167.

---

<sup>i</sup> A gestão de dados é aqui entendida como o conjunto de “práticas, manipulações, melhorias e processos que asseguram que os dados de investigação são de qualidade e que são bem organizados, documentados, preservados, sustentáveis, acessíveis e reutilizáveis” (Corti et al., 2014, p. 2).

<sup>ii</sup>De acordo com a *Digital Preservation Coalition*, entende-se por preservação digital o conjunto de atividades necessárias para garantir o acesso continuado à informação digital, mesmo em caso de falhas ou mudanças tecnológicas.

---

## Acesso aberto à informação científica: diretrizes, políticas e modelos de repositórios científicos para Moçambique

---

**Horácio Francisco Zimba**

Professor Auxiliar, Escola de Comunicação e Artes, Universidade Eduardo Mondlane, Moçambique.

[horacio.zimba@uem.mz](mailto:horacio.zimba@uem.mz)

**Ranito Zambo Waete**

Assistente Universitário, Escola de Comunicação e Artes, Universidade Eduardo Mondlane, Moçambique.

[varela@uem.mz](mailto:varela@uem.mz)

**Aidate Mussagy**

Professora Auxiliar, Departamento de Ciências Biológicas, Faculdade de Ciências, Universidade Eduardo Mondlane, Moçambique.

[amussagy@uem.mz](mailto:amussagy@uem.mz)

### **Resumo**

Esta comunicação discorre sobre as iniciativas de acesso aberto a informação científica com ênfase em propostas de diretrizes, políticas e acções institucionais e/ou governamentais de diferentes países a nível mundial. Identifica as orientações e recomendações emanadas por organismos internacionais como a UNESCO, a IFLA, em busca de um pano de fundo para propor diretrizes, políticas e acções concertadas, que possam ser aplicadas para o contexto de Moçambique. Para consubstanciar as reflexões, primeiro, faz-se uma radiografia da situação actual do país no que diz respeito a existência de iniciativas institucionais e/ou governamentais de acesso aberto à informação científica, tendo como principal fonte de dados, os documentos institucionais e o quadro legal do sector de ensino superior, ciência, tecnologia e inovação de Moçambique. Em segundo lugar, exploram-se as iniciativas institucionais atinentes a estabelecimento de plataformas de acesso aberto, especificamente, revistas e repositórios institucionais de acesso aberto, no contexto

nacional. Em terceiro lugar, resgatam-se as contribuições surgidas nas discussões havidas no I Seminário de Acesso Aberto à Informação Científica, realizado nos dias 9 e 10 de Maio de 2016, na Universidade Eduardo Mondlane, em Maputo, visando o estabelecimento de propostas de modelos de diretrizes e políticas de repositórios científicos de acesso aberto, no contexto moçambicano.

**Palavras-chave:** Acesso Aberto. Informação Científica. Políticas de Informação. Modelo de Repositórios Científicos. Universidade Eduardo Mondlane. Moçambique.

## **Open access to scientific information: guidelines, scientific repositories policies and models for Mozambique**

### **Abstract**

This article discusses the open access initiatives to scientific information with emphasis on institutional and or governmental directives, policies and actions from different countries of the World. The paper identifies the guidelines and recommendations by international bodies such as UNESCO, IFLA, in order to search for a baseline platform to propose directives, policies and concrete actions that may be applicable to the Mozambican context. First, and to substantiate the analyses, an overview of the country's current situation is carried out seeking for existing institutional and or governmental open access initiatives to scientific information. The main data sources are institutional documents and the Mozambican legal framework for high education, science, technology and innovation. Second, the paper explores institutional initiatives aiming at the establishment of open access platforms, mainly scientific journals and institutional open access repositories, within the national context. Third, the results of the discussions from the first Open Access to Scientific Information Workshop, that took place between 9 and 10 May 2016, at the Eduardo Mondlane University, in Maputo, are used to propose directive models and policies for scientific information open access repositories within the Mozambican context.

**Keywords:** Open Access, Scientific Information, Information Policies, Scientific Repository Model, Eduardo Mondlane University, Mozambique.

### **Introdução**

A iniciativa de acesso aberto que nasceu em finais da década de 1990, tem vindo a crescer e a ganhar adeptos e seguidores em vários países á nível mundial. Este movimento foi despoletado inicialmente, por membros da comunidade científica, que entendiam haver necessidade de se encontrar um mecanismo alternativo de publicação dos resultados de investigação, que pudesse manter a mesma qualidade ao já consagrado sistema de publicação em revistas científicas controladas pelas grandes editoras comerciais a nível mundial. Esta iniciativa surgiu, fundamentalmente, em defesa do acesso livre ao

conhecimento produzido nas universidades e institutos de investigação, financiado por fundos públicos. Visto que no modelo tradicional de publicação científica esse conhecimento, regra geral, tem sido publicado em revistas comerciais que ao final acarretam elevadas somas orçamentais para manter as assinaturas para se ter o acesso a informação produzida e publicada nessas revistas.

Portanto, a crise dos periódicos, ocasionada pelos altos custos de manutenção das assinaturas de acesso às principais bases de dados de periódicos científicos, obrigou a que os membros da comunidade científica se juntassem a favor do acesso aberto, através de publicação dos resultados de investigação em plataformas de acesso livre. Neste contexto, a Convenção de Santa Fé, que ocorreu em 1999, nos Estados Unidos da América (EUA), sobre a *Open Archives Initiative* (OAI), despoletou o surgimento de várias outras iniciativas a nível mundial em prol do acesso aberto a informação científica. Como são os casos das declarações de Budapeste no ano de 2002, Bethesda e Berlin, ambas em 2003, que produziram recomendações atinentes a promoção e disseminação do acesso aberto a nível mundial, tendo como principal orientação a disponibilização em acesso aberto, através de auto-arquivamento em repositórios, e/ou publicação em revistas de acesso aberto, dos resultados de investigação produzidos com financiamento público.

Para além das iniciativas despoletadas por comunidades científicas em contextos de certos países conforme os exemplos acima, é importante notar o papel desempenhado por entidades internacionais como é o caso da UNESCO e da IFLA, que também vem apoiando estas iniciativas através de programas e financiamento de acções de promoção e disseminação do movimento de acesso aberto a informação científica. Neste aspecto, destaca-se, de entre outros, os seguintes documentos: “Diretrizes para o manifesto da IFLA/UNESCO sobre a Internet”, publicado em 2006 e atualizado em 2014; e o livro *Open access to knowledge and information: scholarly literature and digital library initiatives; the South Asian scenario*, publicado pela UNESCO, em 2008, (DAS, 2008).

Estas e outras iniciativas referidas em estudos sobre o movimento de acesso aberto, tem servido de base para alavancar o processo de criação de políticas e diretrizes institucionais e nacionais de acesso aberto, em diferentes países. A Declaração do Estoril sobre o acesso à informação, em 2004, e o Compromisso do Minho, em 2006, são exemplos da iniciativa de acesso livre à informação científica em países lusófonos. Segundo dados extraídos do *The Registry of Open Access Repository Mandates and Policies* (ROARMAP), em Outubro de 2016 existiam 786 registos de políticas e mandatos depositados nesta plataforma, distribuídos pelos cinco continentes da seguinte forma: Europa com maior número de políticas e mandatos registados, 477; seguido das Américas com 200; Ásia com 50; Oceania com 40; e África com 19. Dos 19 registos referentes ao continente africano estavam assim disbrituidos: Este de África – sete; Norte de África – dois; Sul de África – oito; e Oeste de África – dois. Sendo que dos oito registos arrolados para o Sul de África todos

eram referentes a África do Sul, que em outras palavras significa que Moçambique e outros países da região Sul de África não possuíam nenhuma política e/ou mandato registado no ROARMAP até essa data (ROARMAP, 2016).

Uma rápida análise ao directório de registo de repositórios institucionais de acesso aberto (*The Directory of Open Access Repositories* – OpenDOAR e no *Registry of Open Access Repositories* – ROAR), percebe-se que a Europa aparece na dianteira, como era de se esperar, tomando em consideração o número de políticas e mandatos depositados no ROARMAP, seguido das Américas e Ásia. Em quarto lugar encontra-se o continente africano, seguido da Oceania (OpenDOAR, 2016; ROAR, 2016). Os dados do OpenDOAR e ROAR indicam que Moçambique possui dois repositórios, o Repositório “Saber”, activo desde 2009 e contendo neste momento cerca de 3100 registos (RIS, 2016); e o Repositório Multinacional, IAPRA – *Irish African Partnership Research Archive*, do qual Moçambique é membro (ROAR, 2016).

Neste contexto, o presente trabalho, propôs-se a analisar os modelos e práticas de implementação de políticas e directrizes para o estabelecimento de repositórios e revistas de acesso aberto no cenário mundial, tendo em vista, a iniciar um debate, a nível de Moçambique, das diferentes abordagens possíveis de implementar no país. De realçar que este trabalho, surgiu na sequência das discussões havidas no seminário de acesso aberto realizado em Maputo, em Maio de 2016. Este seminário, organizado em parceria com a Universidade de Minho, a Fundação para Ciência e Tecnologia de Portugal e o Instituto Brasileiro de Informação em Ciência e Tecnologia, visava consciencializar a comunidade académica de Moçambique sobre a importância do acesso aberto à informação científica e, discutir as estratégias de gestão e disseminação da informação científica por meio de novas ferramentas tecnológicas.

O seminário contou com a participação de várias Instituições de Ensino Superior e de Investigação Científica nacionais e constituiu um marco importante para o estabelecimento do movimento de acesso aberto à informação científica em Moçambique. Assim, as reflexões contidas nesta comunicação servirão de base para futuras discussões, assim como, para proposta de directrizes, políticas e modelos de revistas e repositórios científicos de acesso aberto, concetâneos com a realidade moçambicana.

### **Procedimentos metodológicos**

Este é um estudo tipicamente de análise documental. Para tal, fez-se um levantamento documental, recorrendo-se a pesquisa nas páginas de Internet das principais Instituições de Ensino Superior e de Investigação (IESI) moçambicanas, com o intuito de (i) localizar documentos institucionais relativos a políticas de investigação e/ou de publicação; (ii) verificar a existência de repositórios institucionais e revistas científicas de acesso aberto e; (iii) identificar as plataformas tecnológicas utilizadas para criação de repositórios e revistas de acesso aberto nessas instituições. Para complementar a coleta de informações

recorreu-se a contactos interpessoais, via telefone, com objectivo de perceber se aquelas IESI que não possuíam esta informação disponível na página de Internet, teria-na em outros formatos.

Para fins deste estudo, são consideradas principais IESI, as quatro universidades públicas, nomeadamente, Universidade Eduardo Mondlane (UEM), Universidade Pedagógica (UP), Universidade Lúrio (UNILURIO) e Universidade Zambeze (UNIZAMBEZE); quatro Institutos de investigação científica, designadamente, Instituto de Investigação Pesqueira (IIP), Instituto de Investigação Agrária de Moçambique (IIAM), Centro de Investigação em Saúde de Manhiça (CISM) e Instituto Nacional da Saúde (INS); quatro instituições de ensino superior privadas: Universidade Politécnica (“A Politécnica”), Universidade Católica de Moçambique (UCM), Universidade São Tomás de Moçambique (USTM) e Instituto de Ciência e Tecnologia de Moçambique (ICTEM); e quatro institutos politécnicos públicos, nomeadamente, Instituto Politécnico de Gaza (ISPG), Instituto Politécnico de Tete (ISPT), Instituto Politécnico de Songo (ISPS) e Instituto Politécnico de Manica (ISPM). No total foram pesquisadas 16 IESI. Segundo dados extraídos da página do Ministério de Ciência e Tecnologia, Ensino Superior e Técnico Profissional (MCTESP), em Outubro de 2016, em Moçambique, existiam 49 IES e 20 Institutos de Investigação Científica (IIC), perfazendo assim um universo de 62 IESI, até a data do levantamento realizado para o presente trabalho.

Portanto, a nível de instituições do Governo, realizou-se uma pesquisa documental sobre políticas públicas, leis e planos estratégicos do sector de ensino superior, ciência e tecnologia, visando perceber a abordagem do Governo, em relação aos mecanismos de produção e disponibilização da informação em modelos e/ou plataformas que seguem os padrões do movimento de acesso livre ao conhecimento.

### **Quadro legal e o acesso aberto em Moçambique**

Analisando os diferentes instrumentos legais que criam e estabelecem os sistemas de ensino superior e de ciência e tecnologia moçambicanos, nota-se a preocupação do Governo no desenvolvimento e implantação de infraestruturas básicas de suporte as actividades de ensino e investigação, assim como, de uma plataforma de produção e disseminação de conhecimento científico e tecnológico necessária para a solução dos problemas da sociedade. No entanto, nenhum destes instrumentos aborda de uma forma clara o posicionamento do Governo em relação as forma de disponibilização da informação e conhecimento científico produzido, via financiamento público, ou seja, quais mecanismos devem ser utilizados para publicar e disponibilizar esse conhecimento.

É importante referir que tanto, a política de ciência e tecnologia, aprovada pelo Conselho de Ministros, Resolução n. 23/2003, de 22 de Julho, assim como a estratégia de ciência, tecnologia e inovação aprovada pelo Conselho de Ministros, em Junho de 2006, com horizonte temporal de 10 anos, reconhecem a importância da necessidade de sistematização e publicação da informação e conhecimento produzido pelas instituições de investigação,

mas não explicitam o paradigma (tradicional e/ou acesso livre) de orientação que deve ser seguida para a publicação dos resultados de pesquisa e/ou investigação, desenvolvidas com financiamento público, no contexto nacional.

A mesma constatação, foi observada em outros instrumentos legais, por exemplo, a lei do ensino superior (Lei n. 27/2009, de 29 de Setembro), os decretos que criam e estabelecem os princípios de funcionamentos da maioria das IESI, inclusive da Academia de Ciências de Moçambique (ACM) e do Fundo Nacional de Investigação (FNI), entre outros instrumentos de regulação do sector (MESCT, 2003; MCT, 2006; MCT, 2007; Premgy, 2012).

### **Políticas e directrizes institucionais de acesso aberto em Moçambique**

No contexto mundial vários países e instituições, tem se mostrado favoráveis à iniciativa do acesso aberto, dando suporte tanto a nível de princípios como de práticas, criando instrumentos legais que mandatam a aplicação desta iniciativa. Contudo, existem muitos países que ainda não aderiram a este movimento, ou seja, que ainda não possuem instrumentos legais e normativos, para o estabelecimento de mecanismos de produção e disponibilização do conhecimento científico em acesso aberto. Moçambique está entre esses países. No entanto, reconhece-se que há países onde a iniciativa do acesso aberto tem sido implementada por vias menos formais, ou seja, com ausência de políticas ou normas claras sobre este assunto, tanto institucionais, como das agências de financiamento. Relatos de experiências indicam que, em alguns países as iniciativas de acesso aberto tem sido lideradas pelas universidades e instituições de investigação, Moçambique, parece estar a seguir esse modelo, visto que, a UEM tem sido a instituição promotora e indutora destas iniciativas a nível nacional.

Assim, em 2015, a UEM foi a IESI, pioneira, na aprovação a nível do Conselho Universitário, de uma Política de Publicação (UEM, 2015), que em reconhecimento aos benefícios das publicações em acesso aberto, esta política, contempla um capítulo sobre o acesso livre à informação técnico-científica. Neste capítulo, a UEM, torna claro que privilegia a produção e disponibilização da informação técnico-científica em conformidade com a iniciativa de acesso aberto e com o paradigma de acesso livre à informação científica que, pressupõe a disponibilização de documentos e cópias *online* de acesso livre – sem restrições de licença para uso para a investigação, ensino e outros propósitos – de artigos de revistas científicas, contribuições de conferências, teses, dissertações assim como informações técnicas (relatórios, ensaios, entre outros).

Para a materialização deste desiderato, na sua política de publicação, a UEM compromete-se a: (i) dispor, em ambiente de acesso livre, de revistas editadas pela instituição ou pelas suas unidades subordinadas; (ii) dispor de repositórios institucionais e temáticos, observando o paradigma do acesso livre; (iii) reconhecer a publicação em ambiente de acesso livre para efeitos de avaliação, progressão e promoção académicas e para o acesso ao financiamento para investigação. Neste âmbito, o autor – docente,



investigador ou estudante da UEM –, por sua vez, obriga-se a conceder o direito de acesso gratuito, irrevogável e irrestrito dos conteúdos informacionais por si produzidos, bem como a conceder a licença para copia-los, usa-los, distribuí-los, transmiti-los e exhibi-los, ressaltando, contudo, o respeito, pela integridade e direitos morais sobre os conteúdos devendo estes serem devidamente reconhecidos e citados (UEM, 2015).

Deste modo, ainda que timidamente, a UEM está a dar os primeiros passos para a materialização da iniciativa de acesso aberto em Moçambique, através de acções concretas que resultaram na criação do Repositório “Saber”, em 2009, e da Revista Científica da UEM, em 2012; e, em Maio de 2016, realizou o primeiro seminário de acesso aberto. Entende-se, portanto, que a UEM, gozando do privilégio que lhe é peculiar a nível nacional, através destas acções, poderá exercer: (i) influência positiva a nível da comunidade científica nacional, tanto a nível das universidades, assim como, de entidades do Governo, indústria e organizações da sociedade civil; (ii) liderança das discussões, em foruns apropriados que, poderão determinar os factores importantes para a preparação e adopção de políticas de acesso aberto, a nível nacional; (iii) e servir de incubadora de iniciativas de criação e estabelecimento de plataformas tecnológicas de acesso aberto, entre outros aspectos relevantes para o desenvolvimento desta iniciativa em Moçambique.

### **Repositórios institucionais e revistas científicas de acesso aberto em Moçambique**

Observando elementos da contemporaneidade concernentes ao acesso público ao conhecimento e à informação científica, tendo como ponto de partida a contextualização do acesso aberto à informação, as revistas científicas (RC) e os repositórios institucionais (RI) de acesso aberto disponíveis em Moçambique afiguram-se como um novo modelo de disseminação da produção científica nacional. Desta forma, a organização e o gerenciamento de revistas científicas e repositórios institucionais, principalmente os desenvolvidos em ambientes interinstitucionais, como instituições de ensino superior e de investigação, são ferramentas importantes para o desenvolvimento do acesso aberto à informação científica, bem como a consolidação de políticas públicas de acesso aberto ao conhecimento, no contexto de Moçambique (Waete; Zimba; Mussagy, 2015; Waete et al., 2014; Waete; Moura; Mangué, 2012).

Entretanto, é preciso reconhecer que a história da produção científica é relativamente recente em Moçambique. Por isso, em termos globais e qualitativos, a produção científica nacional ainda é incipiente se comparado, por exemplo, a de países da região ou de outros centros mundiais com tradição científica. Com a excepção da UEM, fundada em 1962, e dos Institutos de Investigação da Saúde e de Investigação Agrária de Moçambique, que são anteriores a proclamação da independência, em Junho de 1975, todas as outras IESI, foram criadas a partir da metade da década de 1980, e atravessaram várias crises económicas e sociais do país, que condicionaram o seu pleno desenvolvimento (Zimba; Mueller 2010). Alia-se o facto de a comunidade científica também ser constituída, na maioria dos casos,

por investigadores geralmente jovens recém-formados. Outro sim, é o financiamento da investigação, que quase na totalidade resulta de apoios externos. Mas, é preciso realçar que nos últimos dez anos, nota-se já uma grande preocupação, por parte do Governo, no desenvolvimento de acções que visam estabelecer plataformas e infraestruturas primordiais para a implementação do sistema de investigação científica e inovação tecnológica nacional (Matos; Mosca, 2009; Zimba; Mueller, 2010; Waete; Moura; Mangué, 2012). Foi, a partir dessa preocupação, que nascem as iniciativas de criação do Repositório Interinstitucional “Saber” e da Revista Científica da UEM.

O levantamento realizado, em Outubro de 2016, nos directórios de registo de repositórios e revistas científicas de acesso aberto, ROAR, *openDOAR*, assim como, nos sites oficiais das próprias IESI e desenvolvedores do *Dspace*, *E-prints* e OJS, principais ferramentas tecnológicas utilizadas para a criação de RIs e RCs, indicou a existência de três RIs (o Repositório Interinstitucional “Saber”, gerido pela UEM, o Repositório da UCM e o Repositório Multinacional IAPRA – *Irish African Partnership Research Archive* – mantido pela *Dublin City University*, da Irlanda e, de duas RCs *online*, que usam a plataforma OJS, nomeadamente, a Revista Científica da Universidade Eduardo Mondlane (RC-UEM), criada em 2012 e a Revista Electrónica de Investigação e Desenvolvimento (REID), da UCM, lançada em 2013. Em termos de RI e RC activos e disponibilizados via plataformas *online* de acesso livre, os resultados indicam que somente a UEM e a UCM, possuíam essas plataformas instaladas e operacionais até ao período de levantamento de dados nos sites ROAR, *openDOAR*, *Dspace*, *Eprints* e OJS. Mas, por outro lado, a pesquisa feita nos sites oficiais das IESI, indicou existirem iniciativas de criação de revistas de acesso livre, publicados em formato impresso mas também disponibilizados integralmente no formato PDF em seus sites. São os casos da Revista do Instituto Nacional da Saúde, da Revista do Instituto de Investigação Pesqueira e das Revistas da Universidade Pedagógica, apenas para citar alguns exemplo. A razão fundamental, segundo percebeu-se nos contactos feitos a algumas destas IESI que possuem revistas nestes moldes, está relacionada com dificuldades técnicas, em alguns casos, e em outros, também associa-se a exiguidade orçamentária para a criação de condições infraestruturais, para a implementação das plataformas *Dspace* e OJS, em suas instituições.

### O Repositório “Saber”

A iniciativa de criação do Repositório Interinstitucional “Saber” (RIS) de acesso aberto à informação científica em Moçambique, foi lançada em 2009, a partir de um esforço conjunto de três instituições, nomeadamente, a Universidade Eduardo Mondlane, a Universidade Politécnica de Moçambique – então Instituto Superior e Politécnico de Moçambique – e o Centro de Formação Jurídica e Judiciária. Esta iniciativa contou com o apoio financeiro da SIDA/SAREC (Agência Sueca para o Desenvolvimento Internacional) e do Projecto QIF (Fundo de Qualidade e Inovação), do Banco Mundial; e com o apoio técnico da Universidade de Minho de Portugal. E teve como principais objetivos, (i) criar uma plataforma

de RI com a finalidade de recolher, agregar e indexar conteúdos académicos e científicos existentes nas instituições de ensino superior e de investigação em Moçambique que pudessem ser disponibilizados em acesso aberto e; (ii) disponibilizar a investigação, descoberta, localização e acesso a documentos de carácter científico e académicos produzidos em Moçambique e/ou sobre Moçambique. Assim, o RIS foi o primeiro repositório institucional, de carácter interinstitucional, criado com o propósito de atender, não só as instituições de ensino superior, como também aos institutos ou centros de investigação (Waete; Moura; Mangué, 2012). Desta forma, o RIS, previa o depósito de todo tipo de materiais como manuscritos, artigos de periódicos (*pré-prints* e *post-prints*), manuais de ensino não publicados, monografias, partes de livros, relatórios técnicos, dissertações, teses e vídeos, etc. Dados extraídos do *openDOAR* em Outubro de 2016 indicavam que o RIS possuía 3103 registos, consituídos na maioria por trabalhos de conclusão de cursos de licenciatura, e algumas teses e dissertações, e outros documentos (artigos de revistas e material de ensino).

Na configuração do RIS adaptou-se a plataforma *Dspace*, uma das mais conhecidas para criação de repositórios institucionais no meio científico. Estruturalmente, o RIS está organizado em comunidades, subcomunidades e colecções. Cada comunidade representa uma das instituições que possui conteúdo indexado no repositório. As subcomunidades correspondem às faculdades, escolas, departamentos ou outros sectores das instituições que disponibilizam informação científica neste repositório. As colecções são as espécies documentais a serem disponibilizadas pelas diferentes comunidades e subcomunidades. Relativamente, ao depósito de textos e documentos é feito mediante solicitação ao sector que faz a gestão do RIS. Actualmente a gestão do RIS é feita pelo Departamento de Gestão de Informação Digital (DGIG) da Direcção dos Serviços de Documentação (DSD) da UEM. Qualquer IESI, de Moçambique, pode submeter os seus conteúdos. Para se efectivar a inclusão de conteúdos no RIS, a solicitação para tal, deve ser acompanhada por um termo de autorização com a assinatura do autor. A qualidade dos materiais depositados é da responsabilidade das comunidades que publicam.

### **A Revista Científica da UEM**

Reconhecendo que a publicação de resultados e a acessibilidade das publicações é uma pré-condição para a eficiência do processo de investigação, a questão das publicações científicas começou a ser objecto de análise, em Moçambique, em particular na Universidade Eduardo Mondlane (UEM), e alguns docentes e investigadores eram, de opinião que deveriam existir revistas locais de qualidade, para permitir a publicação de trabalhos, que pelo seu carácter poderiam ser de interesse endógeno, e por via disso poderiam não encontrar aceitação para publicação em revistas internacionais (Mussagy et al., 2008; Zimba; Mueller, 2010).

Resultados de um estudo realizado em 2007, sobre a situação das revistas publicadas nas UEM, mostraram a existência, naquela altura, de algumas iniciativas em diferentes unidades orgânicas. Mas, praticamente todas padeciam de vários problemas, de ordem estrutural e sustentabilidade, que conseqüentemente viam as possibilidades de se tornarem revistas com qualidade diminutas ou quase nulas. Na altura da realização do estudo, o movimento em torno da iniciativa do acesso aberto vinha ganhando espaço a nível mundial, e a equipa responsável pela implementação da RC-UEM ciente dos benefícios desse paradigma, propôs a criação da RC-UEM, seguindo a filosofia do acesso aberto e a adopção da Plataforma SEER, customizada pelo IBICT. Assim, a RC-UEM foi criada através de uma deliberação do Conselho Universitário (Deliberação n. 13/CUN/2008). De acordo com esta deliberação, a RC-UEM é uma revista de acesso aberto, publicada bianualmente, com objectivo de difundir os resultados das actividades científicas realizadas por docentes e investigadores da UEM e de outras instituições de ensino superior. É publicada em séries temáticas, nomeadamente: (i) Série Ciências Biomédicas e Saúde Pública; (ii) Série Engenharias, Arquitectura, Urbanismo e Tecnologia; (iii) Série Letras e Ciências Sociais; (iv) Série Ciências Naturais e do Ambiente; (v) Série Ciências Agronómicas, Florestas e Veterinárias; (vi) Série Ciências da Educação; (vii) Série Ciências Jurídicas; (viii) Série Ciências Económicas, Administração e Desenvolvimento (UEM, 2007; Mussagy *et al.*, 2008).

### **Modelo de repositórios e revistas científicas de acesso aberto para Moçambique**

Analisando o contexto embrionário do estabelecimento de políticas públicas e institucionais do sector de ensino superior, ciência e tecnologia em Moçambique, fica claro que, existe uma relação directa que condiciona de certa forma a percepção da comunidade científica, no que diz respeito aos modelos de publicação científica, que o país deve orientar suas sinergias. Existe, no seio da comunidade científica nacional, a ideia de que a publicação em revistas internacionais traz mais prestígio e tem sido o indicador fundamental, na avaliação de projectos de financiamento a nível internacional. Por esta razão, poucos investigadores tem interesse em publicar localmente. Associado a isto, as revistas nacionais tem críticos problemas, (i) estruturais; (ii) ausência de recursos humanos capacitados em matérias editoriais, nos seus diferentes níveis, por exemplo, editores científicos, revisores interessados em prestar seu apoio, pessoal técnico de suporte as diferentes actividades pertinentes para o funcionamento do sistema todo; (iii) funcionando na maioria dos casos de forma amadora; e (iv) improvisado nos procedimentos editoriais. Como consequência, as revistas nacionais sofrem do síndrome de discontinuidade precoce. E a alternativa de publicar em revistas internacionais pode concorrer para a marginalização de produção de conhecimento de aplicação endógeno. Se considerado, que este tipo de trabalho pode não encontrar espaço de publicação em revistas internacionais, pois esta publicação, regra geral, divulga assuntos de interesse para o contexto transnacional (Zimba; Mueller, 2010).

No entanto, o levantamento realizado para efeitos de estudo indicou existir um esforço das principais IESI nacionais no sentido de desenvolver plataformas de publicação e disseminação de informação e conhecimento científico, utilizando o modelo de acesso aberto. Em relação a isto, informações a respeito de políticas ou directrizes, repositórios e revistas científicas de acesso aberto existentes a nível nacional, demonstram o estágio embrionário em que Moçambique se encontra no processo do estabelecimento das iniciativas de acesso aberto. Embora, exista um repositório interinstitucional – o Repositório “Saber” –, activo desde 2009, que congrega a produção científica de seis IESI, percebe-se que grande parte dos registos são da comunidade da UEM. As restantes comunidades contribuem com menos de 100 registos dos 3103 registos existentes no RSI (dados colectados em Outubro de 2016). A UCM possui o repositório institucional e uma revista científica, baseados nas plataformas *Dspace* e *OJS*, respectivamente. E a UEM, através da Direcção dos Serviços de Documentação e da Unidade Editorial da Revista Científica, hospeda e gere o Repositório Interinstitucional “Saber”, assim como, o Repositório Institucional e o Portal de Revistas Científicas da UEM.

Portanto, a experiência da UEM, na implementação de iniciativas de acesso aberto em Moçambique, data de 2010, altura do lançamento oficial do primeiro repositório científico em Moçambique, a que se seguiu o lançamento, em 2012, também da primeira revista científica no contexto nacional. Porque era necessário buscar experiências de outros países, principalmente, os de expressão portuguesa, a equipe de implementação da iniciativa de acesso aberto na UEM, encontrou na Conferência Luso-brasileira de Acesso Aberto (CONFOA), o espaço privilegiado para o efeito. Foi neste contexto que depois de participar nas edições de 2013 a 2015, realizou, em Maio de 2016, o primeiro Seminário de Acesso Aberto à Informação Científica, em Moçambique, organizado em parceria com a CONFOA.

Analisando as experiências e iniciativas de implementação de repositórios e revistas científicas, no período que vai desde a criação do primeiro repositório no país, até ao presente momento, e resgatando as contribuições do debate havido no seminário de acesso aberto de Maputo é possível esboçar pelo menos três cenários de modelos de implementação do acesso aberto em Moçambique: (a) repositório interinstitucional, como é o caso do Repositório “Saber”, onde todas as instituições interessadas poderiam se associar formando comunidades deste que seria o portal do conhecimento de Moçambique; (b) repositórios federados, onde cada instituição teria o seu próprio repositório, que depois seria indexado em um repositório agregador gerido pelo centro responsável por criar as directrizes e políticas orientadoras da iniciativa de acesso aberto a nível nacional; (c) repositórios institucionais independentes, obedecendo a políticas e directrizes definidas pelas próprias instituições.

Dos três cenários, o modelo de repositórios federados, parece oferecer melhor abordagem, para o contexto moçambicano. Apesar de se reconhecer que o modelo de

repositório interinstitucional, envolve menor esforço de desmobilização de recursos financeiros e de infraestrutura, e que por isso é mais fácil de implantar e manter, é necessário ter em conta que, em uma situação, onde não existe uma regulamentação clara sobre os papéis dos actores envolvidos, o modelo de repositório interinstitucional, pode-se tornar inviável, e com resultados pouco satisfatórios. O Repositório Interinstitucional “Saber”, pode ser considerado como exemplo dessa situação, a avaliar pela quantidade e qualidade dos documentos depositados nele, assim como, pela forma de gestão não regulamentada, que tem sido característica do mesmo, passados seis anos desde a sua criação.

### **Condições Finais**

O levantamento de dados sobre as iniciativas de acesso aberto no contexto moçambicano revelou existirem algumas iniciativas institucionais e interinstitucionais, dos quais se destacam a criação do Repositório Interinstitucional “Saber”, a criação das Revistas Científicas da UEM e da UCM, e mais recentemente, a criação dos repositórios institucionais também da UEM e da UCM.

No que diz respeito a directrizes e políticas de acesso aberto, merece destaque, a iniciativa da UEM, que aprovou em 2015, a sua política de publicações, que apresenta um capítulo específico sobre o acesso aberto. Mas, deve-se reconhecer que uma política específica sobre o acesso aberto, que atenda aos requisitos e directrizes internacionais de acesso aberto seria provavelmente o caminho ideal, para o efectivo estabelecimento do movimento de acesso aberto a informação e conhecimento científico produzido na UEM. Ainda sobre o mesmo aspecto, mais do que políticas institucionais a nível das universidades e institutos de investigação, seria importante que o Governo através das entidades de tutela e de financiamento do sector de ensino superior, ciência e tecnologia, aprovasse directrizes, políticas e mandatos específicas para incentivar o depósito, em repositórios científicos nacionais, de toda a produção científica e tecnológica, desenvolvida no contexto nacional e com financiamento público.

A UEM em particular e as entidades tuteladas ou subordinadas do MCTESTP, como sejam, a Direcção Nacional de Ciência e Tecnologia (DNCT), a Direcção Nacional de Ensino Superior (DNES), a Direcção Nacional de Tecnologias de Informação e Comunicação (DNTIC), o Fundo Nacional de Investigação (FNI) e a Empresa Nacional de Parques de Ciência e Tecnologia (ENPCT), podem desempenhar um papel muito importante no estabelecimento de directrizes, políticas e de uma infraestrutura tecnológica e de suporte, para a implementação efectiva das iniciativas de acesso aberto, a nível institucional e central: (i) a UEM, porque possui recursos humanos com experiência de mais de seis anos, na implementação de repositórios, revistas e directrizes e políticas de acesso aberto; (ii) o FNI, visto ser o principal financiador da investigação a nível nacional; (iii) a DNCT e a DNES, principais actores de definição de políticas e estratégias do sector; e (iv) a DNTIC e a ENPCT, principais actores de

implementação da infraestrutura tecnológica necessária para a implantação dos diferentes serviços de informação.

Contudo, é importante realçar que, do levantamento documental realizado nos sites das diferentes IESI, não foi possível identificar a existência de documentos normativos ou de políticas, que apontem para o estabelecimento ou institucionalização do movimento de acesso aberto. Mas é de se reconhecer, que existem acções atinentes a criação de repositórios e revistas de acesso aberto em algumas IESI, mas dada as limitações, relacionadas fundamentalmente ao fraco domínio das ferramentas tecnológicas necessárias para esse efeito, estas acções na maioria das vezes resultam em fracasso absoluto. Isto remete a um desafio que as IESI tem no sentido de encontrar uma plataforma de colaboração interinstitucional, com vista a maximizar o uso dos recursos e a transferência de *know how* existente na UEM, neste domínio, de forma a catalizar o processo de desenvolvimento da plataforma de acesso aberto em Moçambique.

### Referências bibliográficas

BERLIN Declaration on Open Access to Knowledge in the Sciences and Humanities. (2003). [Consult. 14 Out. 2016]. Disponível na Internet: <<https://openaccess.mpg.de/Berlin-Declaration>>.

BETHESDA Statement on Open Access Publishing. (2003). [Consult. 14 Out. 2016]. Disponível na Internet: <<http://legacy.earlham.edu/~peters/fos/bethesda.htm>>.

BOAI – Budapest Open Access Initiative. (2002). [Consult. 14 Out. 2016]. Disponível na Internet: <<http://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese>>.

COMPROMISSO do Minho compromisso sobre o acesso livre à informação científica em países lusófonos. (2006). [Consult. 14 Out. 2016]. Disponível na Internet: <<https://www.ufrgs.br/blogdabc/compromisso-do-minho-pelo-acesso-livre/>>.

DAS, Anup. K. (2008) – *Open access to knowledge and information: scholarly literature and digital library initiatives; the South Asian scenario*. (Eds: Bimal Kanti Sen and Jocelyne Josiah), New Delhi: UNESCO, 2008. [Consult. 14 Out. 2016]. Disponível na Internet: <<http://unesdoc.unesco.org/images/0015/001585/158585e.pdf>>.

DECLARAÇÃO do Estoril sobre o acesso à informação. (2004). [Consult. 14 Out. 2016]. Disponível na Internet: <<http://www.apbad.pt/Downloads/DeclaracaoEstoril.pdf>>.

IFLA. (2006) – *Directrizes para manifesto IFLA/UNESCO sobre a Internet*. [Consult. 14 Out. 2016]. Disponível na Internet: <http://www.ifla.org/files/assets/faife/publications/policy-documents/internet-manifesto-guidelines-pt.pdf> >.

MATOS, Narciso.; MOSCA, João. (2009) – Desafios do ensino superior. In: Brito, L. de *et al.* (org.). *Desafios para Moçambique 2010*. Maputo: IESE, 2009. P. 297–319.



- MOÇAMBIQUE. CONSELHO DE MINISTROS. (2003) – *Decreto 23/2003 de 22 de Julho*. Maputo.
- MOÇAMBIQUE. MINISTÉRIO DE CIÊNCIA E TECNOLOGIA. (2006) – *Estratégia de ciência, tecnologia e inovação de Moçambique*. Maputo: MCT.
- MOÇAMBIQUE. MINISTÉRIO DE CIÊNCIA E TECNOLOGIA. (2007) – *Compilação legislativa sobre ciência e tecnologia de Moçambique*. Maputo: MCT.
- MOÇAMBIQUE. MINISTÉRIO DE ENSINO SUPERIOR, CIÊNCIA E TECNOLOGIA. (2003) – *Política de ciência e tecnologia*. Maputo: MESCT.
- MUSSAGY, Aidade.; MANGUE, Manuel. V.; NHARRELUGA, Rafael.; ZIMBA, Horácio. F. (2008) – *Proposta de criação da revista científica da Universidade Eduardo Mondlane*. Maputo: UEM.
- OpenDOAR. (2016) – *The Directory of Open Access Repositories*. [Consult. 14 Out. 2016]. Disponível na Internet: <<http://www.opendoar.org/find.php>>.
- PREMUGY, Cassamo. I. C. (2012). *Colectânea de legislação do ensino superior*. Maputo: MCT.
- RIS. (2016) – Repositório Interinstitucional Saber. [Consult. 14 Out. 2016]. Disponível na Internet: <<http://www.saber.ac.mz>>.
- ROAR. (2016) – Registry of Open Access Repositories. [Consult. 14 Out. 2016]. Disponível na Internet: <<http://roar.eprints.org/>>.
- ROARMAP. (2016) – *The Registry of Open Access Repository Mandates and Policies*. [Consult. 14 Out. 2016]. Disponível na Internet: <<http://roarmap.eprints.org/view/country/710.html>>.
- THE SANTA Fé Convention For The Open Archives Initiative. (1999). [Consult. 14 Out. 2016]. Disponível na Internet: <[http://www.openarchives.org/sfc/sfc\\_entry.htm](http://www.openarchives.org/sfc/sfc_entry.htm)>.
- UEM (2007) – *Reflexão sobre disseminação/publicações na Universidade Eduardo Mondlane*. Maputo: UEM.
- UEM (2015) – *Política de Publicação da Universidade Eduardo Mondlane*. Maputo: UEM.
- WAETE, R.anito Z.; MUSSAGY, Aidade.; ZIMBA, Horácio. F.; MANGUE, Manuel. V. (2014) – O impacto das revistas científicas e dos repositórios institucionais de acesso aberto em Moçambique: o caso da revista científica da universidade eduardo mondlane e do repositório “Saber” (poster). In: *Conferência Luso-Brasileira sobre Acesso Aberto*, 5, 2014, Coimbra. Actas. Coimbra: Universidade de Coimbra.
- WAETE, Ranito. Z.; MOURA, Maria. A.; MANGUE, Manuel. V. (2012) – Acesso livre à informação científica em países em desenvolvimento: o caso dos Repositórios “Saber” (Moçambique) e Digital da UFMG (Brasil). *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação.*, Vol.Florianópolis, v. 17, Nn. Esp. 2, p. 195–210.



WAETE, Ranito. Z.; ZIMBA, Horácio. F.; MUSSAGY, Aidate. (2015) – Avaliação do nível de percepção e aderência da comunidade acadêmica da UEM à iniciativas de publicações de acesso aberto (poster). In: *Conferência Luso-Brasileira sobre Acesso Aberto, 6, 2015, Salvador*. Actas . Salvador: Universidade Federal de Bahia.

ZIMBA, Horácio. F.; MUELLER, Suzana P. M. (2010) – A dimensão política e processo de institucionalização de ciência e tecnologia em Moçambique. In: *Encontro Nacional de Pesquisa em Ciência da Informação, 11., 2010, Rio de Janeiro*. Actas. Rio de Janeiro : ANCIB. [Consult. 14 Out. 2016]. Disponível na Internet: <<http://enancib.ibict.br/index.php/enancib/xienancib/paper/viewFile/3543/2668>>.

---

## Social Dendro: Aplicação de técnicas das redes sociais à gestão de dados de investigação

---

João Rocha da Silva

Faculdade de Engenharia da Universidade do Porto / INESC - TEC

[joaorosilva@gmail.com](mailto:joaorosilva@gmail.com)

Nelson Pereira

Faculdade de Engenharia da Universidade do Porto

[ei09025@fe.up.pt](mailto:ei09025@fe.up.pt)

### Resumo

A gestão de dados de investigação é atualmente um grande desafio para as instituições de investigação. Os pequenos grupos de investigação, em particular, necessitam de apoio na gestão dos seus dados para serem capazes de corresponder aos requisitos de dados abertos que começam a ser impostos pelas entidades financiadoras. Dado que o trabalho de investigação é tipicamente desenvolvido em ambiente colaborativo, é importante que seja suportado por ferramentas desenhadas para suportar essa colaboração. É neste sentido que tem vindo a ser desenvolvida a plataforma Dendro para gestão de dados de investigação, uma ferramenta open-source, fácil de instalar e usar por pequenos grupos de investigação, que pretende facilitar o armazenamento e descrição dos dados de investigação em preparação para o seu depósito em praticamente qualquer repositório final. Atualmente está a ser desenvolvida uma extensão à plataforma Dendro, que pretende incorporar diversos conceitos das redes sociais na plataforma (incluindo gostos, comentários e partilhas), com o objetivo de tornar a descrição de dados mais dinâmica e fácil de realizar dentro do grupo de investigação. Após o desenho de esboços de interações, o Social Dendro está neste momento em fase de implementação e testes junto de alguns grupos de investigação. Nesta comunicação será apresentado o modelo de dados da plataforma, alguns exemplos de interação e o estado atual de desenvolvimento da

aplicação; terminaremos com algumas breves conclusões e algumas perspectivas de trabalho futuro.

**Palavras-chave:** Gestão de dados de investigação; repositórios de dados; redes sociais; interação; interfaces com o utilizador

## **Social Dendro: Applying social network techniques to research data management**

### **Abstract**

Research data management is currently a large challenge for research institutions. Small research groups, in particular, need support in the management of their data in order to comply with the open data requirements that start to be imposed by funding institutions. Since research activities are typically carried out in a collaborative environment, it is important that it is supported by tools designed to support that collaboration. It is with this in mind that we have been developing the Dendro research data management platform. It is open-source, easy to install and use even by small research groups, which facilitates the storage and description of research data in preparation for their deposit in practically any repository. We are presently developing an extension that intends to incorporate several concepts from social networks into the platform (including likes, comments and shares) with the goal of making data description more dynamic and easy to carry out within the research group. After the initial outline of the interactions with the extension, Social Dendro is currently in development and testing stage with several research groups. In this paper, we will present the platform's data model, some interaction examples and its current state of development. We will finish with some conclusions and perspectives of future work.

**Keywords:** Research data management; data repositories; social networks; interaction; user interfaces

## **Introdução**

A gestão de dados de investigação é cada vez mais uma preocupação dos investigadores, talvez devido à presença de cada vez mais recomendações (COMMISSION, 2013) ou até mesmo imposições (FOUNDATION, 2011) por parte das entidades financiadoras no sentido da disponibilização dos conjuntos de dados que servem e suporte às publicações resultantes dos projetos por elas financiados. Muito devido a estas novas políticas, tem sido possível a assistir a uma mudança na atitude dos investigadores relativamente à gestão dos seus dados de investigação. Em particular, a partilha de dados tem aumentado em grande medida (TENOPIR et al., 2011, TENOPIR et al., 2015).

Apesar destes sinais muito positivos, a descrição e partilha de dados de investigação ainda está longe de estar perfeitamente enraizada nos processos de investigação. Diversos estudos indicam, contudo, que as melhorias devem passar não só pelas ferramentas de

gestão e partilha como também pelo melhor reconhecimento do trabalho dos investigadores que partilham dados (TENOPIR et al., 2015, WALLISROLANDO e BORGMAN, 2013). O problema da gestão de dados é um tema muito complexo, que inclui tanto questões técnicas como políticas, sociais e de propriedade intelectual (BORGMAN, 2012).

Nas instituições de investigação onde existem recursos para tal, a gestão dos dados de investigação pode ser suportada por especialistas denominados curadores de dados (PENG et al., 2016); em outros casos, os investigadores acabam por realizar eles próprios a descrição e depósito dos seus conjuntos de dados. Para obter os melhores resultados, contudo, tanto curadores com investigadores devem participar no processo de descrição e depósito, pois desempenham funções complementares. Os curadores são peritos na área de gestão de dados, mas não propriamente nas áreas científicas de que provêm esses mesmos dados; por outro lado, os investigadores têm profundo conhecimento das áreas científicas em questão, mas não sabem necessariamente como formalizar esses metadados de acordo com as melhores práticas de gestão de dados.

Para que seja possível criar metadados de elevada qualidade, capazes de fomentar a descoberta e reutilização dos conjuntos de dados é, portanto, necessário construir *workflows* que integrem o trabalho dos curadores e dos investigadores, desde o momento da produção dos dados. É portanto necessário motivar a participação dos investigadores através da oferta de soluções que lhes proporcionem vantagens imediatas no seu dia-a-dia, que podem passar pelo simples armazenamento seguro dos dados e sua descoberta mais fácil dentro do grupo de investigação (POSCHEN et al., 2012). Outras vantagens como o indiscutível aumento das citações originada pela partilha de dados base (PIWOWAR e VISION, 2013) são muito importantes mas ainda pouco palpáveis para os investigadores—especialmente quando consideramos que a produção de registos de metadados para partilha com elementos externos implica um investimento muito maior na sua descrição (BORGMAN, 2012).

Como já demonstrado em projetos como o ADMIRAL (HODSON, 2011), o DataStaR (STEINHART, 2011) ou o MaDAM (POSCHEN et al., 2012) a gestão de dados deverá estar presente em todo o processo de investigação, focando-se na organização dos dados desde a sua criação até à sua disseminação, passando pelo seu armazenamento e publicação. A fase de publicação representa o passo final do trabalho dos investigadores, onde os dados produzidos são depositados em repositórios reconhecidos pela comunidade científica, tornando possível a revisão por pares.

O processo de investigação implica a comunicação entre os elementos do grupo, assim como a partilha de ficheiros entre eles. Esta colaboração é tipicamente levada a cabo por e-mail ou conversa online, enquanto a partilha de ficheiros é feita com recurso, por exemplo, a unidades de armazenamento USB, partilha na rede ou soluções de armazenamento na nuvem, como é caso da Dropbox. Apesar de serem rápidos e simples de usar, estes processos podem fazer com que os dados fiquem dispersos e separados das

eventuais notas relativas ao seu contexto de produção. Por outras palavras, os dados armazenados ficam separados dos seus metadados, que acabam por se perder em cadeias de e-mail trocados entre elementos dos grupos de investigação ou outros contatos informais. Ao analisar o tempo gasto na escrita de todas essas mensagens e emails conclui-se que o trabalho de descrição acaba por ser feito, mas não de forma centralizada, sistemática ou interoperável.

O Dendro<sup>i</sup> é uma plataforma de gestão de dados de investigação atualmente em desenvolvimento, que permite aos grupos de investigação armazenar, descrever e partilhar os seus dados de investigação nos principais repositórios (AUTOR, 2014). A plataforma permite aos utilizadores criar projetos, que são semelhantes às pastas partilhadas na Dropbox. Dentro de cada projeto poderão ser depositados ficheiros e criadas pastas. Contrariamente à Dropbox, contudo, há também um forte foco na produção de metadados de qualidade. Para tal, a plataforma usa ontologias genéricas e específicas dos domínios de investigação no seu modelo de dados (AUTOR, 2014). A plataforma também inclui um sistema que ajuda os investigadores a escolher os descritores mais adequados para os seus ficheiros e pastas, tendo por base as suas interações com o mesmo (descritores mais usados no projeto, por exemplo) (AUTOR, 2016).

Para tentar direcionar os esforços dos investigadores para a adequada descrição dos seus dados e incorporá-la nas suas atividades diárias, começaram recentemente a ser propostas algumas abordagens colaborativas assentes em paradigmas das redes sociais. O conceito “Science 2.0 Repositories” traz para o processo de descrição e partilha de dados de investigação algumas metáforas de interação normalmente presentes nas redes sociais. Estas incluem, por exemplo os “gostos”, comentários e partilhas, mas associados a eventos como a criação ou descrição de conjuntos de dados de investigação (ASSANTE et al., 2015). Esta visão de repositório colaborativo e social alinha-se com um importante foco de desenvolvimento do Dendro: a produção de metadados flexíveis e com histórico completo de edição (AUTOR, 2014). Dado que esta informação de auditoria sobre os metadados é parte integral do Dendro, uma extensão social deste tipo pode tirar grande partido da estrutura e modelo de dados já existente.

## **A extensão Social Dendro**

A extensão Social Dendro tem como objetivo trazer para a plataforma Dendro alguns conceitos das redes sociais para melhor suportar a gestão de dados de investigação. A extensão social irá atuar sobre as alterações aos metadados de ficheiros (edição, adição e remoção) bem como a adição, edição e remoção de ficheiros e pastas de um projeto. A cada alteração realizada no contexto de um projeto, será gerado um Post com a informação descrevendo a alteração em questão, e identificando o autor da mesma. Os restantes contribuidores do projeto terão a possibilidade de fornecer feedback sobre cada Post, nomeadamente sob a forma de Gostos, Comentários e Partilhas.

Os Posts são apresentados numa linha temporal, permitindo aos intervenientes no projecto ter uma visão clara do decorrer das atividades de investigação. A possibilidade de comentar esses Posts ou fazer “Gosto” poderá ajudar também a determinar quais os projetos mais ativos, e ao mesmo tempo passar de certa forma pelas etapas da fase de publicação em cada alteração ao projeto, já que o uso dos “Gostos” e Comentários podem ser vistos como uma forma informal de revisão por pares. As Partilhas são também uma forma de disseminação de dados dentro do grupo de investigação. Também se pode assumir como uma ferramenta de auditoria, ajudando a manter registos de autoria de trabalhos e colaborações, complementando o papel do e-mail como ferramenta de documentação das interações dentro do grupo de investigação.

### **Funcionalidades principais da componente social**

As principais funcionalidades implementadas relativamente à componente social do Dendro assentam sobre princípios introduzidos pelas redes sociais, bem como ideias apresentadas no conceito Science 2.0 Repositories. Por outro lado, dado que o modelo de dados do Dendro é totalmente construído sobre ontologias, foi necessário modelá-lo com recurso a estas especificações. O reaproveitamento de conceitos e ontologias já definidas é ponto essencial para a sua correta modelação, pois só assim é possível assegurar a interoperabilidade entre os modelos de dados de diferentes sistemas que caracteriza a web semântica.

Na Figura 1 está representado o modelo de dados da extensão e as suas relações com conceitos do esquema schema.org, que resulta de uma colaboração entre as principais empresas de motor de busca (Google, Bing, Yandex e Yahoo!). Após a sua análise, foram reaproveitadas as classes CommentAction, LikeAction e ShareAction para a definição dos comentários, gostos e partilhas presentes no Social Dendro. O significado da classe SocialMediaPosting está alinhado com a nossa definição para os posts registados no Social Dendro e apresentados na timeline social.

Neste diagrama, a classe MetadataChange representa uma alteração aos metadados de uma pasta ou ficheiro dentro de um projeto. Uma MetadataChange possui um tipo (adição, edição ou remoção de uma instância de um descritor) e o valor anterior e o novo valor caso existam. Está também associada a um Descriptor: no projeto podem haver diversas modificações feitas ao descritor “Descrição” (Dublin Core) ao longo do tempo, por exemplo—cada uma será uma instância de MetadataChange, associada a esse Descriptor.

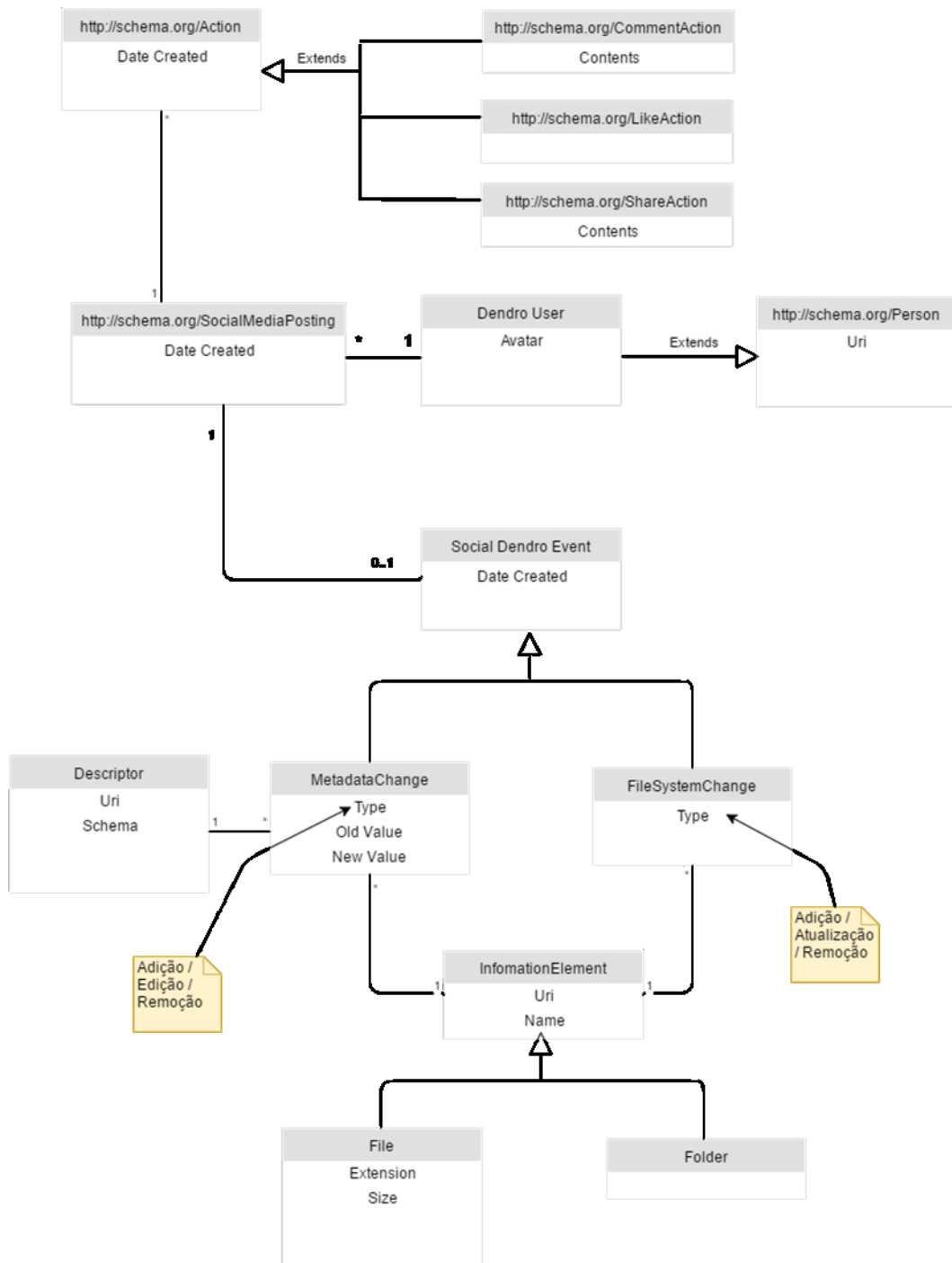


FIGURA 1 - MODELO DE DADOS

Outro exemplo de reaproveitamento de conceitos pré-existentes neste modelo de dados é a representação da hierarquia de pastas e ficheiros, que reutiliza classes definidas na Nepomuk File Ontology<sup>ii</sup>. Desta forma, os ficheiros e pastas de um projeto são representados pela classe InformationElement, as pastas pela classe Folder e os ficheiros pela classe File definidas nessa ontologia. Quando alguma alteração ocorre ao nível do sistema de ficheiros (criação, atualização ou remoção) é registada uma FileSystemChange e

gerado um SocialDendroEvent, que representa todos os eventos de manipulação quer de metadados quer de ficheiros. O sistema da timeline regista também a informação do que aconteceu nesse evento na forma de uma instância de SocialMediaPosting. Todos os utilizadores associados ao projeto da qual foi originada a alteração ao ficheiro ou pasta têm a possibilidade de realizar ações sociais sobre esse post, e o mesmo aparecerá na sua timeline.

### Interface com o utilizador

A extensão foi desenvolvida como uma timeline que apresenta, sobre a forma de posts, informação proveniente das alterações aos conjuntos de dados que foram sujeitos ao processo de curadoria. Como se mostra na Figura 2, o utilizador tem primeiro que aceder à sua área de projetos (1,2 na interface A), sendo-lhe apresentada a interface B. Seguidamente, poderá visualizar a timeline representativa das alterações aos seus conjuntos de dados (4) ou então aceder a uma listagem convencional (3).

A interface gráfica da extensão social foi desenvolvida de forma a ser o mais simples possível, a par com a orientação visual introduzida por aplicações de redes sociais bastante conhecidas. O objetivo principal foi o de tornar o processo de aprendizagem da extensão o mais rápido possível.

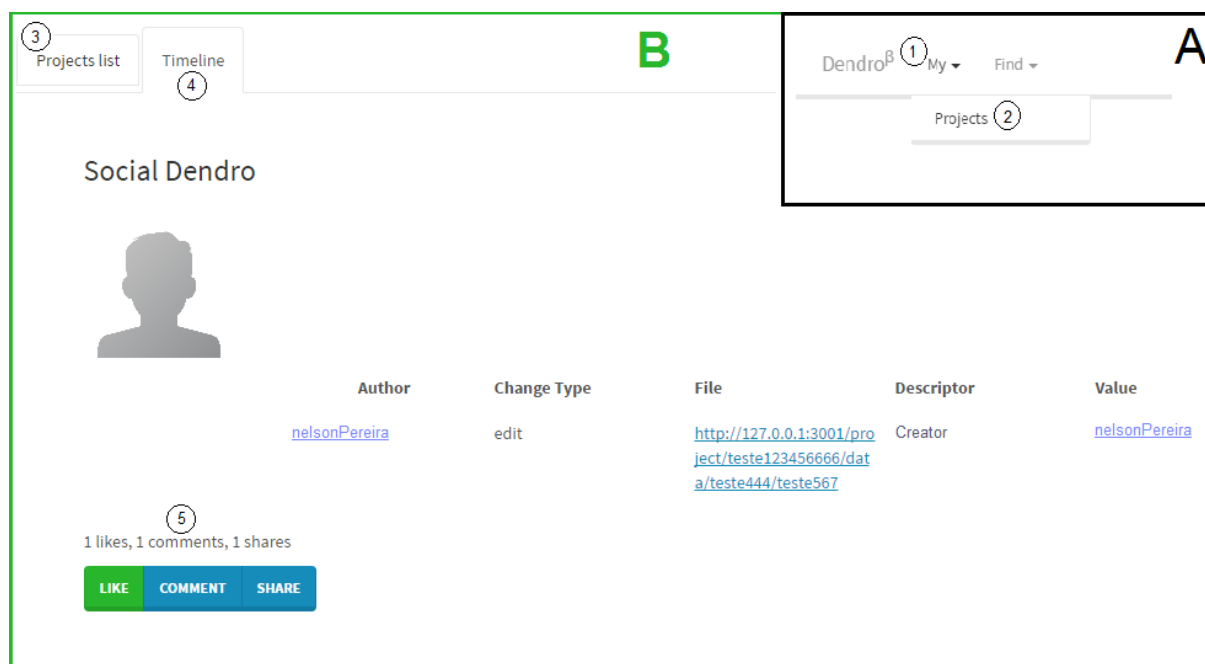


FIGURA 2 – ACESSO À EXTENSÃO SOCIAL DENDRO

Selecionando o separador “Timeline” (4) é possível observar todas as atualizações recentes relativas aos projetos de um utilizador, por ordem cronológica. Cada alteração gera um post respetivo, contendo informação sobre o autor da mesma, o tipo de alteração, bem como o ficheiro, descritor de metadados alterado e o valor do mesmo.



A Figura 3 representa as três possíveis interações a realizar perante um post numa timeline do Social Dendro. É possível interagir com cada post da timeline realizando um comentário (2), um gosto (1), ou uma partilha sobre o mesmo (3), fornecendo feedback ao autor perante as alterações introduzidas ao conjunto de dados de um projeto.

### Estado atual do desenvolvimento

Presentemente, na fase de protótipo, já são criados *posts* a partir das alterações aos conjuntos de dados, estando também implementada toda a componente da interação social com os mesmos (*gostos*, *comentários* e *partilhas*). O foco nas fases seguintes de implementação estará no processo de otimização bem como a realização de testes à usabilidade de forma a avaliar todo o processo de interação. O código fonte encontra-se disponível em regime open-source em [[Ligação github omitida para processo de revisão]].

### Conclusões e trabalho futuro

No estado atual de desenvolvimento, as principais funcionalidades do Social Dendro estão já implementadas. O resultado final encontra-se alinhado com os objetivos iniciais introduzidos no conceito de Science 2.0 Repository.

Serão efetuados testes de usabilidade com dois grupos de utilizadores: o primeiro grupo será constituído por utilizadores que nunca usaram o Dendro e um segundo com utilizadores experientes no seu uso. O objetivo será o de comparar a facilidade com que os dois grupos conseguem estar a par das alterações aos seus projetos, e a comunicação entre

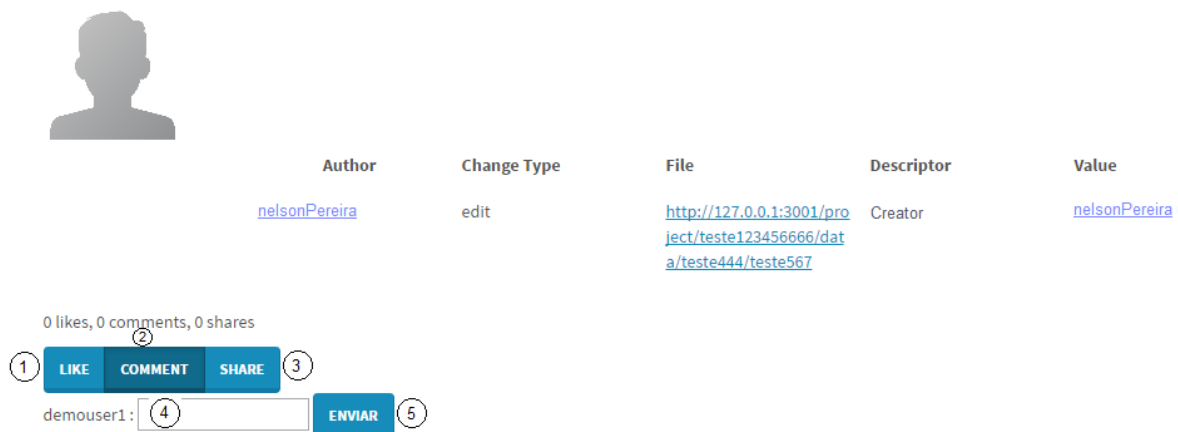


FIGURA 3 – OPÇÕES DE INTERAÇÃO

os diferentes elementos (investigadores) de um projeto, bem como a temporização das diferentes interações com a interface gráfica e observação e contagem do número de erros cometidos pelos utilizadores. Um inquérito com métricas bem estabelecidas de qualidade da experiência do utilizador (MACK e NIELSEN, 1993) será também levado a cabo.

## Referências bibliográficas

ASSANTE, Massimiliano [et al.] – Science 2.0 Repositories: Time for a Change in Scholarly Communication – D-Lib Magazine [Em linha]. Vol: 21, nº (2015),

BORGMAN, C.L. – The conundrum of sharing research data – Journal of the American Society for Information Science and Technology [Em linha]. Vol: 63, nº (2012),

COMMISSION, European – Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 – [Em linha]. (2013), p. 1–14.

FOUNDATION, National Science – Grants.Gov Application Guide A Guide for Preparation and Submission of NSF Applications via Grants.gov. 2011. ISBN: ISBN.

HODSON, Simon – ADMIRAL: A Data Management Infrastructure for Research Activities in the Life sciences.In [Em linha]. 2011,

MACK, Robert ; NIELSEN, Jakob – Usability inspection methods – ACM SIGCHI Bulletin [Em linha]. Vol: 25, nº (1993), p. 28–33. ISSN: 0897916514

PENG, Ge [et al.] – Scientific Stewardship in the Open Data and Big Data Era -- Roles and Responsibilities of Stewards and Other Major Product Stakeholders – D-Lib Magazine [Em linha]. Vol: 22, nº (2016),

PIWOWAR, Heather ; VISION, Todd – Data reuse and the open data citation advantage – PeerJ [Em linha]. Vol: 1, nº (2013), p. e175. ISSN: 2167–8359

POSCHEN, Meik [et al.] – Development of a Pilot Data Management Infrastructure for Biomedical Researchers at University of Manchester – Approach, Findings, Challenges and Outlook of the MaDAM Project – International Journal of Digital Curation [Em linha]. Vol: 7, nº (2012), p. 110–122.

STEINHART, Gail – DataStaR: A Data Sharing and Publication Infrastructure to Support Research – Agricultural Information Worldwide [Em linha]. (2011), p. 26–29.

TENOPIR, C. [et al.] – Data sharing by scientists: practices and perceptions – PLoS One [Em linha]. Vol: 6, nº 6 (2011), p. e21101. Disponível em WWW: <<https://www.ncbi.nlm.nih.gov/pubmed/21738610>>. ISSN: 1932–6203 (Electronic)

TENOPIR, C. [et al.] – Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide – PLoS One [Em linha]. Vol: 10, nº 8 (2015), p. e0134826. Disponível em WWW: <<https://www.ncbi.nlm.nih.gov/pubmed/26308551>>. ISSN: 1932–6203 (Electronic)

WALLIS, Jillian C. ; ROLANDO, Elizabeth ; BORGMAN, Christine L. – If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology – PLoS ONE [Em linha]. (2013), ISSN: 19326203 (Electronic)

---

<sup>i</sup> Disponível em código aberto em: [[link para repositório GitHub omitido para revisão]]

<sup>ii</sup> Mais informações em <http://www.semanticdesktop.org/ontologies/2007/03/22/nfo/>

---

# O cenário dos livros eletrônicos científicos em acesso aberto

---

Tainá Batista de Assis

Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

[taina@ibict.br](mailto:taina@ibict.br)

## Resumo

O Acesso Aberto à Informação Científica preconiza a disseminação dos artigos científicos em acesso aberto, livres de quaisquer custos de acesso e download. Entretanto, os artigos científicos não são o principal canal de comunicação para todas as áreas do conhecimento. Notadamente, as Ciências Sociais e Humanidades utilizam os livros como principal veículo para o registro e disseminação de seus conhecimentos e avaliação dos pesquisadores da área. Nesse sentido, o depósito, em acesso aberto, dos livros científicos em repositórios institucionais é baixo em comparação às outras produções científicas. Por outro lado, observa-se o crescimento de sistemas agregadores de conteúdo que reúnem os livros científicos e os disponibilizam para acesso dos usuários. Assim, esta pesquisa objetiva levantar e identificar as diferentes iniciativas para apontar a importância desses sistemas para possibilitar a maior visibilidade desses documentos.

**Palavras-chave:** Acesso Aberto, Livros científicos, Comunicação científica, Open Access Books

## THE SCENE OF ELECTRONIC BOOKS OPEN ACCESS SCIENTIFIC

### Abstract

Open Access to Scientific Information recommends the dissemination of scientific articles in open access, free of any access and download costs. However, the papers are not the main communication channel for all areas of knowledge. Notably, the Social Sciences and Humanities use the books as the main vehicle for the registration and dissemination of its knowledge and assessment of researchers. In this sense, the deposit in open access, scientific books in institutional repositories is low compared to other scientific productions. On the other hand, there is the growth of systems

aggregators of content that meet scientific books and make them available for user access. Thus, this research aims to raise and identify the various initiatives to highlight the importance of these systems for possibilitar increased visibility of these documents.

**Keywords:** Open Access, Scientific books, Scientific Communication, Open Access Books

## Introdução

O Movimento do Acesso Aberto à Informação Científica preconiza a disseminação dos artigos científicos, em seus textos completos, livres de quaisquer custos de acesso e download por meio da internet. No entanto, algumas áreas do conhecimento, notadamente as áreas das Ciências Sociais e Humanidades, se utilizam, maiormente, de livros como formas de registro e disseminação de seus resultados de pesquisas. Bizzocchi (2015) afirma que as Humanidades têm, em sua natureza, a necessidade de desenvolver discurso não direto, sucinto e objetivo, produzindo, assim, textos que necessitam de mais “linhas”.

Ribeiro (2005) defende que a tendência que se vê nas áreas das Humanidades é a publicação de livros a partir de uma coletânea de artigos, ou a publicação de livros caracterizados pelo autor por terem “começo, meio e fim”, sem repetições, com mediações e modificações nos textos. Ressalta Bizzocchi (2015) que a publicação em formato de livro na área das Ciências Sociais e Humanidades tem maior pontuação em avaliações institucionais, como é o caso nas universidades brasileiras.

Considerando o último aspecto apresentado, no Brasil, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), instituição responsável por fazer as avaliações dos Programas de Pós-Graduação brasileiros, aprovou na 111ª Reunião do CTC, realizada em 2009, critérios comuns para qualificação da produção intelectual veiculada por meio de livros em um modo de avaliação trienal. O documento aprovado nesta reunião destaca ainda a ausência de indexadores para a qualificação dos livros científicos acadêmicos assim como os adotados para periódicos, por exemplo, pelo Institute for Scientific Information (ISI).

A disponibilização de livros científicos em acesso aberto, contudo mostra-se tímida, ao se observar quantitativamente a presença desses recursos nos portais agregadores lusófonos de acesso aberto; o Repositório Científico de Acesso Aberto de Portugal (RCAAP) contém um total aproximado de 8.290 livros e o Portal Brasileiro de Publicações Científicas em Acesso Aberto (oasisbr) totaliza 1.945 livros em acesso aberto aproximadamente.

Assim, esta pesquisa objetiva levantar as iniciativas que têm sido observadas em todo mundo e que reforcem a necessidade de ações mais contundentes no âmbito do

Movimento do Acesso Aberto à Informação Científica para ampliação da disponibilização desse tipo de material.

### **Movimento do acesso aberto e os livros científicos**

É certo que o Movimento do Acesso Aberto à Informação Científica apresentou e continua provocando uma profunda alteração nos fluxos da comunicação científica e nos modelos de negócio adotados pelas editoras científicas. A chamada crise dos periódicos enfrentada pelas bibliotecas de pesquisa em todo o mundo se tornou o cenário propício para a estruturação de um novo fluxo de compartilhamento e divulgação dos resultados de pesquisa. Surgiram, assim, os primeiros sistemas de informação em acesso aberto que adotavam as tecnologias da filosofia da Iniciativa dos Arquivos Abertos (*Open Archives Initiatives* – OAI) e os primeiros periódicos científicos eletrônicos em acesso aberto. O modelo do Acesso Aberto, segundo Willinsky (2005) é “completamente digital e decididamente não-comercial.” A consulta aos textos completos das pesquisas científicas se dá livre de quaisquer custos de restrições de acesso e download, salvo aqueles documentos protegidos pelos direitos de autor.

Em 1991, o surgimento do primeiro repositório a adotar os preceitos da filosofia dos Arquivos Abertos chamado ArXiv impulsionou a criação de demais sistemas semelhantes em todo o mundo. O ArXiv é um sistema em acesso aberto das áreas da Física e Ciências correlatas que disponibiliza mais de 800 mil publicações. Nesse sentido, considera-se que, desde as primeiras iniciativas, o Movimento do Acesso Aberto preconiza a disponibilização livre e irrestrita aos artigos científicos para atender a demanda do campo de estudo das Ciências Exatas, uma vez que o principal veículo para a divulgação de sua pesquisa é o periódico científica. Entretanto, outras áreas do conhecimento, notadamente as Ciências Sociais e Humanidades, têm como principal canal de publicação e divulgação científica o livro.

Em um estudo de análise dos repositórios institucionais de 13 países em todo o mundo desenvolvido por Westrienen e Lynch, em 2005, identificou os principais tipos de documentos depositados nos sistemas. Enquanto na França e Itália, 80% e 70%, respectivamente, dos documentos correspondiam a artigos científicos, nos países Noruega e Suécia, 90% e 70%, respectivamente, apontavam para livros e teses de doutorado. No mesmo estudo, quando observado as áreas do conhecimento que cobriam os repositórios, a Itália, por exemplo, apresentava maior foco no campo das Ciências Sociais e Humanidades.

Embora o Movimento do Acesso Aberto preze por beneficiar igualmente as diferentes áreas do conhecimento, tem se desenvolvido mais rapidamente no campo das ciências exatas (Suber, 2014). Este autor levanta alguns aspectos relevantes nesse contexto: (i) artigos científicos tendem a serem os primeiros documentos científicos de leitura na área das Ciências Exatas e uma literatura secundária para as Sociais e Humanidades; (ii) nas

Ciências, os livros tendem a sintetizar as pesquisas publicadas em artigos, enquanto que nas Humanidades os artigos tendem a reportar uma história ou uma interpretação do livro (Eve, 2014). Já no ano de 1999, Meadows destacava para outro fator importante. De acordo com o autor, a velocidade de publicação na área das Ciências Sociais e Humanidades é, em geral, lenta.

Hellman (2011) apresenta outros possíveis argumentos que justificam a ainda baixa presença dos livros científicos nos repositórios. Destaca o custo para as taxas de custo para a publicação de um livro em comparação com um artigo científico. O autor aponta que, nos Estados Unidos, enquanto os custos para publicação de artigos científicos em periódicos de acesso aberto estão em torno de US\$ 600 a US\$ 3.000 dólares, as despesas gastas para a edição e publicação de um livro científico por uma editora universitária está estimado em cerca de US\$ 10.000 dólares. Ainda com relação aos custos centrados para a publicação de um livro nos Estados Unidos, Greco e Wharton (2008) apresentaram os gastos na produção por parte das editoras acadêmicas. Segundo dados do trabalho dos autores, em 2008, o faturamento anual das grandes editoras giravam em torno dos US\$ 50 milhões de dólares. Como exemplos, são citadas as editoras da Universidade de Oxford que, na época, apresentava um faturamento anual de aproximadamente US\$ 140 milhões de dólares e a editora da Universidade de Cambridge com cerca de US\$ 60 milhões de dólares.

Apesar de todos os obstáculos e desafios é consenso comum entre os autores que a disponibilização dos livros científicos em acesso aberto contribue positivamente e fortemente para o avanço da ciência em seus campos de estudo.

### **Livros científicos em acesso aberto: o cenário brasileiro**

Embora não tenha sido o foco desta pesquisa, a breve análise do cenário brasileiro acerca do assunto busca aqui também contribuir com dados para discussões futuras. Primeiramente, cabe destacar que, no contexto brasileiro, cada editora científica tem a sua política própria para os autores quanto à disponibilização de seus recursos em acesso aberto, embora a grande maioria faça parte da Associação Brasileira de Editoras Universitárias (ABEU). Atualmente, estão associadas à ABEU 119 editoras universitárias de todos os estados brasileiros.

No contexto governamental, a discussão sobre os critérios de qualificação da produção intelectual dos pesquisadores brasileiros, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), instituição responsável por fazer as avaliações dos Programas de Pós-Graduação brasileiros, aprovou durante a 111ª Reunião do CTC, realizada em 2009, critérios comuns para qualificação da produção intelectual veiculada por meio de livros em um modo de avaliação trienal. O documento aprovado destaca que, assim em outras áreas do conhecimento, os livros constituem em “modalidade de veiculação da produção artística, tecnológica e científica.” O documento reforça também a dificuldade

quanto à avaliação dessa tipologia científica, pois também não há, no contexto internacional, exemplos de iniciativas para esse fim.

Nesse sentido, a CAPES apresentou o instrumento para a avaliação dos livros científicos brasileiros. Os estratos definidos para a classificação dos livros entre os níveis L1 a L4, sendo L1 a menor classificação e a L4 corresponde a estratos superiores. A saber, são os critérios para a avaliação dos livros:

Parte I: Dados de identificação da obra – basicamente são os dados que constam na ficha catalográfica;

Parte II: Avaliação pela Comissão de classificação de livros – compreendem a análise dos dados mínimos (como ISBN, número de páginas...); dos aspectos formais (autoria, premiação, financiamento...); quanto ao tipo e natureza do texto;

Parte III: Avaliação do conteúdo da obra – compreendem os aspectos relacionados com a relevância temática, inovação e potencialidade do impacto.

Assim sendo, a decisão apresentada pela CAPES busca orientar e consolidar as discussões acerca do tema para uma avaliação individual mais coerente para a comunidade científica brasileira.

Considerando a disseminação dos livros científicos brasileiros em acesso aberto, em uma breve análise no Portal Brasileiro de Publicações Científicas em Acesso Aberto (oasisbr)<sup>i</sup> observa-se que os livros também apresentam baixa presença. Os livros correspondem a aproximadamente 0,40% do total de documentos presentes no sistema.

Tendo em vista a quantidade de livros que são produzidos pelas editoras universitárias associadas à ABEU, nem todo o material produzido está presente nos repositórios institucionais de suas instituições. Tal fato justifica o número reduzido desses documentos no oasisbr e aponta para a necessidade de uma maior proximidade entre os responsáveis pelas editoras e os gestores dos repositórios institucionais.

## **Metodologia**

Considerando o objetivo proposto para o desenvolvimento desta pesquisa foram identificadas oito grandes iniciativas em todo o mundo que reúnem e disseminam, em acesso aberto, os livros científicos eletrônicos.

O levantamento para a identificação desses sistemas foi realizado de forma livre no motor de busca *Google*, sendo assim selecionados, àqueles que apareceram na primeira apresentação. A sequência de apresentação dos sistemas nos resultados deste estudo não segue uma ordem de importância e/ou representatividade, estando assim dispostas de forma aleatória ou por proximidade de natureza.



## Resultados

Embora os livros científicos não figurem fortemente nos sistemas de informação em acesso aberto, importantes iniciativas foram identificados por meio do uso do motor de busca *Google*. Este fato aponta para uma perspectiva futura positiva nessa discussão e demonstra a preocupação das instituições em reunirem, em um só sistema de consulta, os livros científicos produzidos sobre os mais variados temas.

Para começar a apresentação dos resultados, o próprio motor de busca *Google* apresenta aos usuários a opção para recuperarem os textos completos de livros, serviço este chamado de Google Books. O projeto teve início em 2004 e disponibiliza as obras no formato Portable Document Format (PDF) que estão em domínio público e/ou que não tenham direitos autorais declarados. O Google realiza o processo de escaneamento dos livros em conjunto com seus editores sócios e estima-se que esse processo custou, no ano de 2007, mais de US\$ 5 milhões de dólares.

O Directory of Open Access Books (DOAB) reúne mais de 5.400 livros acadêmicos revisados por pares de 161 editores. O DOAB é um serviço da Fundação OAPEN dedicada às publicações de monografias com base na National Library in The Hague e com parcerias com OpenEdition, Brill e Springer. Utiliza ferramentas que viabilizam a coleta automática de metadados como o Protocol for Metadata Harvesting (OAI-PMH). Também deixa claramente visível as informações para a coleta de seus bancos por qualquer outra iniciativa.

O Scientific Electronic Library Online (SciELO) disponibiliza a Rede SciELO Livros com o acervo de editoras brasileiras e que segue critérios, políticas e procedimentos para a avaliação e seleção de editoras, coleções e livros que participam do projeto. O Portal apresenta 745 títulos disponíveis, tendo 430 títulos em acesso aberto. Observa-se também que a iniciativa da Rede SciELO Livros interopera e compartilha os objetivos e recursos também com a Rede SciELO Periódicos. O projeto é financiado por um consórcio formado pelas editoras da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), Universidade Federal da Bahia (UFBA) e da Fundação Oswaldo Cruz (FIOCRUZ).

No contexto das editoras comerciais, a própria Springer tem o projeto SpringerOpen Books que oferece a opção aos seus autores de publicação de seus livros em acesso aberto dentro das áreas da ciência, tecnologia e medicina, sob a cobrança de uma taxa de publicação.

A Taylor & Francis, por meio do projeto Routledge Books Open Access, possibilita a publicação de livros em acesso aberto dentro dos assuntos cobertos pela Routledge e outros de Humanidades, Ciências Sociais e Ciências do Comportamento.

Entre as instituições de ensino apresentam-se as experiências de Cambridge, Harvard e Oxford. A primeira apresenta o acesso aberto aos livros por meio do “Programa de Publicação Via Dourada e a Política da Via Verde”; a segunda, tem o projeto de acesso aberto

implementado na instituição desde 2011 e; por último, a Oxford participa da iniciativa OAPEN-UK que dá viabilidade às monografias nas áreas das Ciências Sociais e Humanas.

Desta última iniciativa, destaca-se o trabalho desenvolvido para a promoção das discussões relacionadas com os direitos autorais e a disponibilização de guia para orientar a criação de uma publicação de livros em acesso aberto por pesquisadores dos campos de estudos das Artes, Humanidades e Ciências Sociais.

### **Considerações finais**

Como considerações finais destaca-se a necessidade de se pensar em estratégias junto às editoras, comerciais e institucionais, nos países lusófonos para maior disponibilização em acesso aberto dos livros científicos, sobretudo aqueles produzidos com recursos públicos. Tal ação implicará na análise e possível revisão das políticas impostas pelas editoras para a disponibilização desses recursos em acesso aberto.

Certamente, a maior disseminação dos livros científicos em acesso aberto contribuirá para o avanço da ciência na geração de novos conhecimentos, indicadores e maior transparência da ciência produzida nas diversas áreas do conhecimento.

### **Referências bibliográficas**

- BIZZOCCHI, A. Sobre livros acadêmicos [Em linha]. 2015. Disponível em WWW: <<http://revistalingua.com.br/textos/blog-abizzocchi/sobre-livros-academicos-274896-1.asp>>.
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Roteiro para classificação de livros: avaliação dos programas de Pós-Graduação [Em linha]. 2009. Disponível em WWW: <[https://www.capes.gov.br/images/stories/download/avaliacao/Roteiro\\_livros\\_Trienio2007\\_2009.pdf](https://www.capes.gov.br/images/stories/download/avaliacao/Roteiro_livros_Trienio2007_2009.pdf)>.
- GRECO, A. N.; Wharton, R. M. Should University Presses Adopt An Open Access [Electronic Publishing] Business Model For All of Their Scholarly Books? [Em linha]. 2008. Proceedings ELPUB 2008 Conference on Electronic Publishing – Toronto, Canada – June. Disponível em WWW: <<https://pdfs.semanticscholar.org/147a/a9c4d574cb01d57b105f80daef5fa6a390e.pdf>>.
- HELLMAN, E. Open Access eBooks, Part 1 [Em linha]. 2011. Disponível em WWW: <<https://go-to-hellman.blogspot.com.br/2011/04/open-access-ebooks-part-1.html>>.
- MEADOWS, A. J. A comunicação científica. Brasília: Briquet de Lemos, 1999.
- RIBEIRO, R. J. A questão do livro de pesquisa em Humanidades [Em linha]. 2005. Disponível em WWW:

<[http://www.sbpcnet.org.br/livro/57ra/programas/CONF\\_SIMP/textos/renatojanineribeiro.htm](http://www.sbpcnet.org.br/livro/57ra/programas/CONF_SIMP/textos/renatojanineribeiro.htm)>.

SUBER, P. Preface [Em linha]. 2014. EVE, Martin Paul. Open Access and the Humanities: contexts, controversies and the future. Disponível em WWW: <[goo.gl/eg6DMr](http://goo.gl/eg6DMr)>.

WESTRIENEN, G. van; LYNCH, C. A. Academic institutional repositories: deployment status in 13 nations as of Mid 2005 [Em linha]. 2005. Disponível em WWW: <<http://www.dlib.org/dlib/september05/westrienen/09westrienen.html>>

WILLINSKY, J. Scholarly Associations and the Economic Viability of Open Access Publishing [Em linha]. 2005. Open Journal System Demonstration Journal, 1(1). Disponível em WWW: <[journals.sfu.ca/present/index.php/demjournal/article/view/6/8](http://journals.sfu.ca/present/index.php/demjournal/article/view/6/8)>.

---

<sup>i</sup> Portal oasisbr <[oasisbr.ibict.br](http://oasisbr.ibict.br)>

---

## Integração de plataformas institucionais: o caso EUDAT

---

Ricardo Carvalho Amorim

INESC-TEC/ Faculdade de Engenharia da Universidade do Porto

[ricardo.amorim3@gmail.com](mailto:ricardo.amorim3@gmail.com)

Fábio Silva

Faculdade de Engenharia da Universidade do Porto

[ffjs1993@gmail.com](mailto:ffjs1993@gmail.com)

### Resumo

A qualidade dos dados de investigação é um debate recorrente entre instituições e investigadores. A constante preocupação em garantir que os dados sobrevivem aos seus criadores tem contribuído para o aparecimento de plataformas que, em casos mais avançados, permitem que os investigadores sejam envolvidos diretamente no processo de gestão dos seus dados. A plataforma EUDAT surge como um exemplo de uma uma solução para os inúmeros desafios da investigação, proporcionando ambientes colaborativos, ferramentas de processamento e outras mais valias para instituições que têm dificuldades em adquirir estas ferramentas de forma individual. Contudo, a divulgação de dados de investigação deve ter como princípios não só a sua preservação a longo prazo, mas também a sua potencial reutilização dentro da comunidade científica, pelo que é necessário transpor o conhecimento gerado durante a investigação para os dados associados. O Dendro surge como uma plataforma para apoiar a descrição de dados desde o momento em que são criados, permitindo que os registos produzidos os acompanhem e permitam mais tarde uma interpretação extensiva dos dados. Este caso de estudo apresenta a integração entre o Dendro e o EUDAT com o objetivo de promover um ambiente de gestão de dados abrangente que envolva os diversos responsáveis no processo.

**Palavras-chave:** Dendro, EUDAT, Gestão de dados de investigação, Descrição de dados, Preservação digital

## Integration of institutional platforms: the EUDAT case study

### Abstract

Research data quality is a common discussion among both institutions and researchers. To ensure that data outlive their producers, several platforms have recently emerged that, in some scenarios, allow researchers themselves to contribute to their data management. The EUDAT platform is an example of a solution for several challenges in the research environment, providing researchers with collaborative environments, large-scale data processing tools, and other features to promote both their data quality and long-term preservation. Data publication has also to consider aspects beyond long-term preservation, such as ensuring their potential reuse within the research community and thus, it is necessary to create adequate tools to capture knowledge as soon as it is produced. Dendro is a platform for data management from early stages, with a focus on data description, from the moment researchers start their projects. This case study reports on the integration between EUDAT and Dendro, to create an environment for data management involving all the stakeholders in the process.

**Keywords:** Dendro, EUDAT, Research Data Management, Data description, Data preservation

### Introdução

A gestão de dados de investigação apresenta vários desafios para todos os envolvidos no processo. Apesar de grandes instituições já disporem de ferramentas e plataformas para a gestão dos seus dados nas várias fases, o mesmo não acontece com pequenos grupos de trabalho que por vezes não têm suporte para tal e são encarregues de gerir os dados que produzem no decorrer de vários projetos (Comp, 2016).

A preocupação com a preservação e divulgação de dados de investigação de grupos de pequena ou média dimensão tem vindo a promover iniciativas no sentido de criar plataformas para esse efeito, de entre as quais se destaca o EUDAT. Como repositório de dados, o EUDAT procura oferecer um conjunto de serviços para investigadores de vários domínios cujas instituições possam, por exemplo, ter dificuldades em criar e manter infraestruturas para o efeito (Burns, 2013).

Com o objetivo de publicar dados de investigação e, ao mesmo tempo, criar condições para que estes possam ser reutilizados, a descrição de dados apresenta alguns desafios, ao considerar os inúmeros domínios da investigação. A produção de descrições adequadas para os dados de investigação é por vezes uma tarefa que os investigadores já

incluem no seu dia a dia contudo, a falta de ferramentas adequadas faz com que estes recorram frequentemente a suportes que se deterioram rapidamente, e não dão suporte a processos estruturados de descrição. A plataforma Dendro foi desenhada com o princípio de envolver investigadores diretamente na gestão dos seus dados, desde o momento em que começam a produzir dados. A sua arquitetura permite também que equipamentos ou aplicações externas se possam ligar para cobrir ambientes específicos.

Este caso de estudo apresenta a ligação entre o Dendro e o EUDAT, como forma de estabelecer um ambiente abrangente para a gestão de dados. Este tipo de iniciativas permite alargar o leque de possibilidades para instituições que procurem dar resposta às crescentes necessidades de gestão de dados dos seus projetos e de torna-los inteligíveis para a comunidade científica.

### **Gestão de dados na cauda longa da investigação**

A cauda longa da investigação compreende instituições que lidam frequentemente com vários domínios da investigação, em grupos de relativamente pequena dimensão. O suporte tradicionalmente existente para a gestão de dados em grandes instituições, não é aplicável neste cenário onde cada grupo tem requisitos diferentes. Os requisitos existentes para uma adequada gestão de dados podem envolver, entre outros aspetos, a criação de estruturas capazes de lidar vários formatos de representação de informação, com diferenças organizacionais entre domínios, com diferentes metodologias de produção de dados e, ao mesmo tempo, serem capazes de envolver os produtores de dados neste processo (Massimiliano, 2016). Por apresentar inúmeros desafios, a gestão de dados de investigação neste domínio tem vindo a receber bastante investimento com vista a promover a preservação digital e a reutilização de dados que, de outra forma, perderiam o seu valor. Existem, contudo, ainda alguns aspetos que têm vindo a merecer um maior destaque por influenciarem diretamente a forma como estes objetivos são alcançados. A gestão de dados é, por vezes, uma tarefa deixada para as fases finais de cada projeto onde um especialista-frequentemente o responsável pela gestão do projeto-tem a seu cargo a tarefa de organizar, descrever e depositar os recursos gerados numa plataforma destinada para o efeito. Ao concentrar este conjunto de responsabilidades e tarefas num único ponto ou numa única pessoa, o tratamento e depósito dos dados podem não chegar a acontecer, perdendo-se assim o seu valor e mantendo-se apenas a publicação deles resultante (McLure, 2014).

A gestão de publicações é uma área já muito desenvolvida e conta já com inúmeras plataformas que a suportam e garantem a sua continuidade. Para além disso, está já bem documentada, fazendo parte do conhecimento dos investigadores tarefas como a indexação de publicações, a sua divulgação, a sua citação e conseqüente crédito aos autores. O mesmo ponto de vista pode ser aplicado aos dados: um conjunto de dados que possa ser consultado pela comunidade científica pode também ser citado e igualmente dar crédito aos seus autores e instituições envolvidas na sua produção. O desafio é, contudo, distinto uma vez

que a interpretação de dados requer um conhecimento mais profundo sobre a sua origem e contexto, que desta forma não chegam a ser associados ao seu depósito em repositórios. Por outro lado, instituições que se candidatam para financiamentos comunitários têm vindo a deparar-se com recomendações para a criação de um plano para a gestão dos dados resultantes do projeto e, em alguns casos, a sua existência é já um fator de exclusão. Estes planos incluem muitas vezes a especificação de metodologias para o tratamento e descrição de dados, assim como a identificação de um repositório no qual serão depositados estes recursos.

Pela sua natureza, os dados de investigação podem beneficiar de uma descrição com grande pormenor, quando comparados com publicações. Por permitir incluir descrições características do domínio em questão, a descrição de dados de investigação tem sido considerada uma etapa importante no seu processo de gestão. Para um projeto no domínio da biodiversidade, por exemplo, esta descrição pode indicar referências geográficas para os pontos onde foram recolhidas amostras, as espécies identificadas ou mapas com os percursos realizados em cada saída de campo.

Várias plataformas têm surgido para permitir que os depósitos de dados possam incluir igualmente descrições em vários níveis, para facilitar a interpretação por parte da comunidade. Dado o requisito de receber dados de vários domínios, estas plataformas tendem a oferecer funcionalidades que são de interesse para uma grande quantidade de equipas de investigação, delegando funcionalidades específicas para ferramentas individuais que se podem integrar com o repositório em qualquer altura.

## **A plataforma EUDAT**

Criada no âmbito da promoção da gestão de dados de investigação para instituições europeias, o EUDAT apresenta vários módulos que incluem ferramentas colaborativas, gestão de equipas, processamento em larga escala e plataformas de divulgação de dados.

### **Posicionamento da plataforma**

As plataformas para gestão de dados podem ser agrupadas consoante o seu princípio de funcionamento: plataformas que estão sobre o controlo de instituições, sendo instaláveis em servidores institucionais; e plataformas que fornecem um serviço através do qual as instituições podem contratar diferentes tipos de acesso. Ambas apresentam as suas vantagens quer nos custos de implementação, quer no controlo sobre os dados (Amorim, 2015). Por incluir um conjunto de serviços adicionais juntamente com a componente de repositório, a plataforma EUDAT insere-se numa categoria distinta da dos repositórios para dados de investigação. O EUDAT permite que os dados comecem a ser criados diretamente na plataforma, processados e posteriormente divulgados, juntamente com informações que os autores pretendam anexar.

## Módulos de gestão e produção de dados

Dentro da plataforma, os diferentes módulos estão interligados e permitem troca de recursos entre si. Esta abordagem surge também como uma solução para acompanhar os processos desde o arranque de cada projeto, permitindo que os investigadores interajam com a plataforma assim que começam a produzir os seus dados.

Para fases iniciais, onde os dados são frequentemente atualizados ou alterados, o módulo B2Drop permite que os investigadores acedam a uma área colaborativa para gerir diariamente os seus resultados. Este módulo mantém ao mesmo tempo versões dos dados, permitindo assim acompanhar as alterações feitas ao longo do tempo.

No final de cada projeto, e caso não existam períodos de embargo a respeitar, os dados envolvidos podem ser divulgados para a comunidade. Nesta fase, o B2Share assume o papel de receber estes recursos—que podem ter origem em outras plataformas para além do B2Drop—e proporcionar meios para os tornar inteligíveis e acessíveis ao exterior. Neste processo, o EUDAT dispõe de interfaces para apoiar o investigador no depósito, assegurando que aspetos como períodos de embargo, identificadores únicos e políticas de acesso são especificados antes de tornar o conjunto de dados público. Nesta fase, o investigador pode também adicionar uma descrição aos seus dados, para facilitar a sua interpretação, contudo, esta descrição tem um carácter facultativo e tende a assumir uma representação não normalizada.

Uma vez que todos os módulos podem comunicar entre si para abranger os vários requisitos envolvidos, o B2Share é também responsável por garantir que os recursos depositados são preservados indefinidamente. Nesta fase o módulo B2Safe encarrega-se de implementar políticas de curadoria e preservação digital incluindo a gestão de identificadores, de cópias de segurança e a verificação frequente da integridade dos dados. Tal como acontece com os outros módulos, o B2Safe é capaz de receber recursos de outras plataformas, reconhecendo automaticamente pacotes de repositórios conhecidos como o Dspace e o FEDORA. Esta abordagem permite responder a plataformas que não tenham capacidade para proporcionar um ambiente de preservação com grandes capacidades de armazenamento.

No lado do investigador externo que procura um determinado conjunto de dados, o B2Find assegura que os conjuntos geridos pela plataforma se encontram devidamente registados e indexados. A sua interface para o exterior foi pensada para permitir pesquisas de dados independentes do seu domínio, origem ou comunidade associada. Este módulo assume particular importância por ser responsável por levar os recursos armazenados e tratados até aos potenciais re-utilizadores, fazendo também a ligação com publicações que possam estar associadas a cada conjunto de dados.



Os módulos propostos na plataforma EUDAT podem assumir um papel importante na gestão de dados em equipas de investigação que não disponham atualmente de meios para o fazer. As fases iniciais de cada projeto estão também abrangidas, permitindo o depósito em áreas privadas e a gestão colaborativa de recursos que, de outra forma, ficariam dependentes dos modelos organizacionais da equipa de investigação. Esta abordagem está em linha com as recentes preocupações no sentido de envolver diretamente os investigadores na gestão dos seus dados até à sua divulgação. Esta é considerada uma prioridade, dado o conhecimento sobre os dados estar naturalmente disponível nesta fase e poder ser capturado sem requerer um grande esforço, existindo ferramentas adequadas para tal.

### **Metadados uma melhoria na qualidade dos dados**

A descrição de dados é uma tarefa que, em alguns casos, já faz parte do quotidiano dos investigadores. O contexto em que os dados são produzidos é tradicionalmente registado em suportes como cadernos de laboratório ou blocos de notas pessoais e inclui geralmente metadados que o investigador considera relevantes para a sua compreensão. Apesar de bastante utilizados, estes suportes podem, a longo prazo, comprometer os registos recolhidos uma vez que tendem a deteriorar-se e a sua preservação está dependente do proprietário. A descrição de dados assume assim um papel importante para a sua compreensão e, como produtores de dados, os investigadores são os melhores candidatos para executar esta tarefa.

A inclusão de metadados no depósito em repositórios pode contribuir para melhorar a visibilidade de um determinado conjunto de dados. Ao permitir que as plataformas existentes indexem o seu conteúdo, por exemplo, as pesquisas podem tornar-se mais eficientes e incluir modelos avançados de pesquisa para melhorar a sua precisão. Um investigador no domínio da Química pode, por exemplo, pesquisar por conjuntos de dados que envolvam um determinado reagente, se este estiver devidamente identificado na folha de metadados. Por outro lado, um outro investigador no domínio da biodiversidade pode restringir os resultados da sua pesquisa a uma determinada cobertura geográfica sem que, para isso seja necessário abrir cada conjunto de dados apresentado. Os esforços no sentido de implementar mecanismos de recolha e produção de metadados pelos investigadores têm feito surgir ferramentas e aplicações desenhadas para serem integradas no seu ambiente de trabalho ou substituir suportes que não permitam a flexibilidade desejada para este efeito (Amorim, 2014). Aplicações neste domínio começam também a valorizar a recolha estruturada de metadados, fazendo com que estes possam ser interpretados, mesmo depois de uma eventual descontinuação da aplicação. Apesar desta vantagem, a descrição estruturada de dados é ainda um tema em desenvolvimento, não fazendo parte de muitos dos repositórios existentes.

## Descrição de dados na plataforma Dendro

A plataforma Dendro (Silva, 2014) foi pensada para permitir que a gestão de dados, colaborativa ou não, possa começar com o arranque de cada projeto. Para o investigador, a interface assemelha-se à que já é apresentada em plataformas de armazenamento remoto como Dropbox ou Google Drive. Por outro lado, apresenta também uma interface de preenchimento de campos cujo objetivo é guiar o investigador na descrição dos seus dados, à semelhança do que já é feito em cadernos de laboratório ou blocos de notas. Cada elemento que se adiciona à folha de descrição-descriptor-define um determinado contexto e está associado a um determinado ficheiro ou pasta na plataforma. Apesar de a seleção de descritores depender do investigador, a plataforma é capaz de gerar um perfil de descrição para cada investigador ou equipa de investigação com base em dados recolhidos ao longo das diferentes interações. O resultado consiste na adaptabilidade que a plataforma é capaz de alcançar, independentemente dos domínios que estão envolvidos na sua utilização. O Dendro é capaz de, por exemplo, destacar descritores do domínio da psicologia para grupos de investigação dessa área, sem descurar no entanto, a possibilidade de cada investigador poder individualmente adaptar o perfil às suas necessidades.

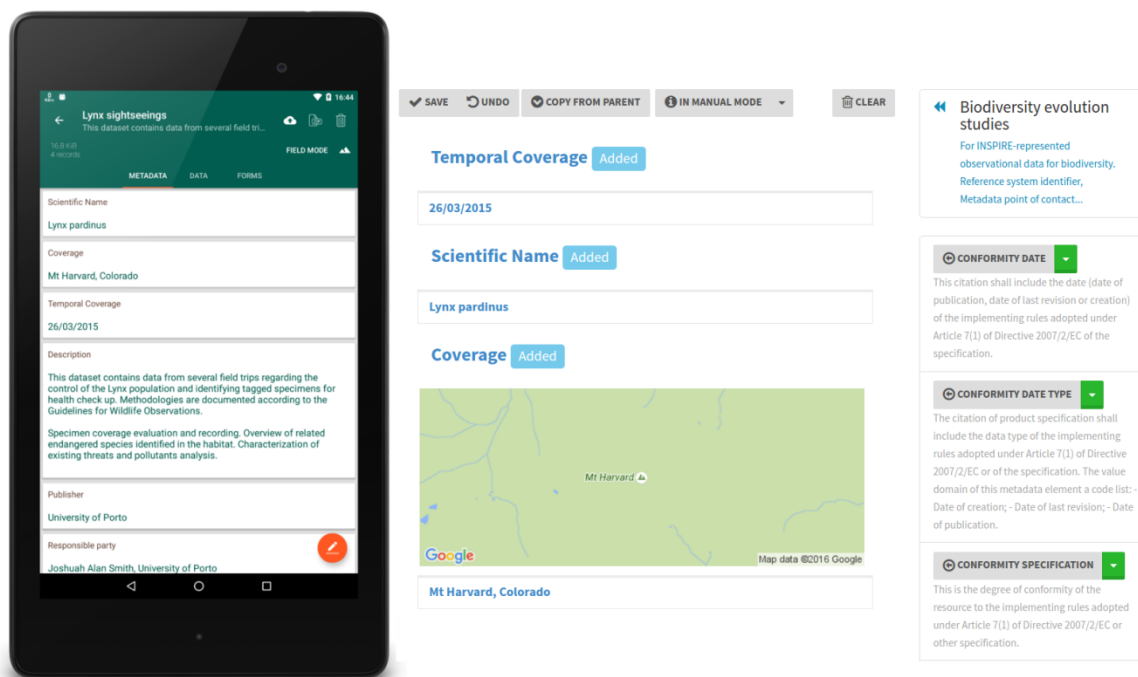


FIGURA 1 VISTA DE DESCRIÇÃO DE UM PROJETO, E INTERFACE EQUIVALENTE NA APLICAÇÃO LABTABLET

Ao construir um perfil adequado para cada domínio, a tarefa de descrição pode tornar-se menos um requisito, e mais uma tarefa guiada, personalizada e por vezes automatizada. Dentro de cada pasta, os investigadores podem copiar a folha de descrição completa e replicá-la para outro local do projeto, alterando apenas as descrições que sejam necessárias. A gestão colaborativa permite também que, se um investigador determinar que para os seus projetos determinados descritores devem ser destacados, estes possam ser

sugeridos aos seus colaboradores. As descrições feitas seguem também um determinado modelo de metadados, promovendo assim a sua interpretação por outros sistemas. Especialistas de informação contribuem aqui para a seleção de modelos de metadados pertencentes a normas existentes no domínio. No domínio da biologia marinha, por exemplo, o perfil da norma INSPIRE já especifica detalhadamente vários níveis de descrição para ajudar os investigadores a descreverem os seus dados. Por outro lado, no mesmo domínio, mas com maior abrangência, o perfil Darwin Core inclui alguns descritores que cobrem conceitos como georeferenciação, identificação de espécies e outros elementos relacionados. Na sua utilização, os investigadores podem selecionar livremente descritores de vários perfis, se considerarem que a sua inclusão é relevante para apoiar a interpretação dos seus dados.

O Dendro pertence a uma categoria de repositórios com um maior foco em capturar dados e os respetivos metadados assim que estes são produzidos. Consiste assim numa plataforma intermédia, que se integra no ambiente de trabalho dos investigadores, podendo ser complementado por ferramentas que produzam dados para estes serem automaticamente depositados. Nesta categoria podem ser considerados, por exemplo, cadernos de laboratório eletrónicos como o LabTablet e outros equipamentos laboratoriais. Dada a sua natureza, o Dendro pode em qualquer altura ser ligado a outras plataformas para exportar os recursos no final de cada projeto. A Figura 1 apresenta o mesmo projeto em duas interfaces distintas: na plataforma Dendro (lado direito) e na aplicação LabTablet (lado esquerdo). Ambas as plataformas têm como objetivo auxiliar na descrição de dados, independentemente do domínio em questão e neste caso concreto o seu funcionamento em conjunto pretende dar resposta a ambientes que exijam, por exemplo, uma maior mobilidade. Os investigadores podem tirar partido das funcionalidades que um dispositivo móvel tem, em particular dos sensores que caracterizam estes dispositivos—georeferenciação, temperatura, entre outros—, e sincronizar os conteúdos com o Dendro em qualquer altura. Depois de encerrarem cada projeto, os investigadores podem passar para a fase em que depositam os seus dados numa plataforma escolhida para o efeito. Idealmente, nesta fase, o investigador ou a sua equipa devem apenas escolher o repositório destino para os seus dados e proceder com a exportação.

### **Interligação Dendro–EUDAT**

A plataforma Dendro foi desenhada com base no princípio de ser uma plataforma para gestão intermédia de dados de investigação. A partir deste ponto, o objetivo é fazer com que dados e metadados possam ser exportados em conjunto para um repositório que tenha em consideração os aspetos relacionados com a preservação digital e a disseminação de dados. Para este efeito, e dadas as capacidades já implementadas da plataforma EUDAT, foi estabelecida a ligação das duas plataformas para que, no final de cada projeto, seja

possível exportar os pacotes de dados e metadados produzidos pelos projetos de investigação e assegurar a sua disseminação e preservação.

The screenshot shows a web interface for a data repository. On the left, there's a 'Folder' section with icons for upload, download, refresh, and share. Below it is a 'Selection' section with 'COPY FROM PARENT' and 'IN MANUAL MODE' buttons. A file list on the left includes 'Up to carrental', 'definitions.pdf', and 'instances.xlsx'. The main content area displays a record for 'Vehicle reservation assignment in car rental' by Sebastião Oliveira at the University of Porto. It includes an abstract, a DOI link, and a table of files. The table has columns for Name, Date, Size, and Download. Below the table are 'Export' and 'Metadata' sections.

Name	Date	Size	Download
files.zip	10 Jul 2016	2.1 MB	Download
metadata.rdf	10 Jul 2016	1.9 kB	Download
metadata.txt	10 Jul 2016	1.7 kB	Download

FIGURA 2 EXEMPLO DE EXPORTAÇÃO DE UM CONJUNTO DE DADOS PARA O REPOSITÓRIO EUDAT

A implementação recorreu à arquitetura com base em extensões, adicionando a opção para o repositório EUDAT no formulário de exportação de datasets. Neste ponto, os investigadores podem escolher vários repositórios—Dspace, CKAN e Figshare são alternativas também disponíveis—sendo reencaminhados para uma interface específica de cada um para acompanhar a submissão. Nesta fase são pedidas credenciais de acesso para depósito, sendo que estas podem ser automaticamente configuradas se os depósitos forem feitos pela instituição e não por investigadores de forma individual. A exportação é feita através do módulo B2Share, que mais tarde se encarrega da sua disseminação para os restantes módulos. Após o depósito nesta plataforma, é enviada uma notificação com o identificador resultante, e o endereço onde se pode aceder ao depósito feito. A partir deste momento, os dados podem ser consultados pela comunidade, e citados de igual forma. A Figura 2 apresenta o resultado de um depósito na plataforma. Para a comunidade científica passa a ser possível aceder aos dados depositados, assim como à sua descrição produzida pelos próprios investigadores do seu domínio. À semelhança do que é feito com publicações, estes conjuntos de dados também podem ser citados, dispondo de um identificador único criado para o efeito.

### Divulgação antecipada de dados de investigação

A publicação de dados é ainda um tema em debate quer para instituições, quer para investigadores. Em alguns casos, a divulgação de dados não pode ser feita sem que terminem restrições de acesso como períodos de embargo. Para estes casos, algumas plataformas já permitem manter os dados privados até que esse período termine, respeitando assim quaisquer imposições legais que possam existir e, ao mesmo tempo, garantindo que os dados já se encontram depositados e prontos a serem divulgados. Por outro lado, é também frequente encontrar projetos em que a divulgação dos dados pode

ocorrer desde as fases iniciais para promover aspetos como a visibilidade do projeto ou da instituição para o exterior. A integração entre o Dendro e o EUDAT teve em consideração estas restrições, e cobre assim três dos possíveis cenários quando se aborda a publicação de dados de investigação:

- **Publicação no final do projeto:** o investigador responsável ou a própria instituição encarregam-se de proceder à exportação dos dados do Dendro para o EUDAT, assim que o projeto termina;
- **Publicação imediata:** permite criar projetos na plataforma Dendro e assegurar que estes estão acessíveis desde o início. Investigadores externos que queiram consultar os dados podem fazê-lo diretamente na plataforma;
- **Publicação de metadados:** os metadados sobre o projeto são disseminados para a comunidade científica. Investigadores interessados em aceder aos dados podem solicitar autorização para o fazer ao responsável pelo projeto após se registarem na plataforma.

Os três cenários visam responder a necessidades particulares que podem existir em vários domínios, assegurando que os investigadores e instituições envolvidas têm o controlo sobre os seus dados e recebem crédito pela sua disponibilização.

Para os casos em que se pretende arrancar com o processo de divulgação desde as fases iniciais (quer para os dados, quer para os metadados), a plataforma Dendro tira partido de mecanismos de disseminação já existentes. A disseminação é feita através do protocolo OAI-PMH, que estipula um conjunto de regras para a ligação entre repositórios e indexadores, facilitando a troca de informações entre os mesmos. A divulgação de dados no lado do EUDAT é feita pelo módulo B2Find. À semelhança dos restantes módulos, este é capaz de se ligar a diferentes componentes do EUDAT para indexar conteúdos que estejam disponíveis. Por outro lado, o B2Find também é capaz de recolher informação de repositórios que sejam compatíveis com este protocolo, redireccionando o utilizador que pesquisa nesta interface para o destino correspondente ao conjunto de dados que pretende consultar.

### **Simulação de transportes: um caso de teste**

A ligação das duas plataformas permite que vários cenários para a divulgação de dados sejam abrangidos. Como uma iniciativa de teste e divulgação destas ferramentas, foram contactados vários grupos de trabalho para identificar conjuntos de dados que estivessem prontos a serem depositados e divulgados. O procedimento de avaliação consistia em duas fases: uma entrevista inicial para avaliar o domínio do investigador e os seus hábitos de gestão de dados; e uma atividade de descrição destes dados na plataforma Dendro e consequente exportação para a plataforma EUDAT como passo final.

Como caso de teste foi selecionado um investigador no domínio da simulação de transportes que dispunha de dados prontos a depositar e que tinha servido de base para publicações na sua área. Durante a entrevista inicial, o investigador revelou que não dispunha de quaisquer indicações sobre como gerir os seus dados, ficando encarregue de o fazer, se pretendesse que estes fossem preservados. Relativamente à sua publicação, estes dados poderiam ser acedidos por outros investigadores, dado que serviam como base para testes de algoritmos desenvolvidos neste domínio, como por exemplo de alocação ou reserva de recursos. Quando questionado sobre como é partilhada a informação neste domínio, o investigador revelou desconhecer plataformas para o efeito, o que implica que para todos os projetos, seria necessário obter parcerias com empresas que pudessem ceder dados devidamente anonimizados para testes neste domínio. Na entrevista seguinte, dados de uma experiência feita durante o projeto mais recente foram depositados e descritos pelo investigador, adicionando descrições que este considerava importantes para facilitar a interpretação pela comunidade científica. No final deste processo de descrição, os dados foram exportados para a plataforma EUDAT, estando agora disponíveis para a comunidade e podendo igualmente serem citados (Oliveira, 2016). No seguimento deste contacto, outros investigadores procederam também ao depósito de conjuntos de dados disponíveis, mantendo agora um registo do seu depósito, tal como acontece com as suas publicações (Camacho, 2016) (Vieira, 2016).

## Conclusões

Dependendo da arquitetura de cada plataforma, a integração de plataformas para a gestão de dados pode trazer benefícios para a sua preservação e divulgação. Fruto do crescente interesse em propor ferramentas que se integrem no dia a dia dos investigadores, plataformas como o EUDAT têm vindo a ser apresentadas e melhoradas. Apesar de todas as funcionalidades que oferecem, estas plataformas ainda revelam alguma dificuldade em lidar com a qualidade dos dados, mais concretamente com a sua descrição. A existência de plataformas como o Dendro, com uma preocupação em capturar e descrever os dados assim que estes são produzidos podem contribuir para facilitar a sua interpretação e reutilização pela comunidade científica. A ligação entre as duas plataformas revelou ser uma solução para investigadores em vários estágios dos seus projetos e, em particular, foi capaz de dar resposta a cenários em que os projetos já terminaram, mas os dados ainda se mantêm sob responsabilidade dos investigadores.

## Referências bibliográficas

(Amorim, 2015) AMORIM, Ricardo Carvalho *et al.* – A Comparative Study of Platforms for Research Data Management: Interoperability, Metadata Capabilities and Integration Potential.

New Contributions in Information Systems and Technologies Advances in Intelligent Systems and Computing. 2015) 101-111. doi:10.1007/978-3-319-16486-1\_10.

(Amorim, 2014) AMORIM, Ricardo Carvalho *et al.* – LabTablet: Semantic Metadata Collection on a Multi-domain Laboratory Notebook. Communications in Computer and Information Science Metadata and Semantics Research. 2014 193-205. doi:10.1007/978-3-319-13674-5\_19.

(Burns, 2013) BURNS, C. Sean, Amy Lana, John Budd – Institutional repositories: exploration of costs and value. D-Lib Magazine 2013 19.1.

(Camacho, 2016) CAMACHO, Rui – Comparative Study of Classification Algorithms Using Molecular Descriptors in Toxicological Data Bases. University of Porto 2016 <http://hdl.handle.net/11304/fc173a9d-18c7-4949-ba72-15f8f5e5806d>

(Comp, 2016) AMORIM, Ricardo Carvalho *et al.* – A comparison of research data management platforms: architecture, flexible metadata and interoperability. Universal Access in the Information Society. 2016 1-12. doi:10.1007/s10209-016-0475-y

(Massimiliano, 2016) ASSANTE, Massimiliano, Leonardo Candela, Donatella Castelli, Alice Tani. – Are Scientific Data Repositories Coping with Research Data Publishing? Data Science Journal 15 2016 doi: <http://doi.org/10.5334/dsj-2016-006>.

(McLure, 2014) MCLURE Merinda, Level Allison V., Cranston Catherine L, Oehlerts Beth, Culbertson M. – Data curation: A study of researcher practices and needs. portal: Libraries and the Academy. 2014 14(2) 139-64. doi:10.1353/pla.2014.0009

(Oliveira, 2016) OLIVEIRA, Beatriz – Vehicle reservation assignment in car rental. University of Porto 2016, <http://hdl.handle.net/11304/12dce69e-6355-4353-b021-8887fed6528c>

(Silva, 2014) João Rocha, Cristina Ribeiro, João Correia Lopes – Ontology-based multi-domain metadata for research data management using triple stores. Proceedings of the 18th International Database Engineering & Applications Symposium. 2014 105-114. doi:10.1145/2628194.2628234

(Vieira, 2016) VIEIRA, Bruno – TweepProfiles Twitter data. University of Porto 2016 <http://hdl.handle.net/11304/93e8b144-5ada-46bb-974f-786b21b1ab06>

---

## Promover o acesso aberto com o *Open Journal System* : duas instituições, duas realidades

---

Paula Seguro-de-Carvalho

Escola Superior de Tecnologia da Saúde de Lisboa, IPL

[paula.carvalho@estesl.ipl.pt](mailto:paula.carvalho@estesl.ipl.pt)

Rui Teófilo

Escola Superior de Educação de Lisboa, IPL

[ruit@eselx.ipl.pt](mailto:ruit@eselx.ipl.pt)

### Resumo

Devido às alterações tecnológicas que ocorreram nos últimos anos foram introduzidas novas formas de adquirir, armazenar e transmitir a ciência. As instituições de ensino superior que têm revistas científicas procuram divulgá-las e dar-lhes visibilidade. Para o fazer muitas têm recorrido à utilização do *Open Journal System* (OJS), desenvolvido pelo *Public Knowledge Project*. Usando este sistema foi possível criar uma plataforma independente para a revista *Saúde & Tecnologia*, da ESTeSL, e para a revista *Da Investigação às Práticas*, da ESELx. O objetivo para a implementação do sistema foi facilitar as tarefas dos editores, dar maior visibilidade às revistas e disponibilizá-las em plataformas internacionais. Ambas as Escolas decidiram usar o OJS para a realização desta tarefa e, neste momento, as duas revistas dispõem de um espaço independente onde estão alojados os números já editados. A ESTeSL também disponibiliza a sua revista no *Directory of Open Access Journals* (DOAJ) e a ESELx na *Scientific Electronic Library Online* (SciELO) que, por sua vez, assegura o processo de submissão ao DOAJ. As vantagens da adoção desta plataforma são a disponibilização de uma plataforma *web* acessível a qualquer um dos elementos envolvidos no processo editorial, a qualquer hora e em qualquer lugar.

**Palavras-chave:** Acesso aberto, *Open Journal System*, Visibilidade



## Promote open access with *Open Journal System*: two institutions, two realities

### Abstract

Over the recent years, due to technology advances, new ways of acquiring, storing and transmitting science have emerged. Higher Education institutions publishing scientific magazines aim to publicize them and to increase their visibility. In order to achieve these goals, many have relied on the use of *Open Journal System* (OJS), developed by the *Public Knowledge Project*. By using OJS, ESTeSL and ESELx made possible the creation of an independent platform for their scientific magazines, respectively *Saúde & Tecnologia* and *Da Investigação às Práticas*. The purpose behind the implementation of this system was to facilitate editor's tasks, to give greater visibility to the magazines and to make them available in international platforms. Both Schools have decided to implement OJS in order to achieve those goals, and presently both magazines have independent space where edited issues are available. ESTeSL magazine also makes available its magazine on the *Directory of Open Access Journals* (DOAJ) and ESELx on the *Scientific Electronic Library Online* (SciELO) which in its turn ensures the process of submission to DOAJ. The great advantage on adopting this web platform is that it allows access to the players involved in the editorial process at anytime and anywhere.

**Keywords:** Open Access, *Open Journal System*, *Visibility*

### Introdução

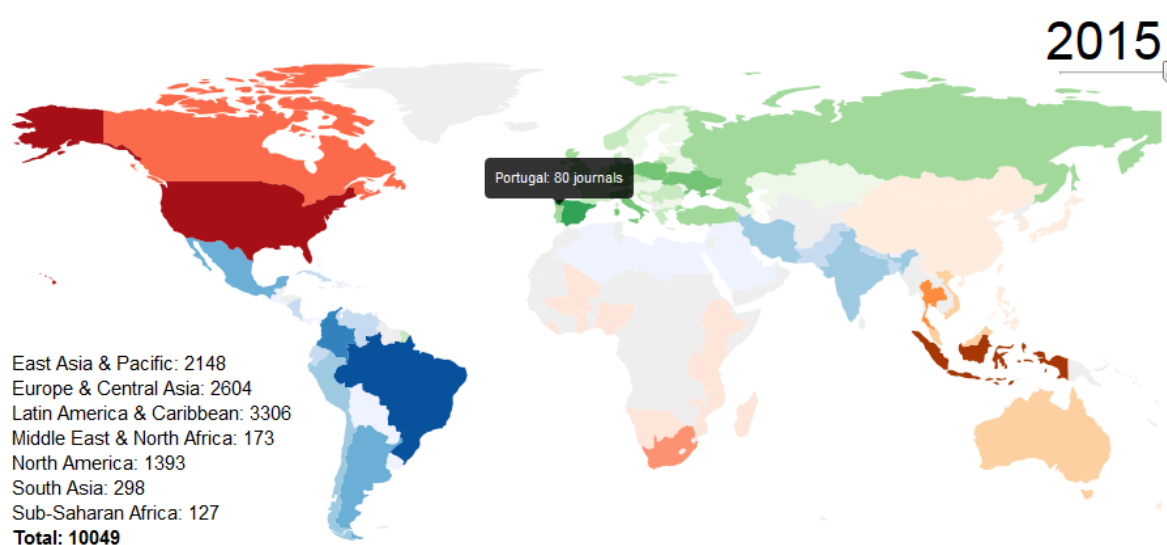
A publicação do conhecimento científico, bem como dos dados experimentais nos quais se baseia, permite à comunidade científica identificar erros, apoiar ou rejeitar teorias e reutilizar os dados para futuros estudos. A ciência tem uma grande capacidade de se autocorriger que provém desta abertura ao escrutínio e ao desafio possibilitado pelo livre acesso ao conhecimento (Boulton *et al.*, 2012, p. 7; Ferreira e Caregnato, 2014, p. 178).

As alterações tecnológicas generalizadas a que se assistiu nos últimos anos trouxeram novas formas de adquirir, armazenar e transmitir a ciência, criando novos hábitos de comunicação e de colaboração entre os cientistas (Boulton *et al.*, 2012, p. 7).

Um dos principais objetivos das instituições de ensino superior que têm revistas científicas é a sua divulgação e visibilidade (Ferreira e Caregnato, 2014, p. 178). Este objetivo levou a que muitas instituições optassem por *softwares* de edição e gestão de revistas científicas (Ataíde Dias, Delfino Junior e Moraes Silva, 2007, p. 76) que permitissem a integração destas revistas em diferentes fontes de divulgação de modo a ganharem maior visibilidade.

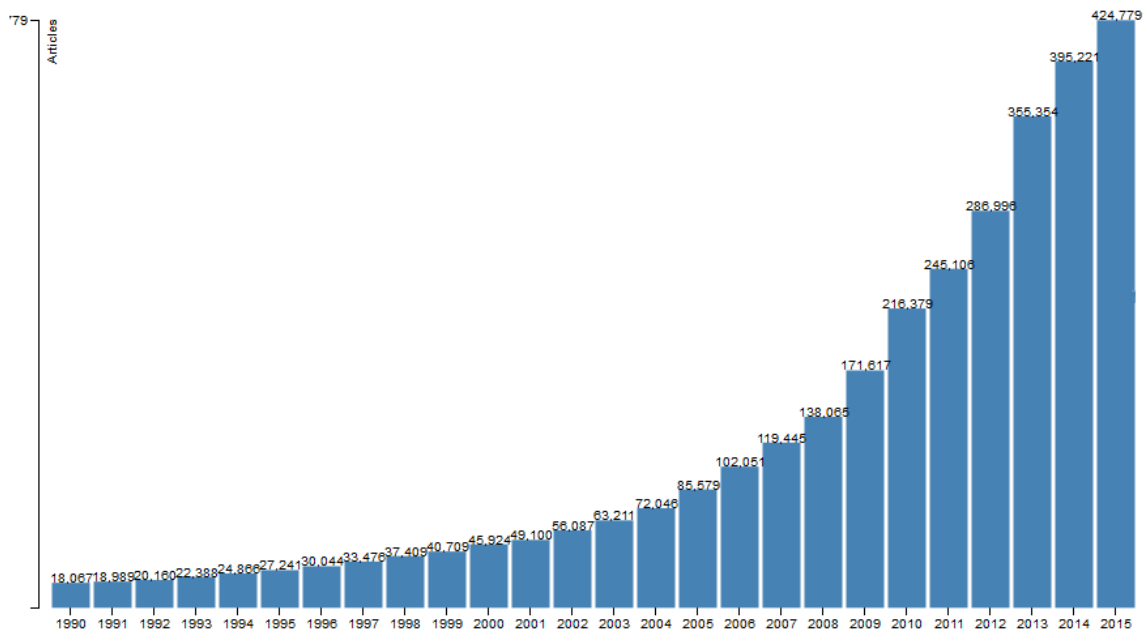
Um dos sistemas que ganhou uma grande visibilidade nos últimos anos foi o *Open Journal System* (OJS), desenvolvido pelo *Public Knowledge Project*, uma iniciativa financiada pelo Governo do Canadá que se encontra situado na Universidade de *British Columbia* (Ataíde Dias, Delfino Junior e Morais Silva, 2007, p. 76). Este sistema foi desenvolvido segundo os princípios de *software* livre, funciona na *web* e oferece uma solução integral que permite gerir todo o processo de publicação científica desde a submissão de um artigo, a seleção de revisores, a definição de lembretes automáticos e de prazos para entrega das correções até à sua publicação e indexação (Ybaseta-Medina e Alarco, 2012, p. 31). Disponibiliza também um vasto conjunto de documentação de apoio em todas as fases, desde a implementação, parametrização e processo de edição (López *et al.*, 2012, p. 219).

Este sistema está amplamente implementado em todo o mundo (Rojas e Rivera, 2011, p. 12), contando com mais de 10.000 revistas registadas no ano de 2015. A maior incidência é na América Latina e Caraíbas, seguida da Europa e Ásia Central. Em Portugal estavam registadas, à data de 2015, 80 revistas (Figura 1). Estas revistas disponibilizam um total de 424.779 artigos em *Open Access* (Figura 2).



**Figura 1.** Localização das revistas que usam o OJS

**Fonte:** <https://pkp.sfu.ca/ojs/ojs-usage/ojs-map/>



**Figura 2.** Total de artigos publicados com o OJS  
 Fonte: <https://pkp.sfu.ca/ojs/ojs-usage/ojs-stats/>

A Escola Superior de Tecnologia da Saúde de Lisboa (ESTeSL) e a Escola Superior de Educação de Lisboa (ESELx), ambas escolas que integram o Instituto Politécnico de Lisboa, abraçaram este desafio de divulgação das suas revistas científicas em acesso aberto numa plataforma e a sua posterior integração em índices e diretórios internacionais que potenciassem a sua divulgação. A ESTeSL optou por submeter a sua revista ao *Directory of Open Access Journals* (DOAJ) e a ESELx fez a submissão à *Scientific Electronic Library Online* (SciELO) que, por sua vez, assegurou o processo de submissão ao DOAJ.

### Objetivo

O objetivo deste trabalho é perceber como é que em duas unidades orgânicas da mesma instituição a implementação do mesmo serviço pode ter abordagens diferentes. Não se pretende fazer a comparação entre as duas revistas, mas dos processos de implementação das suas plataformas.

Criar uma plataforma independente para a revista *Saúde & Tecnologia*, da ESTeSL, e a revista *Da Investigação às Práticas*, da ESELx, teve como objetivo principal facilitar as tarefas dos editores recorrendo a um sistema de registo automático que permitisse um maior controlo e acompanhamento do fluxo editorial da publicação de cada número da revista, contribuindo para o aumento da transparência de todo o processo em consonância com as políticas defendidas no Horizonte 2020 (European Commission: Directorate-General for Research and Innovation, 2016, p. 2). Pretende-se reduzir o tempo e a energia dedicada a este processo complexo e moroso.

Mas, mais importante, pretende-se desenvolver novas possibilidades de disponibilização da informação na *web* que exponencialmente permitirão o aumento do trabalho colaborativo tanto a nível nacional como internacional, contribuindo para aumentar a visibilidade dos conteúdos, dos autores, da revista e da própria instituição.

### Implementação do projeto

Após a decisão de se usar o OJS para a realização desta tarefa, um *software* gratuito e sem custos de edição, ambas as escolas avançaram para o projeto.

Em ambas as Escolas foi solicitado ao serviço de informática a instalação do *software* para serem feitos os primeiros testes. No caso da ESELx, esta versão foi a que perdurou e que foi parametrizada pela Biblioteca com a colaboração do Serviço de Comunicação e Imagem na parte gráfica. No caso da ESTeSL, a parametrização do *software* também ficou a cargo da Biblioteca, mas o alojamento passou para os serviços de informática do IPL que se responsabilizam pelo apoio técnico.

Os bibliotecários das duas escolas responsáveis pelo projeto sentiram a necessidade de fazer formação que permitisse clarificar todas as potencialidades do *software* e tornar o processo de parametrização menos moroso. Visto este ser um processo tão individualizado, os bibliotecários das duas escolas procuraram partilhar as boas práticas e trabalhar em colaboração. Assim, foi necessário atualizar a versão do OJS que existia na ESELx, que era anterior à da ESTeSL e, depois em ambas as revistas, foi feita a integração do OJS com o *Google Analytics* para a obtenção de métricas de utilização.

Em seguida identificaram-se os números da revista e os artigos a incluir, recuperaram-se os PDFs para associar e, recorrendo ao *plugin* de submissão rápida, fez-se a integração de todos os números da revista *Saúde & Tecnologia* no OJS. Após a finalização da introdução de todos os artigos é necessário atualizar os dados de registo da revista no DOAJ para que, recorrendo ao *plugin* de exportação para o DOAJ disponível no OJS, seja feita a exportação dos dados completos da revista para esta plataforma.

A revista da ESELx, *Da Investigação às Práticas*, apresentava características diferentes porque, até ao ano de 2011, havia sido publicada apenas em papel. A partir deste ano foi abandonada a edição em papel e passou a ser publicada exclusivamente em formato eletrónico com um novo ISSN. Desta forma optou-se, nesta fase inicial, por incorporar no OJS apenas os números editados a partir de 2011, recorrendo ao *plugin* de submissão rápida.

Na *interface* com o utilizador, em ambas as revistas, a opção foi de selecionar um dos modelos disponibilizados no sistema e não fazer grandes alterações. Apenas foi parametrizado o *banner* superior de forma a identificar a revista e a instituição a que pertence.

## Resultados

Antes de iniciar este projeto, a revista *Saúde & Tecnologia* (ESTeSL) disponibilizava os seus artigos em PDF na página da revista incluída no *site* da Escola. Esta era a forma possível de garantir o acesso aos textos integrais da revista, mas que levantava vários constrangimentos: pouca visibilidade dos artigos, dos investigadores e da própria revista, não permitindo a realização de estatísticas de utilização.

Neste momento, a revista dispõe de um espaço independente onde estão os 15 números regulares da revista, dois números temáticos e dois suplementos publicados entre 2008 e 2016, com um total de três editoriais, oito artigos de revisão, três recensões críticas e 114 artigos originais.

A revista *Da Investigação às Práticas* (ESELx) quando começou a ser publicada em formato eletrónico disponibilizava os seus artigos em PDF, que podiam ser acedidos através da página do Centro Interdisciplinar de Estudos Educacionais (CIED). Atualmente disponibiliza, em espaço autónomo, os 13 números publicados entre 2011 e 2016 com um total de 13 editoriais, duas recensões críticas e 70 artigos originais. Após a aceitação da revista na plataforma SciELO iniciou-se, numa primeira fase, o moroso processo de conversão dos textos para o formato HTML e, numa segunda fase, a marcação dos artigos de forma a cumprir os requisitos definidos pela plataforma. Fruto deste trabalho, a plataforma da revista passou a disponibilizar também os artigos em formato HTML.

No que diz respeito ao DOAJ, a ESTeSL continua a aguardar um *feedback* relativamente à atualização dos dados da revista e finalização do processo de integração e a ESELx está à espera de confirmação da integração inicial. A integração da revista da ESELX na SciELO ocorreu no mês de agosto com a inclusão de cinco números da revista publicados entre 2014 e 2016.

Recorrendo às estatísticas de utilização do OJS verifica-se que, entre 1 de janeiro e 11 de outubro de 2016, os resumos dos artigos da revista *Saúde & Tecnologia* foram visualizados 2.110 vezes e que foi feito o *download* de 1.677 PDFs dos artigos. Na revista *Da Investigação às Práticas*, os resumos dos artigos foram visualizados 5.164 vezes e foi feito o *download* de 2.478 PDFs dos artigos.

Através do *Google Analytics* conclui-se que, entre 8 de junho e 11 de outubro, a página da revista *Da Investigação às Práticas* foi consultada 1.745 vezes, preferencialmente por utilizadores de Portugal (Figura 3) e a revista *Saúde & Tecnologia* foi alvo de 1.531 visitas, com maior incidência de utilizadores provenientes do Brasil (Figura 4).

Pais	Sessões
	1 745 % do total: 100,00% (1 745)
1. Portugal	1 128
2. Brazil	285
3. United Kingdom	211
4. Sweden	18
5. (not set)	12
6. Spain	10
7. France	9
8. United States	7
9. Angola	6
10. Czech Republic	5

**Figura 3.** Total de consultas da revista *Da Investigação às Práticas* por país (dados recuperados do Google Analytics)

Pais	Sessões
	1 531 % do total: 100,00% (1 531)
1. Brazil	893
2. Portugal	383
3. United Kingdom	197
4. Angola	15
5. United States	12
6. Belgium	6
7. China	3
8. Colombia	3
9. Cape Verde	2
10. Germany	2

**Figura 4.** Total de consultas da revista *Saúde & Tecnologia* por país (dados recuperados do Google Analytics)

## Conclusão

A implementação do sistema de gestão editorial OJS nas revistas *Saúde & Tecnologia* e *Da Investigação às Práticas* é um processo ainda em curso e que terá de enfrentar vários desafios e alguns constrangimentos.

O primeiro desafio foi fazer com que a comunidade científica envolvida nas revistas percebesse a importância e pertinência da implementação deste sistema para a promoção e visibilidade das revistas; a utilização e uma nova ferramenta que pode apresentar alguma complexidade é sempre um fator que contribui para uma certa resistência à mudança (López *et al.*, 2012, p. 227).

Um dos constrangimentos sentido em ambas as instituições decorre da dificuldade de incorporar os diferentes papéis desempenhados por diferentes utilizadores, nomeadamente os autores e os revisores, cuja adaptação ao sistema é, por vezes, difícil

(López *et al.*, 2012, p. 227). Para minimizar esta questão é necessária a construção de um conjunto de tutoriais direcionados para estes utilizadores e disponibilizados no próprio sistema. É também de considerar a implementação de um plano de formação que possa colmatar estas dificuldades (Ataíde Dias, Delfino Junior e Morais Silva, 2007, p. 81).

O maior desafio, que é comum a ambas as Escolas, decorre da falta de recursos humanos para a realização destas tarefas, o que torna o desenvolvimento dos processos pouco ágil e dificulta a definição de metas temporais para a conclusão destes trabalhos iniciais.

Apesar de todas as dificuldades e constrangimentos, as vantagens da adoção desta plataforma são evidentes: estas revistas passaram a dispor de uma plataforma *web* que está acessível a qualquer um dos elementos envolvidos no processo editorial, a qualquer hora e em qualquer lugar. Permite manter um registo de todas as fases do processo e automatizar algumas das tarefas, simplificando a gestão da revista (López *et al.*, 2012, p. 227-228; Ybaseta-Medina e Alarco, 2012, p. 31).

Com a implementação do OJS para a disponibilização das revistas foi possível garantir aos utilizadores a leitura, cópia, distribuição, *download*, impressão dos artigos científicos, dando a possibilidade de virem a ser citados em futuras pesquisas ou em investigações subsequentes (Ybaseta-Medina e Alarco, 2012, p. 31).

Em termos futuros analisar-se-á se a disponibilização das duas revistas na plataforma OJS contribuiu para que tenham alcançado um maior número de membros das respetivas comunidades científicas, contribuindo para o aumento da sua visibilidade. Ao implementar esta plataforma conseguiu-se trazer para as publicações científicas das duas instituições algumas das características que promovem a visibilidade das revistas científicas: estarem disponíveis na *web*, serem disponibilizadas em acesso aberto e estarem presentes em bases de dados internacionais (Ferreira e Caregnato, 2014, p. 178). O passo seguinte será avaliar o verdadeiro impacto do novo formato de funcionamento das revistas *Saúde & Tecnologia e Da Investigação às Práticas*.

## Referências bibliográficas

ATAÍDE DIAS, Guilherme; DELFINO JUNIOR, João Bosco; MORAIS SILVA, José Wendell (2007) – Open journal systems – OJS: migrando um periódico científico eletrônico para um sistema automatizado de gerência e publicação de periódicos científicos eletrônicos. *Informacao & Sociedade: Estudos* [Em linha]. Vol. 17, Nº 2, p. 75-82. [Consult. 25 jan. 2016]. Disponível na internet: <URL:<http://www.ies.ufpb.br/ojs/index.php/ies/article/view/826/1448>>. ISSN 0104-0146

BOULTON, Geoffrey *et al.* (2012) – *Science as an open enterprise* [Em linha]. London : The Royal Society. [Consult. 25 jan. 2016]. Disponível na internet:

<URL:[http://royalsociety.org/uploadedFiles/Royal\\_Society\\_Content/policy/projects/sape/2012-06-20-SAOE.pdf](http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf)>. ISBN 9780854039623.

European Comission: Directorate-General for Research and Innovation (2016) – *H2020 programme: guidelines on open access to scientific publications and research data in Horizon 2020* [Em linha]. Brussels : European Commision. [Consult. 12 mar. 2016]. Disponível na internet: <URL:<https://ai.tecnico.ulisboa.pt/files/sites/52/guidelines-scientific-publicationsresearch-data-in-h2020.pdf>>.

FERREIRA, Ana Gabriela Clipes; CAREGNATO, Sônia Elisa (2014) – Visibilidade de revistas científicas: um estudo no Portal de Periódicos Científicos da Universidade Federal do Rio Grande do Sul. *Transinformação* [Em linha]. Vol. 26, N° 2, p. 177-190. [Consult. 1 abr. 2016]. Disponível na internet: <DOI: [10.1590/0103-37862014000200007](https://doi.org/10.1590/0103-37862014000200007)>. ISSN 0103-3786.

LÓPEZ, Yolibeth; TORRÉNS, Rodrigo; VILORIA, Adriana; RAMÍREZ, Mariela (2012) – OJS: una herramienta de gestión editorial en línea: estrategias para su adopción en la universidad de Los Andes. *Anuario Electrónico de Estudios en Comunicación Social Disertaciones* [Em linha]. Vol, 5, N°1, p. 217-228. [Consult. 19 abr. 2016]. Disponível na internet: <URL:<http://dialnet.unirioja.es/descarga/articulo/4042012.pdf>>. ISSN 1856-9536.

ROJAS, MA; RIVERA, Sandra (2011) – *Guía de buenas prácticas para revistas académicas de acceso abierto* [Em linha]. Santiago de Chile : ONG Derechos Digitales. [Consult. 10 jun. 2016]. Disponível na internet: <URL: [http://www.latindex.org/lat/documentos/descargas/Manual-Buenas\\_Practica\\_Revistas\\_Academicas.pdf](http://www.latindex.org/lat/documentos/descargas/Manual-Buenas_Practica_Revistas_Academicas.pdf)>.

YBASETA-MEDINA, Jorge; ALARCO, J. Jhonnell (2012) – La iniciativa Open Journal Sistem (OJS) y su implementación en la Revista Médica Panacea. *Revista Médica Panacea* [Em linha]. Vol. 2, N° 2, p. 31-32. [Consult. 5 jun. 2016]. Disponível na internet: <URL:<file:///C:/Users/Utilizador/Downloads/32-86-1-PB.pdf>>.



---

## Vocabulários controlados na descrição de dados de investigação no Dendro

---

Yulia Karimova

INESC TEC – Faculdade de Engenharia da Universidade do Porto

[ylaleo@gmail.com](mailto:ylaleo@gmail.com)

João Aguiar Castro

INESC TEC – Faculdade de Engenharia da Universidade do Porto

[joaoaguiarcastro@gmail.com](mailto:joaoaguiarcastro@gmail.com)

### Resumo

Tendo em conta que a reutilização dos dados de investigação contribui para aumentar a visibilidade dos resultados científicos, a gestão de dados é cada vez mais valorizada. Neste contexto a descrição de dados, apesar de exigente, é uma tarefa muito importante. O Dendro, plataforma colaborativa de gestão de dados intermédia, está a ser desenvolvida na Universidade do Porto para auxiliar os investigadores na descrição de dados de investigação assim que começam a ser produzidos para posterior transferência para repositórios. A falta de tempo, ferramentas, experiência ou conhecimento na gestão dos dados, assim como as questões de criação de metadados de boa qualidade estão entre limitações identificadas na descrição e publicação de dados por parte de investigadores. Com este trabalho pretende-se introduzir vocabulários controlados para simplificar e normalizar a descrição de dados na plataforma Dendro. Estes são apresentados sob a forma de listas de escolhas e formalizados como ontologias. Os vocabulários controlados para descrição de dados de investigação foram testados em colaboração com investigadores do domínio da Produção de Hidrogénio. Os resultados de análise de qualidade de metadados sugerem que o uso de vocabulários controlados simplificou o processo de descrição, obtendo-se uma descrição mais completa e correta.

**Palavras-chave:** Gestão de dados de investigação, metadados, vocabulários controlados, qualidade de metadados, Dendro

## Controlled vocabularies in the description of research data on Dendro

### Abstract

Considering that the reuse of research data contributes to an increase in the visibility of scientific results, data management is increasingly valued. In this context, data description, despite being demanding, is an important task. Dendro, a collaborative data management platform, is being developed at the University of Porto to help researchers in the description of research data from their production to the subsequent transfer to repositories. The lack of time, tools, experience, knowledge of data management and supply of good quality metadata are among the limitations identified in the description and publication of data by researchers. This work proposes controlled vocabularies to simplify and normalize the data description in the Dendro platform. They are presented in the form of drop-down lists and formalized as ontologies. The controlled vocabularies for research data description were tested in collaboration with researchers from the “Hydrogen Production” domain. The results of the analysis of the metadata quality suggest that the use of controlled vocabularies has simplified the process of description, obtaining a more complete and accurate description.

**Keywords:** Research data management, metadata, controlled vocabularies, metadata quality, Dendro

### Introdução

Atualmente, com o desenvolvimento de novas tecnologias de informação começam a surgir novas formas e métodos de cooperação científica. Consequência disso são o aparecimento de vários repositórios de dados de investigação que ajudam na promoção, divulgação e valorização dos mesmos. O acesso, partilha e reutilização de dados de investigação, ao diminuir a redundância na criação de novos dados, permite gerar conhecimento científico de forma mais expedita.

Devido à sua diversidade e complexidade, os dados de investigação, são por normas difíceis de interpretar (Smith, Seligman e Swarup, 2008), o que compromete a sua reutilização. Neste contexto os metadados têm um papel fundamental, uma vez que permitem o registo da informação necessária para que se possam compreender os dados, para além de auxiliarem na preservação, localização e recuperação dos objetos digitais. A publicação de dados, sem que as condições para a reutilização dos mesmos estejam assegurada, corresponde a um desperdício de recursos. Por este motivo, os investigadores devem assegurar que os dados são acompanhados por metadados suficientemente detalhados e rigorosos. No entanto, a descrição de dados é uma tarefa exigente e demorada,

que se não for devidamente suportada, pode desmotivar os investigadores a enveredar por atividades relacionadas com a gestão de dados.

A falta de tempo e ferramentas, falta de experiência e conhecimento técnico são alguns fatores que condicionam a partilha e publicação de dados de investigação em diversos domínios de investigação. Torna-se, portanto, importante fornecer aos investigadores ferramentas que facilitem na captura de metadados e potenciem a divulgação dos dados de investigação que estes recolheram.

Apesar de várias normas de metadados estarem disponíveis, estas são complexas, e por vezes inexistentes em certos domínios (Swan e Brown, 2008). Por outro lado, a usabilidade também influencia na utilização de um repositório de dados e pode afetar o envolvimento dos utilizadores (Zhang, Maron e Charles, 2013). Para além disso, muitos projetos não têm uma infraestrutura integrada de gestão de dados, o que muitas vezes resulta no uso de ferramentas para gestão de dados não adequadas. Este problema surge especialmente em projetos pequenos, que têm exigências mínimas na utilização das ferramentas tecnológicas, com recursos humanos e conhecimentos na curadoria digital limitados (Akmon *et al.*, 2011; Borgman, Wallis, Enyedy, 2007).

Tendo em conta a importância dos desafios reconhecidos tanto na gestão de dados de investigação como no processo de descrição de dados, a plataforma Dendro tem vindo a ser desenvolvida na Universidade do Porto. Através do Dendro pretende-se dar o apoio necessário aos investigadores na preparação dos dados de investigação, nomeadamente na captura atempada de metadados, para que os dados sejam transferidos, se possível, para repositórios externos, em condições de serem recuperados e reutilizados por terceiros.

Um dos objetivos deste trabalho consiste na agilização e facilitação do processo de descrição de dados na plataforma Dendro, de forma a motivar o interesse dos investigadores na organização e publicação dos dados que produzem. Neste contexto os vocabulários controlados apresentam-se como uma boa ferramenta, que pode facilitar o processo da descrição de dados e ao mesmo tempo contribuir para uma melhoria na qualidade dos metadados. Em muitos casos, os vocabulários controlados definem o conteúdo admissível para um determinado elemento descritivo e podem ser incorporados nos procedimentos de automatização de forma simples, contribuindo assim para o controlo de qualidade (Bermudez *et al.*, 2011). Com esta ideia em mente este estudo propõe a criação de vocabulários controlados para a descrição de dados em domínios científicos específicos no Dendro.

## Dendro

O Dendro é uma plataforma colaborativa de gestão de dados intermédia, que permite a organização e documentação de dados desde o momento em que começam a ser

produzidos, sobretudo para investigadores na cauda longa da ciência, que se refere aos investigadores individuais e pequenos laboratórios, que não têm possibilidade de realizar gestão de grandes quantidades de dados. Esse tipo de dados é mais difícil de encontrar, reutilizar e preservar (Heidorn, 2008). O Dendro consiste numa interface web e tem a capacidade para exportar pacotes de dados e metadados para repositórios destinados à sua preservação e divulgação.

Esta plataforma facilita a criação de metadados com a utilização de vários esquemas de metadados existentes, como por exemplo *Dublin Core* e *Friend of a Friend*, assim como descritores específicos a domínios de investigação. A plataforma utiliza descritores, baseados em ontologias criadas em colaboração com investigadores dos mais diversos domínios e um curador de dados (Silva *et al.*, 2014). Os investigadores, como especialistas no seu domínio e produtores dos dados, fornecem conhecimento especializado sobre o contexto dos dados a descrever, enquanto o curador tem uma perspetiva mais abrangente sobre a gestão de dados de investigação.

De forma a adicionar novos descritores no Dendro, foram criadas ontologias para vários domínios científicos, tais como: Produção de Hidrogénio, Biodiversidade, Química Analítica, entre outros. Uma vez carregadas no Dendro, estas ontologias contribuem com novos descritores que podem ser combinados pelos investigadores para obter registos de metadados o mais abrangente quanto possível (Castro *et al.*, 2015; Silva *et al.*, 2014).

Decorrente da interação com os investigadores surgiu a necessidade de desenvolver vocabulários controlados, de modo a melhorar a qualidade das descrições e a tornar a tarefa de descrição mais apelativa para os investigadores.

### **Vocabulários controlados**

Em muitos casos, a expressão vocabulário controlado define o conteúdo admissível para um determinado elemento de metadados e pode ser facilmente incorporado nos procedimentos de automatização, contribuindo para o controlo de qualidade, ao fornecer aos utilizadores uma lista de entradas permitidas para os elementos de metadados específicos (Bermudez *et al.*, 2011).

O uso de vocabulários controlados permite ultrapassar as seguintes limitações (National Information Standards Organization, 2005):

- Diferença na interpretação do léxico (variações conceptuais);
- Diferença na utilização de expressões lexicais (variações sociais);
- Expansão da significação do léxico (polissemia);
- Desconhecimento do léxico.

Segundo Hedden (2010) há vários tipos de vocabulários controlados: a lista de conceitos, o ficheiro de autoridade, a taxonomia, os tesouros. Além disto, os vocabulários controlados podem ser abertos (“open-ended”), onde novos conceitos podem ser adicionados ao longo do tempo (Harpring, 2010), e fechados, onde não existe a possibilidade de inserir as novas sugestões.

Algumas implementações de vocabulários controlados são implementadas através de listas de escolhas, que permitem mostrar todas as opções existentes e ajudam ao utilizador escolher um termo num conjunto de conceitos pré-definidos. Normalmente estas listas não incluem sinónimos e são mais fáceis de implementar, sendo adequadas para a melhoria de qualidade na descrição de dados e, portanto, optou-se por seguir esta abordagem no desenvolvimento deste trabalho.

### Seleção do caso de estudo

O envolvimento dos investigadores no desenvolvimento das ferramentas que suportam as atividades de descrição de dados é tido como muito importante. Isto porque, os investigadores são as pessoas com maior aptidão para indicar a informação necessária para permitir interpretar os dados que produziram. O problema dos esquemas de metadados é a sua especialização, o que impede que os investigadores os usem, muitas vezes por não se sentirem familiarizados com a terminologia utilizada. Esta limitação leva a que os investigadores devam ser consultados tanto na seleção de descritores como na definição dos vocabulários controlados correspondentes.

O primeiro dos projetos selecionados, sem ferramentas adequadas para gestão de dados de investigação surge no contexto da Produção de Hidrogénio, num grupo de investigação da Universidade do Porto. Este grupo orienta a investigação à produção instantânea de hidrogénio através da hidrólise catalítica do borohidreto de sódio (*NaBH<sub>4</sub>*) para dispositivos portáteis (telemóvel, tablet, mp3, etc.). Os dados experimentais deste grupo são armazenados principalmente em folhas de Excel. Os sensores, conectados ao reator, estão ligados a um computador com software específico - *LabView*, que grava os dados de temperatura, pressão e outras medições relevantes obtidos durante as experiências no reator num ficheiro *Excel*. A partilha de dados é feita, sobretudo, através de *email*. Em outros casos, os dados são copiados do computador para discos externos.

O trabalho realizado juntamente com este grupo de investigadores incluiu várias entrevistas e experiências de descrição de dados, que permitiram identificar as necessidades existentes na gestão de dados deste grupo. Durante as entrevistas os investigadores afirmaram que a descrição dos dados seria uma grande vantagem, nomeadamente para incentivar a sua partilha, publicação e reutilização dos dados. Esta colaboração levou à definição de descritores como *Additive*, *Catalyst*, *Reactor Type*, entre outros, entretanto

utilizados em experiências de descrição de dados no Dendro, em que estes investigadores foram participantes.

Os resultados desta experiência foram então avaliados para aferir a qualidade dos metadados criados por este grupo. A Tabela 1 mostra que, por exemplo, para o mesmo registo, diferentes investigadores usam diferentes conceitos. Isto é o caso da descrição do tipo de reator usado na experiência (Reactor Type), em que *ovoid* é o mesmo que *Egg Reactor*, o que levanta problemas de consistência nas descrições.

Descritor	Valor de descritor do utilizador 1	Valor de descritor do utilizador 2	Valor de descritor do utilizador 3
Temperature	25°C	24	28°C
Hydration Factor	16	16	15
Reactor Type	RG/RM	ovoid	Egg Reactor / Conical Small Reactor
Reagent	NaBH <sub>4</sub>	Sodium Borohydride	NaBH <sub>4</sub>
Catalyst	Ni-Ru	Nickel-ruthenium	NiRu
Gravimetric Capacity	<5wt%	2,3	1.9wt%
Hydrolysis	Hidrolise clássica	alkali	Classic hydrolysis

**Tabela 1:** Exemplos da diferença na descrição

A análise dos registos de metadados recolhidos reforçou a necessidade da criação de uma ferramenta que além de simplificação de processo de descrição no geral, ajude a melhorar a qualidade dos metadados no Dendro. Por isso mesmo, partiu-se para a elaboração de vocabulários controlados.

### Elaboração de vocabulários controlados

A formalização de vocabulários controlados tirou partido da ontologia desenvolvida para o domínio Produção de Hidrogénio e da avaliação das descrições feitas por parte dos investigadores em Produção de Hidrogénio. Com os investigadores foram definidos os conceitos a incluir no vocabulário controlado para determinado descritor (Tabela 2).

Descritor	Conceitos definidos para vocabulários controlados
Additive	CMC SDS
Catalyst	Co-B Co-B/Ni Co-Mn-B Ni-Ru Pt/C
Hydrolysis	Acid hydrolysis Alkali-free hydrolysis Classic hydrolysis
Reactor Type	EggR - ovoid mini reactor LR - large reactor MRc - conical medium reactor

	<p>Mrf – flat medium reactor                  SRc – conical small reactor                  SRf – flat small reactor</p>
Reagent	<p>KBH4                  LiAlH4                  LiBH4                  NH3BH3                  NaBH4</p>

**Tabela 2:** Descritores e conceitos definidos para vocabulários controlados

De acordo com literatura, existem várias maneiras de modelar vocabulários controlados numa ontologia.

Segundo a *W3C Recommendation – OWL Web Ontology Language Reference* os vocabulários controlados podem ser criados na ontologia como *Annotation Property*. O mesmo acontece no caso da ontologia sobre anatomia humana, da *Foundational Model of Anatomy* (Golbreich, Zhang e Bodenreider, 2006), que tem como base o *NCI-Thesaurus de National Cancer Institute* (Coronado, de e Fragoso, 2004). Já o *OWL FULL* não coloca quaisquer restrições sobre as anotações numa ontologia e o *OWL DL* permite anotações em classes, propriedades, indivíduos e cabeçalhos de ontologias. Além disso, afirma-se que é possível especificar o tipo de valor de um literal numa indicação da *Annotation Property*. Existem cinco *Annotation Property* pré-definidas, que podem ser utilizadas para anotação de *DataProperties* (Bechhofer *et al.*, 2004), estas são: *VersionInfo*; *label*; *comment*; *seeAlso*; *isDefineBy*.

Com base nestes estudos decidiu-se modelar os vocabulários controlados da ontologia de Produção de Hidrogénio através de *Annotation Property – has Alternative*, que é uma das alternativas possíveis para o valor de um descritor como conceito de vocabulário controlado, que por sua vez descreve determinada *Data Property*. A modelação da ontologia com os vocabulários controlados foi realizada utilizando o software Protégé (<http://protege.stanford.edu/>), onde os conceitos definidos foram associados aos respetivos descritores (Figura 1).

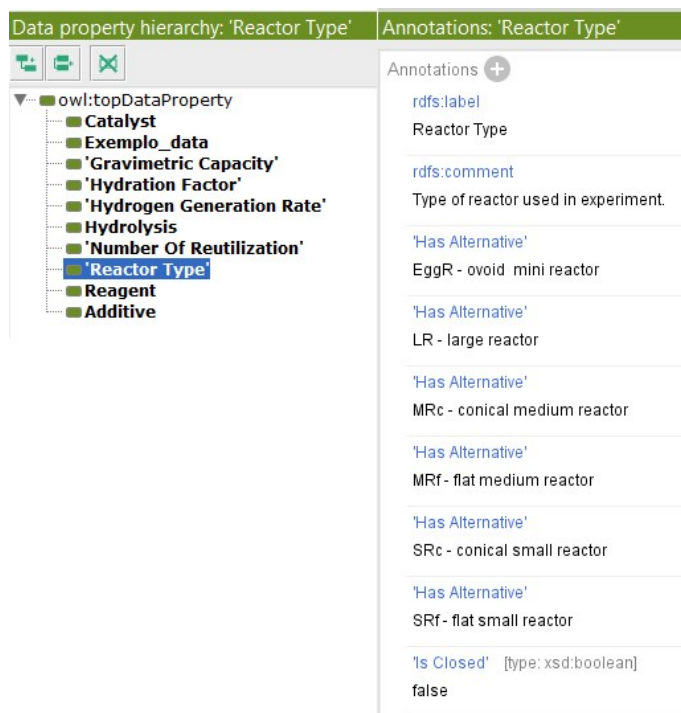


Figura 1: Conceitos de vocabulários controlados de descritor *Reactor Type* no *Protégé*

De acordo com esta modelação os conceitos de vocabulários controlados estão na ontologia e não é no Dendro, o que os permite expandir a utilização da plataforma para qualquer outro domínio de investigação.

Após a implementação da ontologia, os descritores com vocabulários controlados surgem no Dendro, tal como ilustrado na Figura 2.

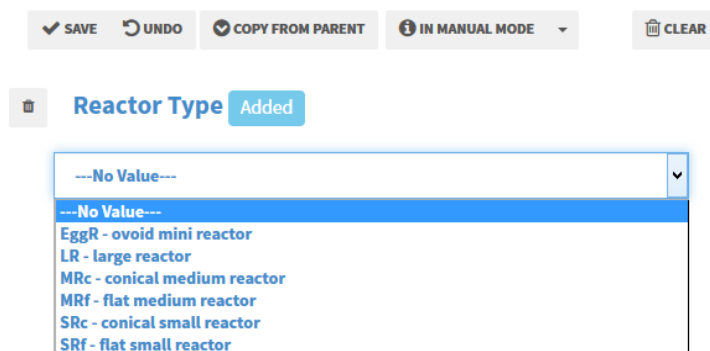


Figura 2: Vocabulários controlados de descritor *Reactor Type* no *Dendro*

## Definição de métricas de qualidade de metadados

De modo a verificar se a implementação de vocabulários controlados facilita o processo de descrição no geral e ajuda a aumentar a qualidade dos metadados, foram definidas métricas de avaliação. Estas métricas servem para determinar a qualidade dos



metadados através da comparação entre descrições feitas sem o uso de vocabulários controlados e descrições com o uso dos mesmos.

De acordo com vários estudos (Alkhattabi, Neagu e Cullen, 2010; Bruce e Hillmann, 2004; Moreira *et al.*, 2009; Ochoa e Duval, 2006; Palavitsinis, 2013; Stvilia *et al.*, 2007) foram identificadas as métricas que melhor se adequam ao nosso caso:

1. **Correctness** – grau em que a linguagem utilizada nos metadados é sintaticamente e gramaticalmente correta;

2. **Completeness** – número de descritores preenchidos em comparação com o número total de descritores;

3. **Conformance to expectations** – grau em que o registo de metadados preenche os requisitos de uma determinada comunidade de utilizadores;

4. **Overall Rating** – pontuação geral do registo de metadados, tendo em conta as métricas anteriores.

E as métricas de avaliação da usabilidade de plataforma Dendro:

5. **Satisfaction** – grau de satisfação de utilizador após da experiência;

6. **Task time** – grau de rapidez da descrição de dados.

Para proceder à análise da qualidade dos metadados no domínio de Produção de Hidrogénio, de acordo com as métricas escolhidas, realizou-se uma série de experiências de descrição de dados na plataforma Dendro.

A primeira análise teve em conta metadados obtidos manualmente em três experiências de descrição de dados, anteriores à implementação dos vocabulários controlados, efetuadas por utilizadores do domínio indicado. Para determinar a qualidade dos metadados a utilização dos conceitos presentes nos vocabulários controlados é tida como uma descrição com cem por cento de qualidade. O registo de metadados, por exemplo, aplicando a métrica *Correctnes* é calculado através da comparação entre o número total de palavras na inserção de um conceito em linguagem natural, com o numero total de palavras do mesmo conceito definido no vocabulário controlado para determinado descritor. O cálculo de *Correctness* não tem em conta a existência de pequenas discrepâncias na escrita, tais como a diferença em letras maiúsculas e minúsculas, vírgulas, acentos, entre outros.

Para além disso foram consideradas as respostas dos investigadores a um inquérito realizado neste contexto. A análise realizada mostrou que a qualidade de descrição sem utilização de vocabulários controlados estava aquém das expectativas, quer seja pela diferença na descrição de dados de investigação, lapsos e registos incompletos. Apesar disto, os utilizadores ficaram satisfeitos com a utilização da plataforma Dendro e não necessitaram de muito tempo para descrever os dados de investigação que produziram.

A segunda ronda de experiências contou com a participação dos mesmos utilizadores, desta vez com recurso a vocabulários controlados. Para possibilitar a comparação de resultados entre as experiências, a análise após a implementação de vocabulários controlados baseou-se nas mesmas métricas e regras definidas na análise anterior. Contudo, a utilização dos conceitos presentes nos vocabulários controlados é tida como uma descrição com cem por cento de qualidade.

Os resultados da segunda análise mostram que a qualidade de descrição de dados de investigação com utilização de vocabulários controlados aumentou. Aliás, todos os valores ficaram muito próximos dos cem por cento. Isto indica que a utilização de vocabulários controlados diminui as diferenças na descrição, erros, e ajudou a obter uma descrição mais completa.

As experiências da descrição dos dados de investigação realizadas no Dendro com utilização de vocabulários controlados deixaram os utilizadores satisfeitos com a usabilidade da plataforma. Para além disso, o tempo que cada investigador precisou para completar a tarefa diminuiu.

### Avaliação do uso de vocabulários controlados

A comparação dos resultados de análises de descrição de dados de investigação mostra que a maioria dos valores de qualidade dos metadados aplicando as métricas definidas subiu. Por exemplo, o valor de qualidade de metadados de descritor *Reactor Type*, aplicando métrica *Correctness* passou de 28% para 100% e aplicando a métrica *Conformance to expectations* aumentou de 36% para 100%. Tal como se compreende, através da Tabela 3, a qualidade da descrição após a implementação dos vocabulários controlados melhorou.

Durante a realização das experiências da descrição os investigadores utilizaram praticamente a mesma quantidade de descritores e não tiveram grandes dificuldades na utilização da plataforma. O nível da satisfação teve um aumento pouco significativo. Contudo, afirmaram que a utilização de vocabulários controlados facilita o processo da descrição. Analisando as respostas de inquéritos e comentários de utilizadores conclui-se que os investigadores preferem criar os metadados com auxílio de vocabulários controlados, pois a comparação dos valores na Tabela 3 obtidos aplicando a métrica *Task Time*, mostram que o tempo médio necessário para realização de uma tarefa diminuiu.

Qualidade		
Descritor	Qualidade antes de implementação de vocabulários controlados ( <i>Correctness</i> )	Qualidade após de implementação de vocabulários controlados ( <i>Correctness</i> )
Reactor Type	28%	100%
Hydrolysis	51%	100%
Catalyst	50%	100%
Reagent	50%	100%
Additive		100%

	Qualidade antes de implementação de vocabulários controlados ( <i>Conformance to expectations</i> )	Qualidade após de implementação de vocabulários controlados ( <i>Conformance to expectations</i> )
Reactor Type	36%	100%
Hydrolysis	69%	100%
Catalyst	75%	100%
Reagent	75%	100%
Additive		50%
	Qualidade antes de implementação de vocabulários controlados ( <i>Completeness</i> )	Qualidade após de implementação de vocabulários controlados ( <i>Completeness</i> )
Média	75%	82%
	Qualidade antes de implementação de vocabulários controlados ( <i>Overall Rating</i> )	Qualidade após de implementação de vocabulários controlados ( <i>Overall Rating</i> )
Reactor Type	46%	94%
Hydrolysis	65%	94%
Catalyst	67%	94%
Reagent	67%	94%
Additive		77%
	Usabilidade sem utilização de vocabulários controlados	Usabilidade com utilização de vocabulários controlados
Satisfaction	4	4.5
	<i>Task Time</i> de tarefa (média) sem utilização de vocabulários controlados	<i>Task Time</i> de tarefa (média) com utilização de vocabulários controlados
	17 min	10 min
	<i>Task Time</i> por descritor (média) sem utilização de vocabulários controlados	<i>Task Time</i> por descritor (média) com utilização de vocabulários controlados
	62 seg	38 seg

Tabela 3: Comparação de resultados de avaliação de qualidade de dados antes e após a implementação de vocabulários controlados

Sendo assim, a comparação dos resultados demonstra que o uso de vocabulários controlados na criação dos metadados facilita o processo de descrição em geral, obtendo-se uma descrição mais completa e correta.

### Conclusões e trabalhos futuros

O estudo realizado sobre gestão de dados de investigação, criação de metadados e importância de qualidade dos mesmos ilustrou vários problemas existentes nesta área. O processo de descrição de dados exige competências, esforço, tempo e ferramentas adequadas, pois só os metadados de qualidade garantem a precisão e acesso completo aos recursos digitais e permitem aos utilizadores finais encontrar e recuperar os recursos que precisam. O interesse e a motivação de investigadores tanto na gestão de dados como na descrição de dados e na escolha de um sistema para criação dos metadados depende de vários fatores, tais como a usabilidade, utilidade, simplicidade, facilidade, entre outros.

O trabalho desenvolvido enquadró o desafio da descrição de dados no processo de desenvolvimento da plataforma de gestão de dados de investigação Dendro da Universidade do Porto. Pretendeu-se encontrar uma solução para agilizar e facilitar todo o processo de descrição de dados de investigação e, assim, contribuir para a melhoria da qualidade dos metadados criados no Dendro.

Em suma, os objetivos deste trabalho foram segmentados em três vertentes. Em primeiro lugar procedeu-se à escolha do domínio Produção de Hidrogénio como caso de estudo, a realização de uma série de experiências de descrição de dados no Dendro e recolhidos os registos de metadados. Em segundo lugar foram elaborados vocabulários controlados para o domínio escolhido e após a sua implementação realizou-se a segunda ronda de experiências de descrição de dados de investigação. Em terceiro lugar definiram-se as métricas para a avaliação de qualidade dos metadados e procedeu-se à análise da qualidade dos metadados criados sem e com utilização de vocabulários controlados, com a finalidade de demonstrar que a implementação de vocabulários controlados facilita o processo de descrição e melhora a qualidade de descrição de dados de investigação no Dendro.

Mediante os resultados obtidos pode-se afirmar que os objetivos foram alcançados. Em particular, a descrição efetuada com uso de vocabulários controlados simplificou todo o processo de criação de metadados, permitiu obter descrições completas e mais rigorosas, com a vantagem de o permitir sem o aumento do tempo necessário para o efeito.

Uma das perspetivas de trabalho futuro é a elaboração de vocabulários controlados para outros domínios de investigação existentes no Dendro e para descritores genéricos, tal como *Language*, *Format*, entre outros. Outro dos objetivos é implementar vocabulários controlados abertos e fechados.

A elaboração de expressões regulares é mais uma estratégia que pode ser utilizada na simplificação do processo de descrição de dados no Dendro. (Grimalovskii, 2013; Skoglund, 2011; Standen, 2010). De acordo com a literatura, as expressões regulares podem reduzir o esforço manual na introdução da informação e ajudar na qualidade de dados. Elas são complexas aquando da sua boa configuração e funcionamento trazem sempre benefícios da sua utilização (Friedl, 2006).

Para terminar, pode-se afirmar que os trabalhos futuros definidos e a continuação de interação e elaboração com os investigadores de diferentes domínios científicos vão continuar a trazer benefícios, tais como o aumento da usabilidade de plataforma Dendro, de forma a incluir os investigadores progressivamente em atividades de gestão de dados de investigação.

## Referências bibliográficas:

AKMON, D. *et al.* (2011) – The application of archival concepts to a data-intensive environment: working with scientists to understand data management and preservation needs. *Archival Science* [Em linha]. Vol.11, N°3, p.329-348. [Consult. 23 jul. 2016]. Disponível na Internet: <URL: <http://link.springer.com/article/10.1007/s10502-011-9151-4>>.

ALKHATTABI, M.; NEAGU, D.; CULLEN, A. (2010) – Information quality framework for e-learning systems. *Knowledge Management & E-Learning: An International Journal* [Em linha]. Vol.2, N°4, p.340-362. [Consult. 20 jul. 2016]. Disponível na Internet: <URL: <http://www.kmel-journal.org/ojs/index.php/online-publication/article/view/21/62>>.

BERMUDEZ, L. *et al.* (2011) – The importance of controlled vocabularies. *The MMI Guides: Navigating the World of Marine Metadata* [Em linha]. [Consult. 19 abr 2016]. Disponível na Internet: < <https://marinemetadata.org/guides/vocabs/vocimportance> >.

BORGMAN, C.L.; WALLIS, J.C.; Enyedy, N. (2007) – Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries. *International Journal on Digital Libraries* [Em linha]. Vol.7, N°1, p.17-30. [Consult. 20 abr. 2016]. Disponível na Internet: <URL:

<https://pages.gseis.ucla.edu/faculty/enyedy/assets/Projects/Teaching%20and%20Learning%20Science/borgman%20enyedy.pdf>>.

BRUCE, Thomas R.; HILLMANN, Diane I. (2004) – The Continuum of Metadata Quality: Defining, Expressing, Exploiting. *Metadata in Practice* [Em linha]. [Consult. 24 jul. 2016]. Disponível na Internet: <URL: <https://ecommons.cornell.edu/handle/1813/7895> >.

CASTRO, João Aguiar *et al.* (2015) – Ontologies for research data description: a design process applied to vehicle simulation. *Proceedings of the 9th Metadata and Semantics Research Conference (MTR 2015)*, CCIS 544, pp.348-354

CORONADO, Sherri.; FRAGOSO, Gilberto. (2004) – Enterprise Vocabulary Development in Protege/OWL: Workflow and Concept History Requirements NCI Center for Bioinformatics. [Em linha]. [Consult. 09 abr. 2016]. Disponível na Internet: <URL: <http://protege.stanford.edu/conference/2004/abstracts/DeCoronado.pdf>>.

DATAMARTIST.COM, Using regular expressions to check data quality. Part 2. James Standen. [Em linha]. [Consult. 21 jun. 2016]. Disponível na Internet: <URL: <http://www.datamartist.com/how-to-use-regular-expressions-to-check-data-quality-part-2> >.

FRIEDL, Jeffrey E. F. (2006) – *Mastering Regular Expressions*. 3<sup>rd</sup> Edition. O'Reilly Media, Inc., 544p. ISBN 978-0-596-52812-6

GOLBREICH, Christine; ZHANG, Songmao; BODENREIDER, Olivier (2006) – The Foundational Model of Anatomy in OWL: Experience and Perspectives. *Web Semantics: Science, Services and Agents on the World Wide Web* [Em linha] Vol.4, N°3, p.181-195. [Consult. 06 abr. 2016]. Disponível na Internet: <URL:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.9032&rep=rep1&type=pdf>>.

GRIMALOVSKII, Alexandr (2013) – Expressões regulares [Em linha].[Consult. 30 abr. 2016]. Disponível na Internet: <URL: <http://www.codenet.ru/webmast/php/regexps.php> >.

HARPING, P. (2010) – *Introduction to controlled Vocabularies : Terminology for Art, Architecture, and Other Cultural Works*. [Em linha]. 1st ed. Los Angeles: Getty Research Institute. [Consult. 14 mar 2016]. Disponível na Internet: <URL: <http://d2aohiyo3d3idm.cloudfront.net/publications/virtuallibrary/160606018X.pdf>>. ISBN 978-1-60606-018-6.

HEDDEN, H. (2010) – Taxonomies and Controlled Vocabularies Best Practices for Metadata. *Journal of Digital Asset Management* [Em linha] Vol.6, N°5, p.279-284. [Consult. 30 abr. 2016]. Disponível na Internet: < URL: <http://link.springer.com/article/10.1057%2Fdam.2010.29>>.

HEIDORN, P.Bryan (2008) – Shedding light on the dark data in the long tail of science. *Library Trends* [Em linha] Vol. 57, N°2, p.280-299. [Consult. 20 set. 2016]. Disponível na Internet: < URL: [https://www.researchgate.net/publication/49175975\\_Shedding\\_Light\\_on\\_the\\_Dark\\_Data\\_in\\_the\\_Long\\_Tail\\_of\\_Science](https://www.researchgate.net/publication/49175975_Shedding_Light_on_the_Dark_Data_in_the_Long_Tail_of_Science) >.

KARIMOVA, Yulia (2016) – *Vocabulários controlados na descrição de dados de investigação no Dendro*. Porto: Faculdade de Engenharia da Universidade do Porto. 120p. Dissertação do mestrado. Disponível na Internet: < URL: <https://repositorio-aberto.up.pt/bitstream/10216/85221/2/140750.pdf> >.

LYNDA.COM. Using regular expressions. Kevin Skoglund. [Em linha]. [Consult. 02 abr. 2016]. Disponível na Internet: <URL: <http://www.lynda.com/Regular-Expressions-tutorials/Using-Regular-Expressions/85870-2.html> >.

MARINE METADATA INTEROPERABILITY. The Importance of Controlled Vocabularies. [Em linha]. [Consult. 03 jun. 2016]. Disponível na Internet: <URL: <https://marinemetadata.org/guides/vocabs/vocimportance> >.

MOREIRA, Bárbara L. *et al.* (2009) – Automatic evaluation of digital libraries with 5SQual. *Journal of Informetrics*. [Em linha] Vol.3:, N°2, p.102-123. [Consult. 06 abr. 2016]. Disponível na Internet: <URL: <http://www.sciencedirect.com/science/article/pii/S1751157708000734>>.

NATIONAL INFORMATION STANDARDS ORGANIZATION (2005) – *ANSI/NISO Z39.19-2005: Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* [Em linha]. NISO: Maryland, USA [Consult. 08 jun 2016]. Disponível na Internet: < URL: [http://www.niso.org/apps/group\\_public/download.php/12591/z39-19-2005r2010.pdf](http://www.niso.org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf)>. ISBN: 978-1-937522-22-3.

OCHOA, Xavier; DUVAL, Erik (2006) – Quality Metrics for Learning Object Metadata. *World Conference on Educational Multimedia, Hypermedia and Telecommunications*. [Em linha]. [Consult. 06 abr. 2016]. Disponível na Internet: <URL: [https://www.researchgate.net/publication/254358241\\_Quality\\_Metrics\\_for\\_Learning\\_Object\\_Metadata](https://www.researchgate.net/publication/254358241_Quality_Metrics_for_Learning_Object_Metadata)>.

PALAVITSINIS, Nikos (2013) – *Metadata Quality Issues in Learning Repositories*. Universidad

de Alcalá: Departamento de Ciencias de la Computación. 295p. Tese de doutoramento. Disponível na Internet: <URL:

<http://dspace.uah.es/dspace/bitstream/handle/10017/20664/Thesis%20Palavitsinis.pdf?sequence=1&isAllowed=y>>.

SILVA, João Rocha Da *et al.* (2014) – Dendro: Collaborative Research Data Management Built on Linked Open Data. *The Semantic Web: ESWC 2014 Satellite Events*. p.483–487

SMITH, K.; SELIGMAN, L.; SWARUP, V. (2008) – Everybody Share: The challenge of data-sharing systems. *ComputerComputer*. [Em linha] Vol.41, p.54–61. [Consult. 06 abr. 2016]. Disponível na Internet: <URL:

<http://ieeexplore.ieee.org/document/4623223/?reload=true&arnumber=4623223>>.

STVILIA, Besiki *et al.* (2007) – A framework for Information Quality Assessment. *Journal of the American Society for Information Science and Technology*. [Em linha] Vol.58, N°12, p.1720–1733. [Consult. 06 abr. 2016]. Disponível na Internet: <URL: [http://myweb.fsu.edu/bstvilia/papers/stvilia\\_IQFramework\\_p.pdf](http://myweb.fsu.edu/bstvilia/papers/stvilia_IQFramework_p.pdf)>.

SWAN, Alma; BROWN, Sheridan (2008) – To share or not to share: Publication and quality assurance of research data outputs. *Report commissioned by the Research Information Network*. [Em linha] June, Vol.56. [Consult. 06 abr. 2016]. Disponível na Internet: <URL: <http://eprints.soton.ac.uk/266742/>>.

W3 RECOMMENDATION. Annotations, ontology header, imports and version information. [Em linha]. [Consult. 07 jun. 2016]. Disponível na Internet: <URL: <https://www.w3.org/TR/owl-ref/#Header>>.

ZHANG, Tao; MARON, Deborah J.; CHARLES, Christopher C. (2013) – Usability Evaluation of a Research Repository and Collaboration Web Site. *Journal of Web Librarianship*. [Em linha]. Vol.7, N°1, p.58–82 [Consult. 06 abr. 2016]. Disponível na Internet: <URL: [http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1061&context=lib\\_fsdocs](http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1061&context=lib_fsdocs)>.

---

## Projeto TAIL—Gestão de dados de investigação da produção ao depósito e à partilha (resultados preliminares)

---

Cristina Ribeiro

INESC TEC—Faculdade de Engenharia da Universidade do Porto

[mcr@fe.up.pt](mailto:mcr@fe.up.pt)

João Rocha da Silva

INESC TEC—Faculdade de Engenharia da Universidade do Porto

[joaorosilva@gmail.com](mailto:joaorosilva@gmail.com)

João Aguiar Castro

INESC TEC—Faculdade de Engenharia da Universidade do Porto

[joaoaguiarcastro@gmail.com](mailto:joaoaguiarcastro@gmail.com)

Ricardo Carvalho Amorim

INESC TEC—Faculdade de Engenharia da Universidade do Porto

[ricardo.amorim3@gmail.com](mailto:ricardo.amorim3@gmail.com)

João Correia Lopes

INESC TEC—Faculdade de Engenharia da Universidade do Porto

[jlopes@fe.up.pt](mailto:jlopes@fe.up.pt)

### Resumo

A gestão dos dados de investigação preocupa neste momento tanto os investigadores como os responsáveis por gestão de ciência e as agências de financiamento. Os investigadores têm consciência do valor dos dados de investigação, enquanto as agências de financiamento estabelecem mandatos para planos de curadoria e partilha de dados como parte dos seus regulamentos. Os responsáveis por políticas de ciência querem também garantir que os resultados obtidos com os seus planos de investimento têm o maior impacto possível. O projeto TAIL, a decorrer de 2016 a 2019, vai construir um portfólio de exemplos de gestão de dados em diversos



domínios que poderão ser usados pelos investigadores para avaliar o esforço requerido e as compensações a obter com esta atividade. O projeto tem como base o trabalho realizado no estudo dos fluxos de trabalho dos investigadores usando a plataforma Dendro e as interfaces móveis para a recolha de dados e metadados de que é exemplo o LabTablet. Estes resultados preliminares informam os processos a usar na publicação de conjuntos de dados existentes em repositórios nacionais e internacionais, no desenho de modelos de metadados para a descrição pormenorizada dos domínios e no alinhamento com as infraestruturas europeias e nacionais.

**Palavras-chave:** Gestão de dados de investigação; repositórios de dados; modelos de metadados; infraestruturas de investigação; cadernos de laboratório eletrónicos

## **The TAIL project: Research data management from creation to deposit and sharing (preliminary results)**

### **Abstract**

Research data management is currently a concern for researchers, science managers and funding agencies. Researchers are aware of the value of research data, while funding agencies have data curation and data sharing plans as part of their mandates. Policy agents are committed to obtain the best possible impact for the results from their funding. The TAIL project, running from 2016 to 2019, will build a portfolio of data management cases in several domains, to be used by researchers to assess the effort required by this activity and the rewards it may bring. The project is based on the work in research data management workflows using the Dendro platform and mobile platforms for data and metadata collection such as the LabTablet. These preliminary results anticipate the processes to use in the publication of existing datasets in national and international repositories, in the design of metadata models for detailed domain-specific description, and in the alignment with national and European research infrastructures.

**Keywords:** Research data management, data repositories, metadata models, research infrastructures; electronic laboratory notebooks

### **Introdução**

A gestão dos dados de investigação (designação internacional: Research Data Management, ou RDM) preocupa neste momento tanto os investigadores como as agências de financiamento e os responsáveis por políticas de ciência. Do lado dos investigadores, há a consciência do valor dos dados criados ou reunidos em contexto de investigação, especialmente quando começa a ser notado que os artigos com dados associados têm mais citações e que a divulgação de dados interessantes atrai a colaboração. As agências de financiamento estão a estabelecer os mandatos para planos de curadoria e partilha de dados como parte dos seus regulamentos e, tanto a nível nacional como internacional, os

responsáveis por políticas de ciência querem garantir que os resultados obtidos com os seus planos de investimento têm o maior impacto possível.

A gestão dos dados ainda não está estabelecida como parte integrante dos projetos de investigação. Este processo está mais adiantado em áreas como as ciências da vida, em que projetos como o do genoma humano foram construídos sobre bases de dados dedicadas desde o início. Os grupos pequenos na chamada cauda longa da ciência não têm ainda a tecnologia para gerir os seus dados, mas têm potencial para gerar grandes quantidades de dados valiosos e únicos.

Os resultados em RDM requerem a consideração de questões tecnológicas e conceptuais. Os problemas tecnológicos têm a ver com armazenamento, mas também com as interfaces para organizar, descrever e depositar os dados. Os problemas conceptuais têm a ver com a descrição. Os dados armazenados numa infraestrutura, por muito sólida que seja, serão inúteis sem a organização e descrição que só os criadores podem fornecer e que permitirá a outros especialistas do domínio reutilizá-los.

Durante 3 anos, o projeto TAIL irá fazer a ponte entre os mandatos de dados com que os grupos de investigação se confrontam e as ferramentas, fluxos de trabalho e modelos de descrição disponíveis. No final teremos um conjunto de grupos de investigação que passaram pela experiência de gerir com sucesso os dados que criaram e que também colheram os benefícios de terem os seus dados publicados, pesquisáveis e citados. Estas histórias de sucesso faltam na comunidade e as suas lições são essenciais para o progresso na gestão dos dados de investigação. Ao longo do projeto, haverá ações promovidas em colaboração pelas universidades participantes, para acompanhar o teste das ferramentas adotadas tendo em vista o seu uso em serviços estáveis de suporte aos investigadores.

Nesta apresentação mostraremos os resultados preliminares do projeto TAIL, nomeadamente os resultados já obtidos nas seguintes tarefas:

1. Fluxos de trabalhos dos investigadores usando a plataforma Dendro: ao alargar o número de grupos usando o Dendro e ao observá-los nas suas tarefas habituais, estamos a recolher informação importante para o aperfeiçoamento das tarefas de organização e descrição de dados.

2. Interfaces móveis para a recolha de dados e metadados: o LabTablet é uma aplicação móvel que permite a recolha de metadados de forma expedita e a geração automática de valores para alguns descritores usando sensores disponíveis no dispositivo móvel. Estamos a expandir o âmbito do LabTablet, usando-o também para adquirir dados que são colecionados e agregados no Dendro.

3. Publicação de conjuntos de dados existentes: nos grupos de investigação existem dados cujo valor e potencial para reutilização são reconhecidos. Para estes dados

estamos a desenvolver ações de publicação em repositórios internacionais, tendo em vista a avaliação do seu percurso em termos de pesquisas, descargas e reutilização.

4. Modelos de metadados: a atividade de descrição é tanto mais efetiva quanto mais os metadados criados tenham um bom reconhecimento nas comunidades. O método que propomos, criando ontologias no Dendro, vai ser testado em domínios com normas bem estabelecidas e em outros que não as têm ainda.

5. Alinhamento com as infraestruturas europeias: nos parceiros do TAIL temos grupos diretamente envolvidos nas iniciativas para infraestruturas especializadas por domínios. Os requisitos identificados por esses grupos são uma fonte importante para o projeto.

Estes são resultados de trabalho em curso e a sua evolução nos primeiros meses deste projeto permite uma reflexão preliminar sobre o estado das tecnologias que podem facilitar o avanço da gestão de dados de investigação. Para além disso vai permitir aferir tanto a recetividade dos grupos de investigação às novas tarefas com que são confrontados como a sua perceção sobre os ganhos a obter com a conformidade a uma política de gestão de dados centrada nos investigadores.

### **Gestão de dados de investigação**

Os dados de investigação são criados e usados em contextos diversos. Podem ser gerados especificamente para um projeto de investigação, como dados de sensores captados numa experiência, ou entrevistas para uma análise. Mas também incluem dados recolhidos sistematicamente para algum fim e que também são usados em investigação, como dados meteorológicos ou os registos de acesso a um serviço de computação. Os dados podem ser documentos comuns, como um conjunto de páginas web, colecionadas ad-hoc para avaliar o desempenho de uma ferramenta de pesquisa. Esta diversidade torna a gestão de dados de investigação uma tarefa difícil de definir, para a qual os investigadores não têm processos bem estabelecidos nem uma visão clara sobre a sua utilidade.

Muitos dados produzidos em grandes projetos são curados em infraestruturas disciplinares. O NCBI nas ciências da vida e o ICPSR nas ciências sociais são exemplos de infraestruturas maduras que suportam as tarefas de curadoria de dados de comunidades bem estabelecidas. Estas infraestruturas são usadas pelos investigadores tanto para pesquisar e obter dados como para contribuir com novos resultados de investigação.

Na chamada “cauda longa” da ciência, onde muitos grupos de pequena dimensão desenvolvem uma parte substancial do trabalho da comunidade, não existe ainda uma solução bem estabelecida para a tornar os dados visíveis, descritos de forma satisfatória, depositados e pesquisáveis. Os esforços nesta área estão ainda ao nível de projeto, como ilustrado por duas iniciativas com financiamento europeu, o EUDAT (GENTZSCHLECARPENTIER e WITTENBURG, 2014, LECARPENTIERMICHELINI e WITTENBURG,

2013) e o OpenAIRE (MANGHI et al., 2012). Ambos têm trabalho substancial no desenho de serviços e no envolvimento com grupos da cauda longa, mas as infraestruturas que propõem são ainda ligadas a projetos.

O trabalho que desenvolvemos tem como objetivo promover a publicação de dados em instituições de investigação, partindo de dois pressupostos: 1) os investigadores são os atores principais, pelo que os fluxos de trabalho devem ser simplificados do seu ponto de vista; e 2) as ferramentas multidisciplinares podem ser desenhadas e adaptadas para áreas específicas com um esforço razoável. Começámos com grupos pequenos com dados interessantes e sem muito tempo ou financiamento para curadoria de dados. Nos casos tratados os investigadores reconhecem os benefícios da gestão de dados, envolvem-se no processo de preparação de dados para publicação e obtêm resultados verificáveis da sua investigação. Procuramos o equilíbrio entre o uso de ferramentas genéricas e a satisfação dos requisitos de cada grupo de investigação.

As ferramentas disponíveis para os investigadores são decisivas para a sua motivação para a gestão de dados. É de esperar que as ferramentas que simplificam o trabalho necessário na gestão de dados e produzem resultados claros sejam mais facilmente adotadas (RIBEIRO et al., 2015). Compreendendo isso, demos especial atenção à preparação dos dados, nomeadamente à organização dos conjuntos de dados e à recolha de metadados. O nosso objetivo foi reduzir o atraso entre a captura dos dados e a sua organização e descrição, que é determinante na perda de dados valiosos.

As experiências que realizámos foram apoiadas num conjunto de grupos em diversas áreas cujos investigadores se comprometeram a realizar algumas tarefas de gestão de dados. Os contactos preliminares em cada grupo identificaram a natureza e objetivos dos seus dados, enquanto se apresentaram aos investigadores os conceitos da gestão de dados. O método usado a seguir adotou as recomendações da literatura (MAYERNIK, 2011, TENOPIR et al., 2011). Fez-se a seleção de modelos de dados específicos dos domínios, a identificação de conjuntos de dados relevantes para publicação e a avaliação de ferramentas de gestão de dados no contexto das atividades regulares dos grupos (DAF, 2016).

O trabalho que descrevemos adotou uma estratégia minimalista para a cauda longa da ciência partindo de alguns pressupostos e tendo em vista responder a uma questão. As suposições foram que os investigadores na cauda longa estão geralmente pouco informados sobre gestão de dados e não têm serviços que os apoiem na preparação de dados para depósito e publicação. A questão é se as ferramentas podem ajudar a integrar a gestão de dados no processo de investigação, resultando em mais dados chegarem à fase de depósito.

Para estes grupos na cauda longa, os nossos resultados são uma plataforma *open source* para gestão de dados de investigação (Dendro), um fluxo de trabalho que inclui a análise de conceitos do domínio e a sua organização em ontologias, a criação automática de

metadados com dispositivos móveis e a avaliação das ferramentas e dos fluxos de trabalho com investigadores de vários domínios.

É importante reunir evidência relativamente a esta proposta de fluxos de trabalho e ferramentas genéricos. Para esse efeito estamos a expandir as experiências para um público mais vasto, tendo em conta as políticas e prática nas instituições de investigação e a possibilidade de ligação a repositórios internacionais com capacidade de preservação. Considerando que os conceitos de gestão de dados estão a evoluir, bem como as recomendações em várias disciplinas, temos também em vista a ligação entre as soluções para a cauda longa e as que estão a aparecer em ambientes disciplinares. Soluções testadas em muitos domínios podem ser valiosas em projetos disciplinares, identificando conceitos e requisitos comuns. Inversamente, os requisitos de projetos de curadoria disciplinares podem revelar-se suficientemente genéricos para serem incorporados nas plataformas mais ligeiras desenvolvidas para a cauda longa.

### **Atores na gestão de dados de investigação**

A gestão de dados ainda procura o modelo certo para se embeber nos fluxos de trabalho de investigação. Vários atores estão identificados, mas para alguns deles o caminho a seguir não é claro e as relações com outros precisam de ser consolidadas.

O nosso contexto é o de uma unidade de investigação grande, com um leque grande de disciplinas. O primeiro ator é a própria universidade, que tem de lidar com um grande número de problemas relacionados com dados, nomeadamente direitos, ética, armazenamento e auditoria.

No nível imediatamente a seguir estão os departamentos e grupos de investigação, que partilham as preocupações com os dados e na prática têm de fornecer as soluções operacionais para armazenamento, acesso e auditoria.

Os investigadores vêm a seguir e são os atores principais, com os papéis de criadores e utilizadores. Um investigador pode gerar ou recolher um conjunto de dados, partilhá-lo com outro investigador num projeto colaborativo e espera-se que prossiga para a publicação de dados em repositórios e sua associação a artigos publicados.

Os atores seguintes são os gestores de investigação. O seu trabalho pode ser distanciado dos investigadores se as suas preocupações principais forem as oportunidades de financiamento e a apresentação de resultados de projetos. Cada vez mais, porém, vemos gestores de investigação que são especialistas do domínio a participar na gestão de todos os resultados de investigação, incluindo os dados.

As agências de financiamento são um ator muito influente. Adotando a visão da ciência aberta, em particular para a ciência financiada com fundos públicos, as agências de

financiamento estão cada vez mais a requerer que os resultados de investigação, incluindo os dados, estejam disponíveis publicamente para outros grupos os usarem (EOSC, 2016). Esta atitude tem como consequência tornar os resultados publicados mais facilmente verificáveis e reutilizáveis.

Dado que estamos ainda longe de estar generalizado o acesso a infraestruturas de gestão de dados, os desenvolvedores de aplicações são também atores importantes. Precisam de familiaridade com os conceitos de gestão de dados, eles próprios ainda a evoluir, e criar boas metáforas para a ligação entre pessoas e tecnologia.

Os resultados de investigação tradicionais, como artigos de revista e relatórios de projetos, têm uma prática bem estabelecida de disseminação, envolvendo revistas, editoras e mais recentemente repositórios institucionais, que fornecem visibilidade e possibilidades de referência cruzada. Tudo isto é incipiente na publicação de dados. Os agregadores de dados e os coletores de metadados são também chamados como atores.

A preservação é uma dimensão importante na gestão de dados, pelo que as capacidades e características dos arquivos têm de estar de acordo com os requisitos de preservação (CAPLAN, 2009, COUNCIL OF THE CONSULTATIVE COMMITTEE FOR SPACE DATA, 2002). Os arquivos, institucionais ou disciplinares, são por isso atores importantes, estando o seu papel ainda mal estabelecido em muitas áreas. Os arquivos de dados podem ser muito volumosos, pelo que a seleção e a descrição apropriada se tornam essenciais. Estas funções são consumidoras de tempo, mas está provado que se tornam mais comportáveis quando realizadas ao tempo da geração dos dados, e com ferramentas apropriadas, do que mais tarde no processo. A seleção e descrição passam aqui a ser tarefas embebidas no processo de gestão de dados desde o início, em vez de serem tarefas à posteriori.

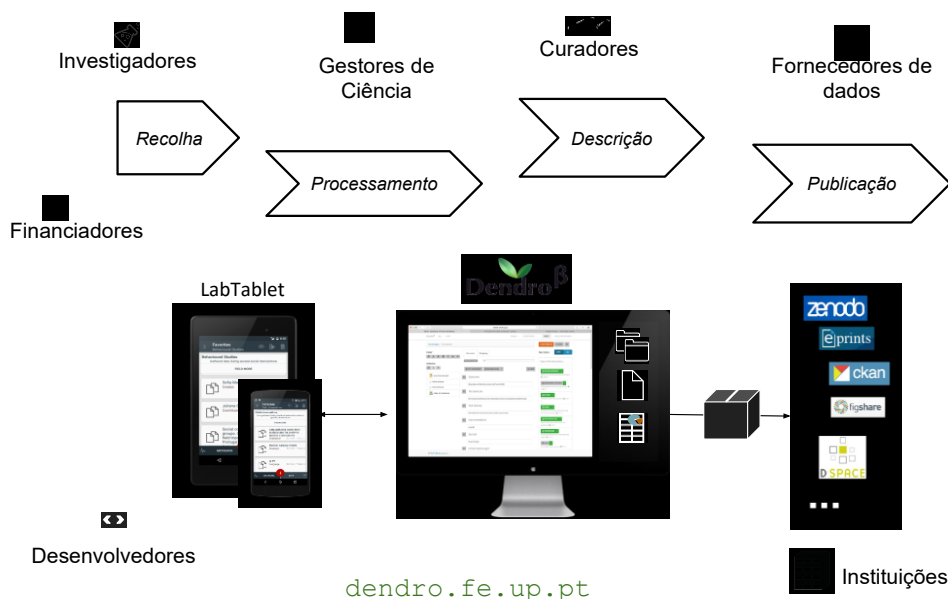


FIGURA 1. ATORES NA GESTÃO DE DADOS DE INVESTIGAÇÃO

## Modelos adaptados à realidade dos grupos

A gestão de dados inclui atividades com diferentes âmbitos num grupo de investigação. Num extremo, a gestão de dados de investigação está fortemente ligada com a publicação: os investigadores colocam um pouco de esforço extra na preparação de materiais adicionais—dados recolhidos, módulos de software, versões digitais de dados apresentados num artigo—para atrair mais atenção sobre os seus resultados, para conformar ao mandato de uma agência de financiamento ou para satisfazer os requisitos de uma revista. No outro extremo, a gestão de dados pode fazer parte de um ambiente de *e-science* muito completo, em que as atividades do dia a dia são registadas, os dados são gerados e explorados de forma colaborativa e os resultados de investigação são preparados.

O trabalho realizado começou com um estudo de avaliação na Universidade do Porto, com a colaboração de 8 grupos de investigação, usando recomendações existentes e cobrindo aspetos como a consciência da necessidade de curadoria de dados, as necessidades prementes relativas a dados existentes, as soluções correntes para armazenamento, o valor atribuído a dados legados e as necessidades de apoio em ações de gestão de dados (RIBEIRO e FERNANDES, 2011). Como resultado deste estudo, o trabalho começou com base na seguinte hipótese. Os repositórios institucionais cresceram e criaram uma grande comunidade, que promoveu o auto-depósito e o acesso aberto, devido à disponibilidade de plataformas de repositórios muito sólidas, como DSpace, Fedora e EPrints. Para a gestão de dados de investigação se tornar parte do fluxo de trabalho num grupo de investigação, precisamos também de ferramentas convenientes. Para fazer a comunidade adotar estas ferramentas, elas devem permitir um benefício imediato aos investigadores, e ainda criar uma expectativa de melhoria na eficácia da publicação e disseminação da sua investigação (AMORIM et al., 2016).

A preocupação com a gestão de dados na cauda longa é relativamente recente e diversas entidades fornecem apoio ao depósito de dados multidisciplinares (ANDS, 2016, DANS, 2016, DASH, 2016, DATAONE, 2016, DCC, 2016). No entanto, nas áreas em que conjuntos de dados grandes estão no centro da investigação há um registo mais longo de iniciativas de gestão de dados, como sejam bases de dados que identificam contribuições individuais e têm uma forte ligação às publicações. O NCBI e o ICPSR são exemplos disto nas ciências de vida e nas ciências sociais, respetivamente (ICPSR, 2016, NCBI RESOURCE COORDINATORS, 2013). Noutras áreas, à medida que crescem os grupos internacionais, a partilha de dados se torna mais importante e a avaliação da investigação requer a ligação a fontes de dados, há novas infraestruturas internacionais a serem montadas. Um bom exemplo são os projetos promovidos pelo ESFRI, o *European Strategy Forum on Research*

*Infrastructures*, que está a apoiar uma rede de infraestruturas de investigação de interesse europeu a longo prazo. O ESFRI suporta correntemente 29 projetos em fase de implementação e 21 em projeto, nas grandes áreas da Energia, Ambiente, Saúde e Alimentação, Ciências Físicas e Engenharia, Inovação Cultural e Social e e–infraestruturas.

### **Organização de dados e criação de metadados com Dendro e LabTablet**

A adoção da gestão de dados depende da existência de processos claros para os investigadores relativamente a recolha de dados, organização, descrição e publicação ou depósito. Ao entrevistarmos investigadores acerca das práticas de gestão de dados, tornou-se claro o fosso entre os processos usados na preparação de artigos e comunicações (embora dependentes do domínio) e as rotinas necessárias para organizar dados e para os tratar como resultados de investigação. O nosso trabalho concentrou-se por isso em duas frentes: o contacto com os investigadores, identificando tarefas que os possam levar a comprometer-se com a gestão de dados, e o desenho e implementação de ferramentas para suportar as suas atividades. As ferramentas de gestão de dados têm um campo de aplicação largo, incluindo a recolha, limpeza, processamento, organização, descrição, depósito, pesquisa. À medida que a gestão de dados se estabelece, espera-se que as plataformas de repositórios, disciplinares ou genéricas, se tornem comuns. As ferramentas de recolha e processamento serão provavelmente dependentes do domínio. Escolhemos por isso investir nas tarefas de organização e descrição de dados, na interface entre os investigadores e os gestores de repositórios (CASTROROCHA DA SILVA e RIBEIRO, 2014, CASTRO et al., 2015, CASTROROCHA DA SILVA e RIBEIRO, 2013, DA SILVA et al., 2014). As ferramentas de preparação e descrição de conjuntos de dados contribuem para um discurso consistente sobre gestão de dados com os investigadores, e uma mesma ferramenta pode fazer a ligação a diversos repositórios. A descrição completa e apropriada ao domínio é essencial para a pesquisa, uma vez que muitos dados têm pouco ou nenhum conteúdo indexável.

O Dendro é uma plataforma de gestão de dados baseada em ontologias e desenvolvimento na Universidade do Porto e recentemente publicada em código aberto (ROCHA DA SILVA, 2016, ROCHA DA SILVA et al., 2014). Os investigadores são o seu público, e o objetivo é ajudá-los a depositar e partilhar dados tanto dentro do seu grupo de investigação como com elementos externos, de forma controlada. O Dendro supõe a disponibilidade de repositórios para a preservação a longo prazo e concentra-se na organização e descrição dos dados, usando a metáfora da *Dropbox* e oferecendo possibilidades de descrição sofisticadas. No Dendro um grupo de investigação cria “projetos”, na forma de pastas partilhadas em que cada colaborador deposita ficheiros, cria pastas e descreve os seus recursos com descritores genéricos ou dependentes do domínio. A



orientação para a publicação revela-se na exportação de dados e metadados para as principais plataformas de repositórios nos momentos escolhidos pelos autores.

Além de suportar a criação colaborativa de metadados, o Dendro ajuda os investigadores a escolher os conjuntos de descritores que são mais relevantes para os dados do seu domínio. Duas componentes contribuem para isso: um modelo de dados flexível baseado em ontologias que cresce à medida que se juntam ontologias específicas dos domínios e um sistema de ordenação de descritores que aprende com as interações passadas na plataforma. O modelo de dados permite exportar registos como dados abertos ligados (*Linked Open Data*) para interoperabilidade, mas vai para além das soluções correntes como o OAI-PMH nas facilidades de interrogação e recuperação. O Dendro permite ligar a produção de dados com a sua preservação a longo prazo de acordo com um plano de gestão de dados, por exemplo.

Os dispositivos móveis têm evoluído para incluir funcionalidades avançadas que os tornam apropriados para diversas tarefas no âmbito da investigação. Além da capacidade de armazenamento, estes dispositivos podem estar ligados à Internet em permanência e incluem sensores capazes de fornecer informação sobre o contexto do utilizador.

O LabTablet é uma ferramenta na linha dos cadernos de laboratório electrónicos, aplicações que tiram partido dos sensores no dispositivo móvel para ajudar os investigadores a descrever os seus dados (AMORIM et al., 2015, AMORIM et al., 2014). Nalguns casos a descrição com o dispositivo móvel substitui o processo realizado em papel, com uma maior garantia da associação entre dados e metadados. O dispositivo móvel contribui para a integração de dados e metadados desde o início de um projeto, distribuindo o esforço de criação de metadados pela sua duração. Isto contribui para evitar que a descrição seja uma tarefa consumidora de tempo e realizada numa altura em que nem todos os pormenores do processo estão presentes.

Os metadados também são valiosos antes do depósito dos dados: os parceiros de um projeto colaboram e precisam de trocar dados; se estes têm descrições associadas, é menos provável que sejam mal interpretados e a colaboração fica facilitada. O resultado desta abordagem é que os conjuntos de dados têm um registo de metadados quando o projeto acaba. Isto garante que os investigadores são envolvidos desde o início e que não será necessário muito esforço para depositar ou publicar dados no final do projeto.

### **Fluxo de trabalho dos curadores usando ontologias ligeiras**

Na história recente da curadoria e publicação de dados, vemos duas linhas de evolução nos modelos de metadados. O primeiro considera que os dados de investigação

são basicamente um resultado de investigação tal como os artigos, e que as descrições têm de seguir os modelos bem estabelecidos para publicações. Esta vista tem o mérito de facilitar a abordagem aos investigadores, tornando o depósito de dados semelhante ao das publicações, e promovendo repositórios em que as publicações e os dados são apenas distinguidos pelo tipo de recurso. A desvantagem é que não se investe na descrição de dados e em tornar os dados compreensíveis fora do grupo de investigação onde foram criados; um investigador que pretenda reutilizar os dados terá provavelmente de contactar os criadores para perceber os dados. Esta abordagem é seguida em iniciativas tais como o projeto OpenAIRE, promovendo o repositório Zenodo para recolher e interligar artigos, conjuntos de dados, software e os projetos em que tiveram origem (OPENAIRE, 2016).

A segunda linha, onde se enquadra este trabalho, parte da hipótese de que os grupos de investigação podem construir competências para descrever os seus dados com descritores mais expressivos. Isto requer ferramentas, tal como as que desenvolvemos, mas também um fluxo de curadoria com algum suporte especializado. Na nossa abordagem, isto é capturado como o fluxo do curador e envolve a análise dos conceitos essenciais para o domínio, a sua representação numa ontologia, a incorporação da ontologia na plataforma de gestão de dados e a publicação da própria ontologia. A evolução da descrição baseada nas ontologias também favorece a identificação, dentro de um domínio, de descritores apropriados que podem evoluir no sentido da adoção na comunidade, ou mesmo da normalização.

### **Conclusões e trabalho em curso**

O projeto TAIL vai desenvolver-se numa época em que a gestão de dados de investigação está na ordem do dia, com muitas iniciativas internacionais e uma crescente consciência da sua importância nas várias comunidades científicas. A equipa do projeto parte de um ponto em que existe trabalho feito na identificação dos requisitos dos investigadores, no desenvolvimento de ferramentas e no seu teste em condições reais. O trabalho em curso vai desenvolver-se em três linhas. Na primeira vamos continuar o trabalho junto dos investigadores, aumentando o número de grupos envolvidos, experimentando com repositórios internos às instituições e com outros externos e construindo uma base de casos que possam servir de referência a outros. Na segunda vamos trabalhar junto de algumas infraestruturas do roteiro nacional, identificando as necessidades de comunidades muito estruturadas e estudando soluções adaptadas. Na terceira vamos continuar o trabalho sobre modelos de metadados e processos para facilitar a sua definição, além de desenvolver métricas para avaliar a qualidade dos metadados produzidos.

## Referências bibliográficas

AMORIM, Ricardo Carvalho [et al.] (2014) – LabTablet: Semantic Metadata Collection on a Multi-domain Laboratory Notebook – Springer Communications in Computer and Information Science [Em linha]. Vol: 478, nº (2014), p. 193–205. ISSN: 978-3-319-13673-8

AMORIM, Ricardo Carvalho [et al.] (2016) – A comparison of research data management platforms: architecture, flexible metadata and interoperability – Universal Access in the Information Society [Em linha]. (2016), p. 1–12. Disponível em WWW: <<http://dx.doi.org/10.1007/s10209-016-0475-y>>. ISSN: 1615-5297

AMORIM, Ricardo [et al.] (2015) – Engaging researchers in data management with LabTablet, an electronic laboratory notebook. In Symposium on Languages, Applications and Technologies, SLATE'2015. 2015.

ANDS – ANDS- Australian National Data Service [Em linha]. Disponível em WWW: <URL: <http://ands.org.au/>>.

CAPLAN, Priscilla (2009) – Understanding PREMIS: an overview of the PREMIS Data Dictionary for Preservation Metadata 2009.

CASTRO, J. A. ; ROCHA DA SILVA, J. ; RIBEIRO, Cristina (2013) – Designing an Application Profile Using Qualified Dublin Core: A Case Study with Fracture Mechanics Datasets. In International Conference on Dublin Core and Metadata Applications. 2013. 2013. p. 47--52.

CASTRO, J. ; ROCHA DA SILVA, J. ; RIBEIRO, C. (2014) – Creating lightweight ontologies for dataset description: Practical applications in a cross-domain research data management workflow. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL and TPDF) DL 2014. 2014. ACM Press, 2014.

CASTRO, João Aguiar [et al.] (2015) – Ontologies for Research Data Description: A Design Process Applied to Vehicle Simulation. In GAROUFALLOU, Emmanouel, HARTLEY, Richard J. & GAITANOU, Panorea – Metadata and Semantics Research: 9th Research Conference, MTSR 2015. Springer, 2015. p. 348–354. ISBN: 978-3-319-24129-6.

COUNCIL OF THE CONSULTATIVE COMMITTEE FOR SPACE DATA, Systems – Reference Model for an Open Archival Information System (OAIS) 2002.

DA SILVA, João Rocha [et al.] (2014)– The Dendro research data management platform- Applying ontologies to long-term preservation in a collaborative environment. In 11th International Conference on Digital Preservation iPRES 2014 [Em linha]. iPRES, 2014, Disponível em WWW: <URL: <http://dcpapers.dublincore.org/pubs/issue/view/165>>.

DAF – Data Asset Framework [Em linha]. Disponível em WWW: <URL: <http://www.data-audit.eu/>>.

DANS – Data Archiving and Networked Services [Em linha]. Disponível em WWW: <URL: <http://www.dans.knaw.nl/en>>.

DASH – Dash– Data Sharing made easy [Em linha]. Disponível em WWW: <URL: <https://dash.cdlib.org/>>.

DATAONE – DataONE [Em linha]. Disponível em WWW: <URL: <https://www.dataone.org/>>.

DCC – DCC– Digital Curation Centre [Em linha]. Disponível em WWW: <URL: <http://www.dcc.ac.uk/>>.

EOSC – European Open Science Cloud [Em linha]. Disponível em WWW: <URL: <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>>.

GENTZSCH, W. ; LECARPENTIER, D. ; WITTENBURG, P. (2014) – Big Data in Science and the EUDAT Project. In 2014 Annual SRII Global Conference. 2014. p. 191–194.

ICPSR – Interuniversity Consortium for Political and Social Research [Em linha]. Disponível em WWW: <URL: <https://www.icpsr.umich.edu/icpsrweb/>>.

LECARPENTIER, D. ; MICHELINI, A. ; WITTENBURG, P. (2013) – The building of the EUDAT Cross–Disciplinary Data Infrastructure. In EGU General Assembly Conference Abstracts. 2013.

MANGHI, Paolo [et al.] (2012)– OpenAIREplus: the European Scholarly Communication Data Infrastructure – D–Lib Magazine [Em linha]. Vol: 18, nº 9/10 (2012), Disponível em WWW: <<http://www.dlib.org/dlib/september12/manghi/09manghi.html>>. ISSN: 1082–9873

MAYERNIK, Matthew Stephen (2011) – Metadata realities for cyberinfrastructure: Data authors as metadata creators – ProQuest Dissertations and Theses [Em linha]. (2011), p. 338.

NCBI RESOURCE COORDINATORS – Database resources of the National Center for Biotechnology Information – Nucleic Acids Research [Em linha]. Vol: 41, nº Database issue (2013), p. D8–D20. Disponível em WWW: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531099/>>. ISSN: 0305–1048

OPENAIRE – The OpenAIRE project. 2016. Disponível em WWW: <<https://www.openaire.eu/>>.

RIBEIRO, Cristina [et al.] (2015) – Motivators and Deterrents for Data Description and Publication: Preliminary Results (Short Paper). In On the Move to Meaningful Internet Systems: OTM 2015 Workshops 2015. Disponível em WWW: <URL: [http://dx.doi.org/10.1007/978-3-319-26138-6\\_55](http://dx.doi.org/10.1007/978-3-319-26138-6_55)>. p. 512–516.

RIBEIRO, Cristina ; FERNANDES, Maria Eugénia Matos (2011) – Data Curation at U.Porto: Identifying current practices across disciplinary domains – IASSIST Quarterly [Em linha]. Vol: 35, nº 4 (2011), p. 14–17.

ROCHA DA SILVA, J. (2016) – The Dendro RDM platform. 2016. Disponível em WWW: <<https://github.com/feup-infolab-rdm/dendro>>.

ROCHA DA SILVA, João [et al.] (2014) – Dendro: Collaborative Research Data Management Built on Linked Open Data. Springer International Publishing, 2014. Disponível em WWW: <[http://dx.doi.org/10.1007/978-3-319-11955-7\\_71](http://dx.doi.org/10.1007/978-3-319-11955-7_71)>.

TENOPIR, Carol [et al.] (2011)– Data Sharing by Scientists: Practices and Perceptions – PLoS ONE [Em linha]. Vol: 6, nº 6. Disponível em WWW: <<http://dx.plos.org/10.1371/journal.pone.0021101>>.

---

## Bibliotecas, repositórios institucionais e direitos autorais

---

Bianca Amaro

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

[bianca@ibict.br](mailto:bianca@ibict.br)

### Resumo

O presente trabalho tem como objetivo descrever aspectos da Lei dos Direitos Autorais brasileira relacionados com a proteção dos direitos patrimoniais do autor. Apresenta a relação existente entre as atividades realizadas nas bibliotecas e a inadequação da Lei de Direitos Autorais vigente. Trata das dificuldades que a lei autoral apresenta quanto a prática do autoarquivamento e com o gerenciamento, por parte das bibliotecas, dos repositórios institucionais da produção científica. Também são apontados exemplos do tratamento das questões autorais, a utilização da lei vigente e das licenças Creative Commons em relação aos repositórios institucionais como um todo, como também quanto às obras depositadas.

**Palavras-chave:** Direitos autorais. Biblioteca. Repositório institucional. Acesso Aberto. Creative Commons. Brasil

### Libraries, institutional repositories and copyright

#### Abstract

This paper aims to describe aspects of the Brazilian Copyright Act related to the protection of the author's rights. It shows the relationship between the activities in libraries and the inadequacy of the current Copyright Act. It deals with the difficulties that copyright law has as the practice of self-archiving and management, by libraries, institutional repositories of scientific production. Also are pointed examples of the treatment of copyright issues, the use of the existing law and the Creative Commons licenses in relation to institutional repositories as a whole, as well as to works deposited.

**Keywords:** Copyright, Library, Institutional repository, Open Access, Brazil

## Introdução

A Lei de Introdução às normas do Direito Brasileiro nos deixa claro, em seu artigo 3º que "ninguém se escusa de cumprir a lei, alegando que não a conhece". Entretanto, pode-se afirmar com segurança que este preceito legal não é seguido em vários âmbitos da sociedade. Não é raro encontrarmos setores que exercem suas atividades desconhecendo o arcabouço jurídico que lhes dizem respeito. Sem lugar a dúvidas esta é a realidade no âmbito bibliotecário, no que tange à lei dos Direitos Autorais (LDA). Seguramente esta falta de conhecimento se dá por questões de formação, ou seja, pela disciplina Direitos Autorais não estar presente nas grades curriculares das Universidades brasileiras. Entretanto Magnus (2009) ressalta que "o bibliotecário é o profissional que trabalha diretamente com o produto protegido pelo Direito Autoral, ou seja, toda e qualquer informação criada por um autor, seja ela um livro, um artigo, uma música, um filme, uma resenha, um resumo, dentre outras formas de criação".

ALDA ao tratar dos Direitos Patrimoniais do Autor e de sua Duração, Art. 28. determina que cabe ao autor o direito exclusivo de utilizar, fruir e dispor da obra literária, artística ou científica e em seu Art. 29 dispõe que depende de autorização prévia e expressa do autor a utilização da obra, por quaisquer modalidades. Esta necessidade de autorização prévia e expressa perdurará por muitos anos, mais especificamente, segundo o artigo Art. 43. "Será de setenta anos o prazo de proteção aos direitos patrimoniais sobre as obras anônimas ou pseudônimas, contado de 1º de janeiro do ano imediatamente posterior ao da primeira publicação."

De acordo com Lemos (2005 apud OLIVEIRA, 2011) o direito autoral tradicional funciona como um grande "não", pois se alguém desejar utilizar uma determinada obra tem de pedir autorização prévia ao autor ou detentor de direitos autorais". Assim sendo, este trabalho tem como objetivo analisar a lei dos Direitos Autorais brasileira à luz das atividades realizadas no âmbito de uma biblioteca e dos repositórios institucionais de acesso aberto no Brasil. O método utilizado nesta pesquisa foi a análise descritiva da LDA frente às atividades realizadas nas bibliotecas que possuem repositórios institucionais e suas práticas de gerenciamento do tema Direitos Autorais em relação às obras depositadas.

## A lei autoral brasileira

Apresentaremos abaixo uma seleção de obras que são protegidas, segundo a lei brasileira (Lei 9.610/98), e que estão diretamente relacionadas com os repositórios institucionais da produção científica:

"Art. 7º São obras intelectuais protegidas as criações do espírito, expressas por qualquer meio ou fixadas em qualquer suporte, tangível ou intangível, conhecido ou que se invente no futuro, tais como:

I – os textos de obras literárias, artísticas ou científicas;

II – as conferências, alocações, sermões e outras obras da mesma natureza;

[...]

X – os projetos, esboços e obras plásticas concernentes à geografia, engenharia, topografia, arquitetura, paisagismo, cenografia e ciência;

XI – as adaptações, traduções e outras transformações de obras originais, apresentadas como criação intelectual nova; [...]"

Conforme dito anteriormente, segundo a letra da lei, é necessária a obtenção de autorização prévia e expressa do autor para a realização das seguintes atividades: o empréstimo de livros, a realização de cópias integrais, sendo só permitida a reprodução de pequenos trechos, a realização de microfilmagens ou digitalizações, como também a inclusão de obras em bases de dados. Abaixo, também apresentamos algumas formas de uso que afetam diretamente as atividades realizadas na gestão e interação direta com os repositórios institucionais.

"Art. 29. Depende de autorização prévia e expressa do autor a utilização da obra, por quaisquer modalidades, tais como:

I – a reprodução parcial ou integral;

[...]

VI – a distribuição, quando não intrínseca ao contrato firmado pelo autor com terceiros para uso ou exploração da obra;

[...]

VIII – a utilização, direta ou indireta, da obra literária, artística ou científica, mediante:

[...]

i) emprego de sistemas óticos, fios telefônicos ou não, cabos de qualquer tipo e meios de comunicação similares que venham a ser adotados;

IX – a inclusão em base de dados, o armazenamento em computador, a microfilmagem e as demais formas de arquivamento do gênero;

X – quaisquer outras modalidades de utilização existentes ou que venham a ser inventadas."

## **A relação das Bibliotecas com os Direitos Autorais**

A LDA prevê exceções e limitações à proteção das obras. Entenda-se aqui os termos limitações e exceções como termos que possibilitam o livre uso das obras. Nestes casos não é necessária a obtenção de autorização prévia do titular dos direitos autorais para sua utilização.

Segundo Mizukami et al. (2008), "Exceptions and limitations are widely recognized as an essential part of the balance between public and private interests inherent in copyright law, deeply linked to the underlying rationales for copyright itself." Ao analisar a LDA vigente constata-se a inexistência de qualquer tipo de exceção ou limitação para o trabalho das



bibliotecas, arquivos ou museus. Uma vez que não existe nenhuma limitação ou exceção, parte do trabalho realizado nesses ambientes infringe diretamente a lei.

Ao seguir a letra da lei, ficam comprometidos os seguintes serviços tipicamente realizados pelas bibliotecas: empréstimo do material do acervo; digitalização de obras, mesmo com fins de preservação e realização de fotocópias parciais das obras e se formos mais a fundo até mesmo o uso, por vários usuários, dentro da biblioteca de uma determinada obra.

Deste fato decorrem problemas que os bibliotecários até mesmo desconhecem que têm, tanto com os suportes físicos como com os digitais. Há também grande insegurança, por parte dos bibliotecários, no tratamento dos e-books em termos contratuais e de utilização. Contudo vemos um aspecto positivo nessa dificuldade, qual seja estar levando os bibliotecários a uma maior atenção aos contratos e ao tratamento dos aspectos autorais, neles contidos. É importante frisar que o desconhecimento da lei dos Direitos Autorais por parte da comunidade bibliotecária é uma realidade. Entretanto faz-se necessário ressaltar que esse desconhecimento da LDA ocorre, fato, em todas as áreas do conhecimento.

Ocorre que apesar de ser indiscutível a importância das bibliotecas para o mundo da Ciência, dentre outros tantos âmbitos, não se tem, até o momento, um equilíbrio entre o interesse público e o interesse privado. A inexistência de limitações e exceções na Lei de Direitos Autorais brasileira é inequívoca prova disso. Faz-se extremamente necessário que a comunidade de bibliotecários, como também o público em geral, se mobilize e consiga obter uma mudança na LDA a seu favor.

### **A relação das Bibliotecas com os Direitos Autorais**

A cada dia todos os atores envolvidos com as atividades científicas, sejam os pesquisadores, os bibliotecários ou os dirigentes institucionais reconhecem os ganhos de implantar, organizar e dar maior visibilidade à produção científica nacional. No País, a prática do autoarquivamento ainda não é uma realidade. Não obstante, caso o autoarquivamento venha a se tornar uma realidade os autores pesquisadores, em sua grande maioria, também se encontrarão com problemas. Os pesquisadores para finalizarem o ciclo de uma pesquisa devem tornar os seus resultados públicos. Em grande parte das áreas do conhecimento esses resultados são publicados em revistas e livros científicos. No afã de encerrar a sua pesquisa e tê-la validada por seus pares, em geral, os pesquisadores terminam por assinar contratos de edição sem ter o conhecimento que está abrindo mão de seu direito de utilização de seu texto. Assim sendo entre os autores pesquisadores, no mundo da produção científica, a transferência dos direitos patrimoniais é uma prática comum. Desta realidade decorre o problema do autoarquivamento, pois o próprio autor deverá solicitar permissões para a realização do depósito e autorização do acesso em um repositório institucional.

A questão relacionada com a transferência de direitos está prevista na LDA, conforme artigo abaixo:

Art. 49. Os direitos de autor poderão ser total ou parcialmente transferidos a terceiros, por ele ou por seus sucessores, a título universal ou singular, pessoalmente ou por meio de representantes com poderes especiais, por meio de licenciamento, concessão, cessão ou por outros meios admitidos em Direito[...].

Desta forma abre-se a possibilidade de criar-se uma figura que passa a gozar da proteção da obra como se autor fosse. A partir do ato de transferência, até mesmo o autor criador terá que obter a autorização prévia e expressa para utilizar, fruir ou dispor da sua obra. Isto quer dizer que o autor criador abrirá mão de seus direitos patrimoniais e a toda proteção que lhe é concedida. Não obstante também permanecerem, de maneira inalienável, os seus direitos morais.

### **Serviços de auxílio ao depósito em repositórios institucionais**

No Brasil, nos dias atuais, os depósitos nos repositórios são, maiormente, efetuados pelas bibliotecas. Desta forma, o problema de obtenção de licenças junto aos titulares dos Direitos Autorais foi transferido para as bibliotecas, que são, frequentemente, as gestoras dos repositórios brasileiros. O processo de depósito, não sendo por autoarquivamento, se vê ainda dificultado porque antes de obter as licenças prévias e expressas junto aos titulares de Direitos de Autor, o bibliotecário tem a tarefa de localizar a produção científica de seus pesquisadores.

A fim de respeitar as políticas editoriais das revistas e assim não ter problemas posteriores relacionados com os Direitos Autorais, muitos gestores de têm consultado diretórios especializados no tema antes de realizar o depósito da produção científica em seus repositórios. Ao longo dos anos vem sendo criados diretórios de políticas editoriais, cujas informações são fornecidas pelos próprios editores, em níveis nacionais e internacionais. Vale ressaltar as iniciativas nacionais do Diretório de Políticas Editoriais das Revistas Científicas Brasileiras – Diadorim, do Brasil, e o serviço Blimunda, de Portugal.

O Diadorim, que nasceu de uma das iniciativas em prol do acesso aberto do Programa Brasileiro de Acesso Aberto à Informação Científica do Instituto Brasileiro de Informação em Ciência e Tecnologia, é um serviço de informações que tem como objetivo identificar, sistematizar e disponibilizar as informações das políticas estabelecidas pelas editoras dos periódicos brasileiros, referentes ao armazenamento dos artigos nos repositórios institucionais, como também a sua permissão de acesso. Dada a realidade brasileira, acima descrita, onde o depósito é realizado, com muita frequência, pelos gestores dos repositórios e não pelos autores, o Diretório Diadorim não especificou o seu usuário-alvo.

O serviço Blimunda, que nasceu no âmbito do Repositório Científico de Acesso Aberto de Portugal (RCAAP), tem como principal objetivo recolher as políticas das editoras e

revistas científicas portuguesas a fim de auxiliar os autores a realizarem o autoarquivamento em repositórios institucionais. Interessante notar que este serviço segue estritamente o proposto pela Declaração de Budapeste (BOAI), que recomenda a adoção de duas estratégias, sendo que uma delas, a via verde, que se traduz no autoarquivamento em “*open electronic archives*”, que posteriormente vieram a conformar os denominados repositórios institucionais. A informação levantada pelo serviço Blimunda é enviada ao diretório internacional SHERPA/RoMEO.

É relevante também destacar a importância do serviço SHERPA/RoMEO da Universidade de Nottingham – Inglaterra, que atingiu um alcance e referência internacional no mundo do Acesso Aberto. O serviço RoMEO apresenta a política geral dos editores de revistas e periódicos revisados por pares, que foi sugerida pelos usuários. Neste serviço é possível encontrar informações sobre as políticas editoriais de revistas de todo o mundo.

Ainda que esses mencionados serviços tenham sido criados como suporte ao depósito de documentos nos repositórios institucionais o seu uso deve ser estimulado junto aos autores para que estes conheçam as políticas editoriais das revistas de sua área e assim ter maiores condições de eleger onde publicar os seus resultados de pesquisa.

### **O uso das licenças autorais nos repositórios institucionais brasileiros**

Dada a complexidade de gerenciamento dos Direitos Autorais, por parte dos gestores dos repositórios institucionais, observaram-se duas tendências para o tratamento da questão autoral nos repositórios: a utilização de licenças Creative Commons e a declaração de “Todos os direitos reservados, conforme Lei nº 9.610” / “Os itens no repositório estão protegidos por copyright, com todos os direitos reservados, salvo quando é indicado o contrário”.

Essas tendências decorrem do nítido desconhecimento sobre Direito Autorais por parte tanto dos gestores dos repositórios, como dos autores. Os repositórios institucionais brasileiros têm sido implantados sob a égide do Acesso Aberto e seus responsáveis são conscientes de sua importância estratégica para que os resultados científicos circulem livre de barreiras. Por este motivo percebemos a falta de percepção do que as licenças chamadas “fechadas” (todos os direitos reservados) realmente significam. Como apontado anteriormente a possibilidade de que editores possam assumir o papel de autores, tornando-se titulares do direito, não é percebida nem pelos gestores e nem, em vários casos, pelos próprios autores criadores. Quando um gestor de repositório opta pela licença onde todos os direitos são reservados, na verdade, desconhece que está protegendo o titular dos Direitos de Autor e não o autor criador. Trata-se de uma decisão, em teoria, em prol dos autores que na verdade pode vir a causar aos seus usuários sérios problemas autorais, uma vez que este fará uso dos documentos presentes no repositório, acreditando que poderá usufruir de um material que é licenciado de forma tal que o seu acesso e uso estão permitidos de maneira ampla.

Os repositórios brasileiros têm, em sua maioria, adotado o uso das licenças Creative Commons (CC). O seu uso disseminado advém da prática contumaz da replicação das decisões tomadas por repositórios já criados, por parte dos gestores de novos repositórios. Não se pode dizer que se trate de uma opção consciente, pois em grande parte dos repositórios institucionais observa-se a informações inconsistentes. Há casos em que o repositório apresenta-se como gerido sob a licença CC, mas, posteriormente na descrição do item, encontramos a declaração que "Os itens no repositório estão protegidos por copyright, com todos os direitos reservados, salvo quando é indicado o contrário". São, notadamente, informações paradoxais.

Interessante também observar a descrição dos Direitos Autorais de diversas maneiras em um mesmo repositório tais como "Direitos Autorais para artigos publicados nesta revista são do autor, com direitos de primeira publicação para a revista. Em virtude de aparecerem nesta revista de acesso público, os artigos são de uso gratuito, com atribuições próprias, em aplicações educacionais e não-comerciais." e "Os leitores desta revista científica podem ler, baixar, copiar, distribuir, imprimir, buscar ou fazer um link para textos completos e usá-los para qualquer propósito legítimo, sem ônus financeiro."

A fim de obtermos um panorama nacional do uso das licenças de direitos autorais buscamos analisar um repositório por cada região, que possuem um número representativo da sua produção científica institucional.

Neste sentido, encontramos na descrição das licenças de Direitos Autorais de um repositório, uma série de possibilidades de licenciamento tais como: "Licença Padrão (nome da Instituição) – utilizada pelos materiais cujo detentor de direitos autorais patrimoniais seja a Instituição, ou seja, documentos produzidos pelos servidores no âmbito de suas atividades profissionais, ou que tenha negociado a cessão dos direitos com os autores; Autorização – licença utilizada para os casos em que o autor permanece como detentor dos direitos patrimoniais, porém tenha autorizado a sua disponibilização no repositório; Licenças Creative Commons – sempre que o objeto digital depositado possuir uma licença creative commons atribuída, esta deve ser indicada, pois já explica o uso definido pelo detentor dos direitos patrimoniais; Domínio Público – segundo a lei brasileira, as obras caem em domínio público 70 anos após a morte do autor. Tais obras podem ser utilizadas e adaptadas livremente pela população; Licença Comum – o repositório poderá ainda utilizar obras que não possuem nenhuma das licenças anteriores, mas que possuam uma licença e termos de uso próprios, que permite [s/c] sua disponibilização em repositórios de acesso aberto, todas referentes a um mesmo repositório".

Há também o caso de uso onde está indicado que "O autor é titular dos direitos autorais dos documentos disponíveis neste repositório e é vedada, nos termos da lei, a comercialização de qualquer espécie sem sua autorização prévia.", mas somente dentro da descrição do registro do item descrevê-lo como utilizador da licença CC-BY-NC-SA.

Interessante notar a especificação do público a quem não se violará os Direitos de Autor, presente em um repositório "A equipe do (nome da Instituição) busca não violar o direito de propriedade das editoras comerciais[...]".

Em dois repositórios encontramos uma licença que faz menção a determinados aspectos, tais como a distribuição não-exclusiva, o uso da licença CC 3.0 Unported (licenças que não estão associadas a uma jurisdição específica), em relação ao uso comercial e a permissão para a realização de modificações à obra condicionada à utilização, posterior, sob a mesma licença.

Em poucos casos encontraremos a especificação da versão da licença CC utilizada. Trata-se do desconhecimento das versões das licenças e as alterações que elas podem vir a trazer. Percebe-se também o desconhecimento iconográfico de representação das licenças. Encontra-se, com frequência, a utilização do símbolo que representa que o conteúdo está protegido pela LDA – © juntamente com o símbolo das licenças Creative Commons (CC).

## **Conclusões**

Por meio do presente trabalho pudemos verificar a inadequação da Lei dos Direitos Autorais brasileira à realidade das bibliotecas, por não prever nenhum tipo de limitação ou exceção para esse espaço difusor de conhecimento relacionado às suas atividades intrínsecas, como também às suas novas práticas digitais de organização, disseminação e preservação dos diferentes itens bibliográficos.

Verificamos a importância de serviços de apoio para a realização de depósitos nos repositórios nacionais e internacionais e como esses serviços podem também ser úteis aos autores.

Averiguamos que o desconhecimento das leis conduz à utilização de informações jurídicas conflitivas ao gerir um repositório, podendo levar à utilização de diferentes licenças e que algumas vezes são antagônicas. Foi possível perceber também o desconhecimento, de fato, das licenças Creative Commons e as nuances de sua aplicação. Os gestores dos repositórios institucionais, na verdade e em sua grande maioria, não compreendem a possibilidade das diversas combinações possíveis para a formação de uma licença e terminam por crer que somente informar a versão CC já é o bastante para determinar quais direitos são reservados aos titulares dos Direitos de Autor.

Não há sombra de dúvidas que os temas relacionados com os Direitos Autorais devem ser mais difundidos em todas as áreas do conhecimento e com especial atenção junto à comunidade de biblioteconomia. Desta forma se promoverá o correto uso das obras sem que se firam os dispositivos legais. Esse conhecimento ao ser adquirido possibilitará a criação de uma maior consciência de que é fundamental um maior equilíbrio entre os interesses públicos e privados, principalmente em prol do avanço científico. E, sem lugar a dúvidas, para tanto, urge uma modernização na legislação nacional que trata do tema.

## Referências bibliográficas

BRASIL. Lei n.º 9610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. [Em linha] [Consult. 08 Set. 2016]. Disponível na internet < [http://www.planalto.gov.br/ccivil\\_03/leis/L9610.htm](http://www.planalto.gov.br/ccivil_03/leis/L9610.htm) >.

CREATIVE COMMONS BR. [Em linha] [Consult. 08 Set. 2016]. Disponível na internet < <https://br.creativecommons.org/> >.

BUDAPEST OPEN ACCESS INITIATIVE (2002) – Declaração de Budapeste. [Em linha]. Budapest. [Consult. 04 Set. 2016]. Disponível na internet < <http://www.budapestopenaccessinitiative.org/read> >.

DIRETÓRIO DE POLÍTICAS EDITORIAIS DAS REVISTAS CIENTÍFICAS BRASILEIRAS. [Em linha]. [Consult. 11 Set. 2016]. Disponível na internet < <http://diadorim.ibict.br/> >.

ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA. Repositório Institucional da ENAP. [Em linha] [Consult. 09 Set. 2016]. Disponível na internet < <http://repositorio.enap.gov.br/> >.

MAGNUS, Ana Paula Medeiros. Direito Autoral e Creative Commons: a participação do bibliotecário na promoção do acesso aberto à informação. Porto Alegre: Universidade Federal do Rio Grande do Sul, Departamento de Ciências da Informação, 2009. Trabalho de conclusão de curso. [Consult. 15 Mai. 2016]. Disponível na internet: < <http://www.lume.ufrgs.br/handle/10183/22705> >.

MIZUKAMI, Pedro Nicoletti [et al.]. Exceptions and limitations to copyright in Brazil: a call for reform. In: SHAVER, L. (Org.). Access to knowledge in Brazil: New Research on Intellectual property, innovation, and development. Yale, 2008. p. 67–122. [Consult. 15 Mai. 2016]. Disponível na internet: < [http://isp.yale.edu/sites/default/files/publications/A2KBrazil\\_bkmk.pdf](http://isp.yale.edu/sites/default/files/publications/A2KBrazil_bkmk.pdf) >.

OLIVEIRA, Flásleandro Vieira de. A licença Creative Commons e a Comunicação Científica: a questão dos periódicos eletrônicos de acesso aberto. Niterói: Universidade Federal Fluminense, Programa de Pós-Graduação em Ciência da Informação, 2011. Dissertação de mestrado. [Consult. 15 Mai. 2016]. Disponível na internet: < [http://www.cprm.gov.br/publique/media/Diss\\_Flasleandro\\_Oliveira.pdf](http://www.cprm.gov.br/publique/media/Diss_Flasleandro_Oliveira.pdf) >.

SHERPA/ROMEO – Publisher copyright policies & self-archiving. [Em linha] [Consult. 10 Set. 2016]. Disponível na internet < <http://www.sherpa.ac.uk/romeo/index.php> >.

UNIVERSIDADE DE SÃO PAULO. Biblioteca Digital da Produção Intelectual da Universidade de São Paulo. [Em linha] [Consult. 10 Set. 2016]. Disponível na internet < <http://www.producao.usp.br/> >.

UNIVERSIDADE FEDERAL DE GOIÁS. Repositório Institucional da UFG. [Em linha] [Consult. 08 Set. 2016]. Disponível na internet < <http://repositorio.bc.ufg.br>>.

UNIVERSIDADE FEDERAL DO PARÁ. Repositório Institucional da UFPA. [Em linha] [Consult. 09 Set. 2016]. Disponível na internet < <http://repositorio.ufpa.br/>>.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE. Repositório Institucional da Universidade Federal do Rio Grande do Norte. [Em linha] [Consult. 09 Set. 2016]. Disponível na internet < <https://repositorio.ufrn.br/>>.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL. Repositório Digital da Universidade Federal do Rio Grande do Sul. [Em linha] [Consult. 09 Set. 2016]. Disponível na internet < <http://www.lume.ufrgs.br/>>.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Repositório Institucional da Universidade Tecnológica Federal do Paraná. [Em linha] Consult. [09 Set. 2016]. Disponível na internet < <http://repositorio.utfpr.edu.br/>>.

# O RI no contexto da gestão da produção científica: Proposta de um questionário para aferição de níveis de satisfação e novas necessidades

---

Maria João Pinto

Universidade Católica Portuguesa, Porto

[mjpinto@porto.ucp.pt](mailto:mjpinto@porto.ucp.pt)

Sofia Fernandes

Universidade Católica Portuguesa, Porto

[sfernandes@porto.ucp.pt](mailto:sfernandes@porto.ucp.pt)

## Resumo

O presente artigo tem como objetivo apresentar uma proposta de um questionário para avaliar qualitativamente o Repositório Institucional no âmbito da gestão e da avaliação da produção científica institucional. O questionário pretende, assim, aferir as necessidades e o nível de satisfação da comunidade académica relativamente aos indicadores extraídos a partir do Repositório e da sua utilização.

A importância desta proposta prende-se, essencialmente, com a necessidade de perceber se o Repositório é uma ferramenta estratégica na gestão e avaliação da produção científica e quais os resultados pretendidos no contexto da investigação.

A proposta do questionário foi elaborada pelos profissionais de informação da Biblioteca e Gestão da Informação, da Universidade Católica Portuguesa, Centro Regional do Porto, diretamente envolvidos na gestão do repositório institucional e apresentada internamente à equipa da biblioteca para validação.

Apresentamos, assim, os passos de construção do questionário para possível aplicação em organizações cujos profissionais de informação pretendam aferir as necessidades e a satisfação destas relativamente aos dados das publicações científicas por parte da comunidade académica.

**Palavras-chave:** Repositório Institucional, Questionário, Avaliação, Indicadores



## **The IR in the context of management of scientific production: Proposal for a questionnaire to seek needs**

### **Abstract**

This paper aims to present a proposal for a questionnaire to qualitatively evaluate the Institutional Repository in the management and assessment of institutional scientific production. The questionnaire thus seeks to assess the needs and the level of satisfaction of the academic community for the indicators extracted from the repository and use.

The importance of this proposal relates essentially to the need to understand whether the repository is a strategic tool in the management and evaluation of scientific literature and what the intended results in the investigation.

The purpose of the questionnaire was drawn up by the professionals of the Biblioteca e Gestão da Informação, of the Universidade Católica Portuguesa, Centro Regional do Porto, directly involved in the management information of the institutional repository and presented internally validation for library staff.

Here, therefore, the construction of the questionnaire steps for possible application in organizations whose information professionals wishing to assess the needs and satisfaction of those concerning the facts of scientific publications by the academic community.

**Keywords:** Institutional Repository, Survey, Evaluation, Indicators

### **Introdução**

No meio académico e científico sente-se cada vez mais a necessidade de avaliar e medir os resultados da atividade científica. Estes resultados permitem monitorizar a atividade científica, ao nível interno, de forma a permitirem a aplicação de indicadores de produtividade dos investigadores e qualidade das publicações. Ainda, permitem a tomada de decisões e orientações estratégicas de investigação.

As organizações que desenvolvem investigação, nomeadamente as existentes no Ensino Superior necessitam de dados sobre a investigação realizada para complementar a informação sobre o ensino para efeitos de acreditação por parte de entidades externas. No caso de existirem centros de investigação que recorram a programas e agências de financiamento, a importância deste tipo de informação torna-se essencial. Fatores como a internacionalização e a situação económica vigente aumentam a competitividade e elevam os requisitos das candidaturas aos projetos de investigação no que diz respeito às publicações científicas.

Os Repositórios Institucionais (RI) permitiram que as instituições que desenvolvem investigação, e que precisam de dados sobre os resultados desta, conseguissem reunir as publicações e extrair indicadores, mesmo sem um sistema de informação próprio para a gestão da investigação desenvolvida. Por outro lado, instituições que adotaram um sistema

de informação puderam “alimentar” o sistema a partir do RI e/ou complementar a informação que tinham do RI devido à interoperabilidade possível entre os sistemas.

Desta forma, as bibliotecas e os profissionais de informação, enquanto gestores dos RI com competências na área da gestão da informação, viram-se envolvidos no complexo panorama da avaliação da Ciência.

Dentro das organizações, as necessidades de informação relativa às publicações concretizadas no exercício da investigação provêm de investigadores individuais, dos centros de investigação, das unidades de ensino respetivas e do departamento de qualidade interna das organizações. Estes, por sua vez, terão eventualmente solicitações de informação por elementos e instituições externas.

Não obstante a sua importância estratégica, no caso da Católica Porto ainda não foi possível a implementação de sistema de informação para a gestão e avaliação dos resultados das atividades científicas provenientes da investigação. De facto, pretende-se um sistema que permita essa gestão e seja interoperável com o RI mas que se encontra ainda em fase de estudo e de desenvolvimento. Na verdade, a partir do RI, a Biblioteca foi capaz de extrair e fornecer indicadores relevantes para investigadores, centros de investigação e para a gestão, avaliação e visão estratégica da universidade. Este trabalho de colaboração no âmbito da investigação nasceu da iniciativa da Biblioteca e envolve o processo desde a recolha das publicações, depósito, validação e extração de indicadores.

A diversidade e a distinção entre as áreas de ensino/investigação poderão ser fatores a ter em conta no que diz respeito às necessidades de informação, pelo que bibliotecas que façam a gestão dos repositórios institucionais deverão ter em conta que as especificidades das áreas de Ciências Naturais e Exatas, Humanidades, Ciências Sociais e Artes poderão eventualmente influenciar o tipo de informação que necessitam, nomeadamente ao nível da especificação das tipologias de publicação.

Por outro lado, a comunicação entre os profissionais de informação e a comunidade interna é um fator importante na perceção e no envolvimento de todos os agentes no processo para que os dados extraídos e análises efetuadas correspondam aos objetivos estratégicos das instituições.

### **Objetivo(s)**

Como gestores de repositórios institucionais, o objetivo da proposta do questionário é compreender se estamos a extrair informação relevante de acordo com as necessidades dos diversos interessados no que diz respeito a indicadores de investigação. Pretende-se, assim, um questionário que permita aferir o grau de satisfação dos vários intervenientes e perceber novas necessidades de informação e análises passíveis de aplicar.

### **Metodologia**

Esta proposta foi desenvolvida no âmbito do serviço da Biblioteca e Gestão da Informação (BGI) da Universidade Católica Portuguesa, Centro Regional do Porto, com vista à

futura implementação para obtenção de respostas para melhoria dos serviços prestados no que diz respeito aos indicadores extraídos dos RI. Não obstante, a proposta visa a possível aplicação por outras bibliotecas que sintam as mesmas necessidades.

O primeiro passo foi a identificação do público-alvo para aplicação do questionário, ou seja, ao longo do tempo quem necessitou e pediu informação relativa às publicações dentro da comunidade interna.

Após a identificação do público-alvo, seguiu-se o levantamento de solicitações efetuadas ao serviço da BGI e respetiva classificação de tipo de pedido, assim como o objetivo com o que era feito, sempre que possível determinar. A partir da análise da informação recolhida, foi possível definir os temas, as linhas de orientação e elaborar as questões.

Os passos anteriores permitiram a concretização do questionário que foi distribuído pela equipa da BGI para análise e sugestões de melhoria. A equipa reuniu-se para a validação final da proposta e, em conjunto, foi definido um modelo de distribuição.

## Resultados

O levantamento dos pedidos de informação relativa às publicações científicas disponíveis no RI permitiu a identificação do público-alvo do questionário e a definição de perfis: Produtores de Informação (docentes e investigadores), Consumidores de Informação (docentes e investigadores, centros de investigação, instituição) e (Re) Utilizadores de Informação (instituição – sistema de qualidade/melhoria contínua).

O levantamento de solicitações à BGI também forneceu a informação necessária para percebermos os tipos de pedidos até ao momento de forma a agruparmos em temas e constituirmos linhas de orientações para a elaboração das questões:

- Informação (dados extraídos a partir do RI);
- Análises de Informação (Indicadores e análises de informação extraída do RI; Informação e indicadores complementares);
- Comunicação (Colaboração e envolvimento entre os perfis de utilizadores e os gestores do RI; Novas necessidades, comentários e sugestões de melhoria).

A elaboração das questões foi realizada de acordo com as seguintes dimensões: atualização, adequação (aos objetivos/fins), usabilidade, utilidade e satisfação.

Apresentamos as questões elaboradas no âmbito das linhas de orientação definidas anteriormente, de acordo com os perfis e com as dimensões. Foram identificados os seguintes tipos/níveis de resposta: Sim, Algumas Vezes, Não e Não Aplicável. Na penúltima questão foi incrementada uma resposta de escolha e a última é de texto livre.

Informação Extraída
<b>Produtores de Informação</b>
As coleções em que as publicações estão agrupadas no RI correspondem às tipologias de publicação

O RI no contexto da gestão da produção científica: Proposta de um questionário para aferição de níveis de satisfação e novas necessidades

utilizadas na área de investigação?
As publicações depositadas são validadas e encontram-se disponíveis no prazo desejável?
Na sua opinião, a disponibilização das suas publicações em RI proporciona vantagens enquanto investigador?
<b>Consumidores de Informação</b>
As coleções em que as publicações estão agrupadas no RI correspondem às tipologias de publicação requeridas internamente?
Correspondem às tipologias de publicação requeridas externamente?
A informação extraída a partir do RI encontra-se atualizada de acordo com as diversas necessidades estratégicas e institucionais?
Tem demonstrado utilidade/vantagens?
<b>(Re) Utilizadores de Informação</b>
A Informação extraída a partir do RI permite responder às necessidades?
Encontra-se atualizada de acordo com o desejável?
Tem demonstrado utilidade/vantagens?
<b>Análise de Informação</b>
<b>Produtores de Informação   Consumidores de Informação   (Re) Utilizadores de Informação</b>
Os indicadores e as análises preparadas e fornecidas pela BGI a partir do RI são adequados e suficientes para a informação necessária?
As análises preparadas proporcionam informação de valor acrescentado?
As fontes de informação complementares (JCR, Scopus, p.ex.) permitem análises relevantes?
Existem outros indicadores ou análises que poderiam ser de interesse incluir?
<b>Comunicação</b>
<b>Produtores de Informação   Consumidores de Informação   (Re) Utilizadores de Informação</b>
Na sua opinião, a colaboração da BGI tem sido útil no que diz respeito à informação extraída e preparada a partir do RI?
Ainda na sua opinião, a o processo encontra-se bem conduzido?
Os prazos de tempo de execução são razoáveis e cumpridos de acordo com as necessidades/tempo disponível de execução?
Na sua opinião, um maior envolvimento da BGI em atividades (ou acesso a relatórios, p.ex.) poderia fornecer melhores resultados?
Caso a resposta anterior seja positiva, quais serão os motivos que dificultaram ou impossibilitaram

<p>um maior envolvimento?</p> <p>a) Não houve necessidade até ao momento.</p> <p>b) Falta de tempo/recursos por parte da comunidade de investigadores.</p> <p>c) Falta de disponibilidade da BGI.</p> <p>d) Não houve conhecimento de competências ou recursos da BGI que pudessem ser úteis.</p>
<b>Comentários e Sugestões</b>
<b>Produtores de Informação   Consumidores de Informação   (Re) Utilizadores de Informação</b>
P.f. deixe o seu comentário face aos tópicos abordados neste questionário e/ou sugestões de melhoria.

**QUADRO 1 – QUESTÕES ELABORADAS PARA O QUESTIONÁRIO**

A equipa da BGI mostrou-se em concordância com as questões definidas e foi decidido um modelo de distribuição por correio eletrónico, a partir do endereço eletrónico da biblioteca, com aviso de leitura e com prazo de preenchimento de 1 mês. Foi ainda pensado que após 2 semanas, seria enviado um e-mail a relembrar o prazo e a necessidade de resposta.

### **Considerações Finais e Perspetivas Futuras**

A construção deste questionário permitiu-nos reunir e perceber toda a informação de que dispomos acerca das necessidades de informação extraída a partir do Repositório. Pretende-se a aplicação futura deste questionário para aprofundarmos esse conhecimento e aferirmos o grau de satisfação do processo de gestão de informação científica e de novas necessidades/melhorias.

De forma a aumentarmos a possibilidade de respostas ao questionário, foi decidido que este questionário seria enviado em Janeiro por motivos de maior disponibilidade da comunidade académica.

### **Referências bibliográficas**

AZEVEDO, Carlos; AZEVEDO, Ana– Metodologia Científica: Contributos práticos para a elaboração de trabalhos académicos. Porto: Azevedo, Carlos, 1994. ISBN 972-9114-11-0

CASSELLA, M. (2010) –Institutional repositories: an internal and external perspective on the value of IRs for researchers’ communities. *Liber Quarterly*. Vol. 20 (2010), p.210-225.

JOINT, Nicholas (2008) – Current research information systems, open access repositories and libraries. *Library Review*. ISSN 0024-2535. Vol. 57, nº8, p. 570-575. doi:10.1108/00242530810899559

---

## Construção de um repositório de dados oceanográficos

---

Ricardo Amorim

INESC TEC, Faculdade de Engenharia, Universidade do Porto  
[ricardo.amorim3@gmail.com](mailto:ricardo.amorim3@gmail.com)

João Castro

INESC TEC, Faculdade de Engenharia, Universidade do Porto  
[joaoaguiarcastro@gmail.com](mailto:joaoaguiarcastro@gmail.com)

Inês Garganta

INESC TEC, Faculdade de Engenharia, Universidade do Porto  
[ei10162@fe.up.pt](mailto:ei10162@fe.up.pt)

Artur Rocha

INESC TEC, Faculdade de Engenharia, Universidade do Porto  
[artur.rocha@inesctec.pt](mailto:artur.rocha@inesctec.pt)

Gabriel David

INESC TEC, Faculdade de Engenharia, Universidade do Porto  
[gtd@fe.up.pt](mailto:gtd@fe.up.pt)

### Resumo

O artigo descreve a conceção e arquitetura de um repositório de dados de investigação de campanhas oceanográficas. O ponto de partida é o modelo de dados OGC Sensor Observation Service, complementado com uma extensão que visa responder às especificidades do caso de aplicação, as campanhas do projeto BIOMETORE, coordenado pelo Instituto Português do Mar e da Atmosfera. Escolheu-se uma implementação de software aberto do SOS, a qual é complementada com dados relativos às campanhas, equipas e documentos. Definiram-se os esquemas de metadados relevantes, por exemplo, para a classificação das espécies. Desenvolveu-se uma aplicação de carregamento dos dados baseada nos formulários usados pelos investigadores, com uma versão auxiliar móvel. E serão desenvolvidos serviços de

visualização dos dados e dos resultados processados, aptos a serem integrados em agregadores de dados marinhos.

**Palavras-chave:** Gestão de dados de investigação, Aquisição de metadados, Dados oceanográficos, Diretiva INSPIRE

## **Building an oceanographic data repository**

### **Abstract**

The paper describes the overall design and architecture of a research data repository for oceanographic campaigns. The starting point is the OGC Sensor Observation Service (SOS) data model, complemented by an extension intended to answer the specifics of the application case, the campaigns of the BIOMETORE project, coordinated by the Portuguese Sea and Atmosphere Institute. An open source SOS implementation has been chosen and it has been complemented with data on the campaigns, the teams, and the documents. The relevant metadata schemes have been defined, for instance, the species classification scheme. A data management application has been developed, based on the forms used by the researchers, with an auxiliary mobile version. Visualization services will be implemented, both for data and for processed results, able to be integrated in marine data aggregators.

**Keywords:** Research data management, Metadata acquisition, Oceanographic data, INSPIRE directive

## **Introdução**

A investigação oceanográfica, tal como muitas outras observações da Natureza, produz dados em grande medida únicos, pois as campanhas de recolha são irrepetíveis, tal a quantidade de variáveis não controladas pelos investigadores. A isso acresce o custo habitualmente elevado das campanhas, pelo navio, equipamentos e equipas que mobilizam. O valor destes dados deriva não só do contexto de produção como também do seu uso potencial. Destinam-se a contribuir para monitorizar e suportar decisões relativas ao Bom Estado Ambiental das Águas Marinhas Europeias. É, portanto, imperioso que estes dados possam ser partilhados e replicados da forma mais alargada possível.

Acontece que tradicionalmente as técnicas de recolha, armazenamento e disseminação dos dados dos projetos de investigação oceanográfica não têm permitido uma partilha significativa. Os dados são muitas vezes registados em formulários em papel, cuidadosamente preparados, e posteriormente digitalizados ou copiados para folhas de cálculo, mas a ênfase tem estado mais na publicação dos artigos com os resultados da investigação do que na dos dados recolhidos. Existem também fluxos de dados provenientes de sensores e de satélites, recolhas de amostras de solo, de água, de plantas e de animais, diários de pesca, percursos de navios, gravações de fotografias, vídeo e som e resultados de análises laboratoriais.

Este problema reflete-se no projeto BIOMETORE (IPMA, 2015), um projeto de grande dimensão coordenado por investigadores do Instituto Português do Mar e da Atmosfera. A equipa de investigadores neste projeto é multidisciplinar e inclui geofísicos, biólogos e químicos que, em oito campanhas, produzem quantidades significativas de registos que devem ser tornados públicos, num suporte digital online, de forma a facilitar a sua reutilização e o seu processamento por diversos destinatários. É esse o objetivo do projeto complementar SeaBioData (INESCTEC, 2015).

O projeto SeaBioData visa organizar e armazenar todos os tipos de dados mencionados e os respetivos metadados de contexto, descritivos e técnicos. Pretende ainda disponibilizá-los, diretamente ou após processamento, em serviços de visualização e de interoperabilidade. Apresentam-se em seguida as principais decisões de arquitetura do sistema.

### Arquitetura proposta

O ponto de partida, dada a complexidade dos dados, e também a existência da diretiva europeia INSPIRE para o desenvolvimento de uma Infraestrutura para a Informação Geoespacial (European Commission, 2007), aplicável ao registo de informação ambiental georreferenciável, foi a adoção do modelo de dados OGC Sensor Observation Service (Bröring et al., 2014). Este modelo assenta num conjunto de abstrações suficientemente genéricas para poderem ser úteis em múltiplos tipos de dados: a observação, as propriedades da observação, o processo utilizado, a referência geográfica, etc.

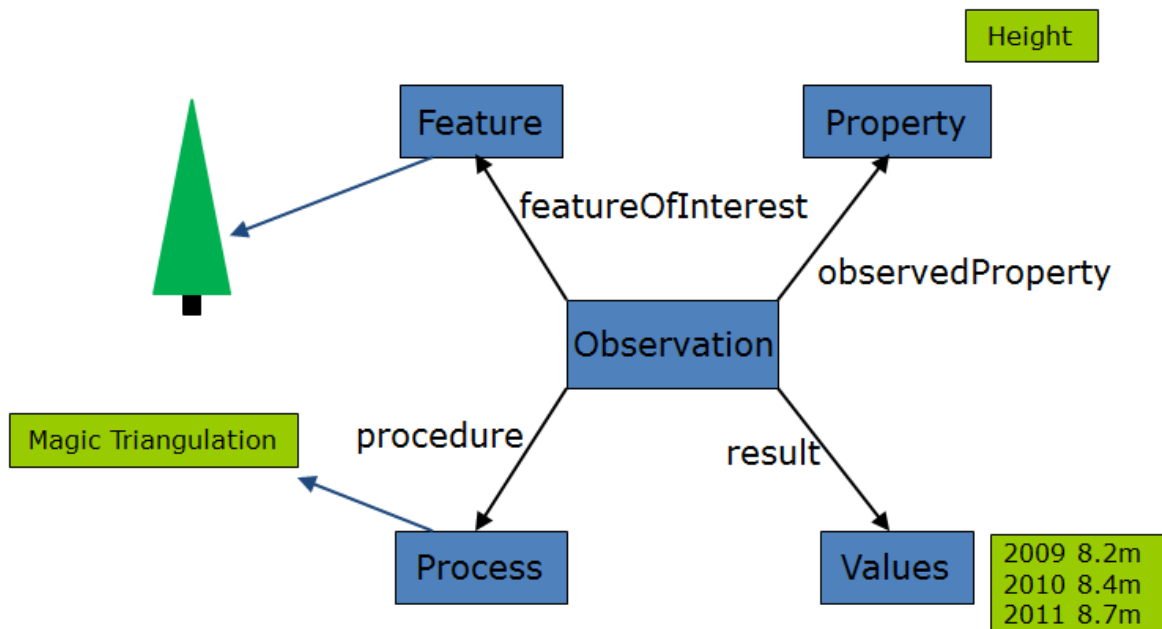


Figura 1: O conceito de observação (INSPIRE, 2011)

A figura 1 está centrada no conceito de observação. A observação é sobre um aspeto de interesse (Feature), habitualmente com uma expressão geográfica: um ponto, uma linha,



uma área, etc. Na figura trata-se de uma árvore, com determinadas coordenadas geográficas. A observação é realizada de acordo com um determinado procedimento (Process), que pode incluir referências ao método de obtenção e ao equipamento específico utilizado. Na figura, o método utilizado foi a triangulação mágica. A observação pode abranger diversas propriedades (Property), cada qual com o seu valor medido (Value). A propriedade observada foi a altura da árvore e poderia ter sido observada apenas num instante, produzindo um único valor. Neste exemplo, optou-se por mostrar uma série de três valores medidos, em três anos consecutivos.

Se, como é frequente no caso de aplicação da oceanografia, a uma observação corresponderem várias amostras, por exemplo, para procedimentos laboratoriais subsequentes, é possível considerar cada uma como observação subordinada à primeira, mudando o aspeto de interesse e considerando as respetivas propriedades. Por exemplo, se forem recolhidas folhas a várias alturas da árvore, cada uma passaria a ser o aspeto de interesse de uma observação subordinada, relacionada com a observação principal na árvore.

Usar este modelo constitui o primeiro passo para garantir a interoperabilidade semântica com outros sistemas baseados no mesmo modelo. Optou-se, em seguida, por utilizar a implementação em software aberto 52° North do SOS, que suporta nas suas estruturas de dados a maior parte dos conceitos e possui API REST e de serviços Web.

No entanto, as especificidades do projeto Biometore, aconselharam a acrescentar uma extensão a este modelo, também com API REST similar à existente, para armazenar metadados relativos aos projetos, às campanhas, às equipas, aos documentos, etc. Com efeito, a produção dos dados de investigação é realizada num contexto específico, que é relevante conhecer, de modo a contribuir para a correta interpretação das condições da obtenção dos dados e para a autenticidade da informação.

Desta extensão local (ver figura 2), salienta-se a tabela de projetos, que regista os detalhes de todos os projetos no âmbito dos quais foram produzidos dados. Os projetos mais complexos organizam-se em work packages. Tanto os projetos como os work packages são realizados por equipas de investigadores (*user*). Os investigadores estão afiliados em departamentos (*department*) de organizações (*organization*) e podem ser registados contactos tanto para os investigadores como para as organizações (*ci\_contact*, *ci\_address*, *ci\_telephone*). A tabela *ci\_responsibleparty* faz a ligação entre os investigadores e o seu papel concreto num work package de um projeto. É esta tabela que se relaciona com a tabela de campanhas (*campaign*), as quais são vistas como estruturando as atividades de realização de observações. Com esta informação, será possível definir políticas de acesso aos dados, acompanhando as várias fases de desenvolvimento de um projeto.

A preparação de uma campanha inclui uma série de atividades que convém registar, com o objetivo de facilitar o trabalho de registo durante o decorrer da campanha e,

simultaneamente, melhorar a qualidade dos respetivos dados. Estas atividades incluem o planeamento das estações (*stations*) a visitar e a identificação dos procedimentos (*procedureslist*) a seguir em cada uma dessas estações.

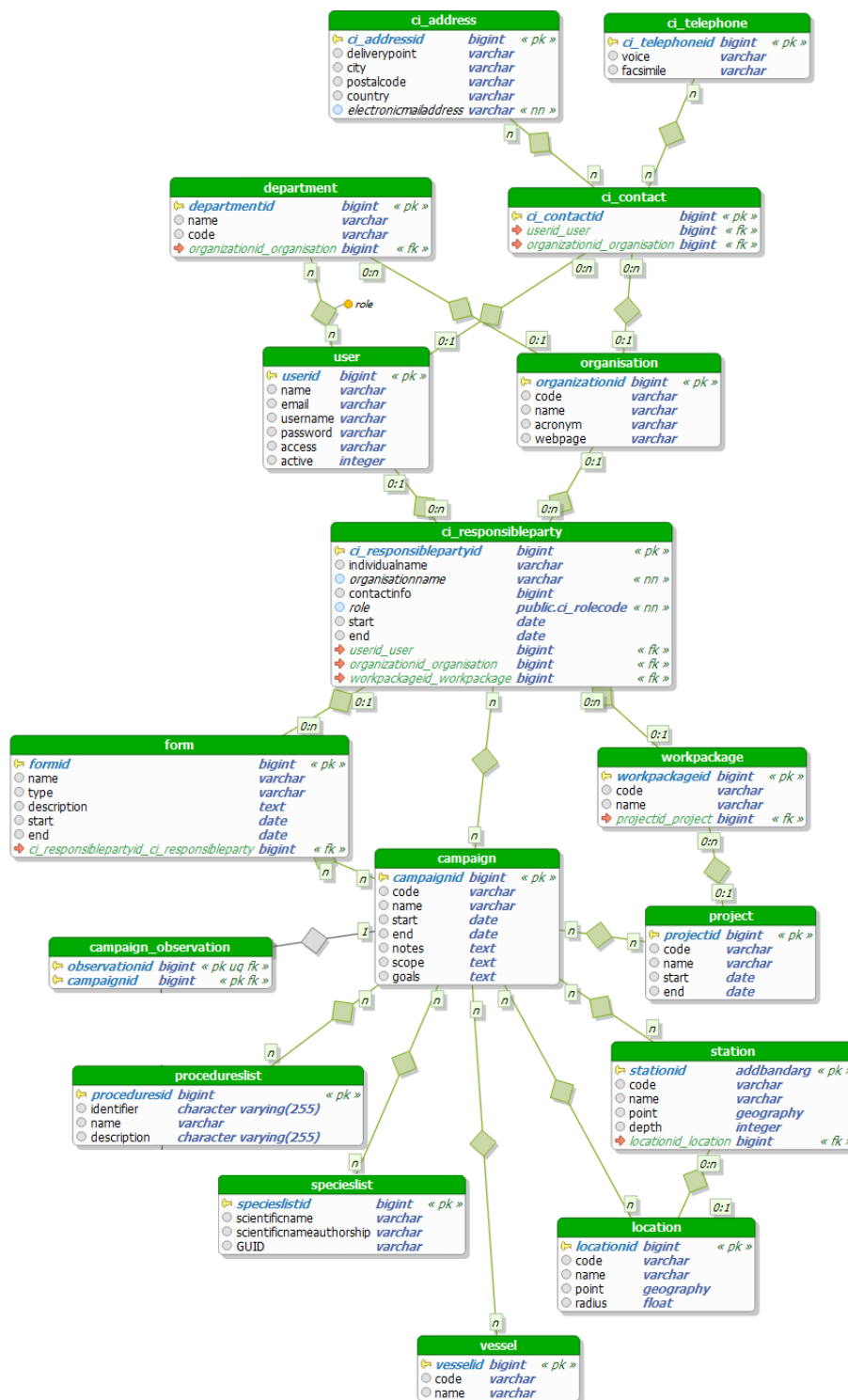


Figura 2: Modelo da base de dados local do SeaBioData (Garganta, 2016)

Um procedimento inclui um método e eventualmente equipamentos. Um exemplo de uma tabela de equipamentos é a tabela de navios (*vessel*), mas toda a lista de logística a

utilizar deve estar disponível, de forma a facilitar a seleção de um equipamento específico para uma instância de um procedimento. Os procedimentos são ainda associados a formulários (*form*), definidos a priori. Da experiência tida, obtiveram-se resultados importantes, a informação dos procedimentos e dos formulários deve ser flexível, admitindo evoluções que se adaptem às decisões tomadas durante a própria campanha. Isso pode significar a inclusão ou exclusão de propriedades para novas observações segundo o mesmo procedimento.

Devido às condições das operações de recolha ou às preferências dos investigadores, uma parte significativa dos registos começa por ser em formulários em papel, os quais são posteriormente transcritos para o repositório. Mas, e ainda na perspetiva de aumentar a autenticidade da informação, é requisito local que esses documentos primários sejam digitalizados e os respetivos ficheiros associados ao repositório. Neste caso, optou-se por armazenar os ficheiros numa hierarquia de pastas que mapeia a organização do repositório, por projeto, campanha e procedimento e, dentro disso, por ordem cronológica. Os metadados relevantes ficam na base de dados. Desta forma é possível eliminar o papel, sem perda de informação.

O armazenamento de ficheiros é também relevante para os vários tipos de dados que não sejam considerados adequados para carregamento como valores de propriedades de observações. Este é o caso das fotografias de aspetos de interesse, que podem ser consideradas como propriedades de uma dada observação, sendo o caminho até aos respetivos ficheiros registado enquanto tal. A fotografia pode por sua vez ser considerada como um aspeto de interesse ela própria, se for, por exemplo, sujeita a um processamento para se determinar a contagem de espécies. Os metadados técnicos da fotografia podem ficar na tabela de ficheiros. Algo de semelhante se passa com os vídeos, com os registos sonoros e até com os ficheiros CTD, caso se opte por não carregar as suas linhas individualmente.

Um outro elemento distintivo deste projeto é a preocupação com a definição prévia dos esquemas de metadados mais adequados a cada situação. A definição do modelo de metadados seguiu a estrutura da norma ISO 19115 para a descrição de dados geográficos, garantindo assim a descrição de dados em conformidade com a regulamentação INSPIRE para os metadados. Contudo, tendo em conta o perfil multidisciplinar do projeto BIOMETORE, foi identificada a necessidade de incluir descritores de forma a permitir a classificação de espécies e a descrição de processos metodológicos associados às amostras. Assim reutilizou-se um subconjunto de conceitos da Ecological Metadata Language e da Darwin Core. Todos os conceitos identificados para o perfil de metadados foram propostos em reuniões junto dos investigadores, para que estes tivessem a oportunidade de validar, excluir ou propor novos conceitos. Para além disso, para simplificar e reduzir o esforço dos investigadores na descrição de dados foram desenvolvidos vocabulários controlados tanto

previstos em normas existentes como definidos de acordo com as prioridades dos investigadores.

### Alimentação do repositório

O melhor momento para registar metadados é o do registo dos próprios dados. Por isso, incluiu-se no projeto a preparação de um conjunto de tablets, com a aplicação LabTablet. As ferramentas para a descrição de dados incluem o caderno de laboratório eletrónico, Labtablet, e a definição de um modelo de metadados. No domínio da descrição de dados, existem várias ferramentas que dão apoio aos investigadores e se integram no ambiente laboratorial ou de campo. Neste contexto, os cadernos de laboratório eletrónicos são ferramentas essenciais para a captura de metadados que permitem, entre outros aspetos, a perceção do contexto de produção do conjunto de dados. O LabTablet é um caderno de laboratório eletrónico com ênfase no suporte para metadados compatíveis com normas existentes em cada domínio. Neste contexto, a aplicação também dá suporte ao registo de recolha de amostras, seguindo formatos para a recolha que estão já estabelecidos no domínio.

Está a ser desenvolvida uma aplicação para funcionar como interface de registo e pesquisa simples para os investigadores, a qual deve estar o mais próxima possível dos formulários que estes criaram e conhecem. Esta aplicação é encarada como interface local sobre dois componentes, a base de dados interoperável do SOS da 52° North e a extensão local para aspetos específicos do projeto BIOMETORE.

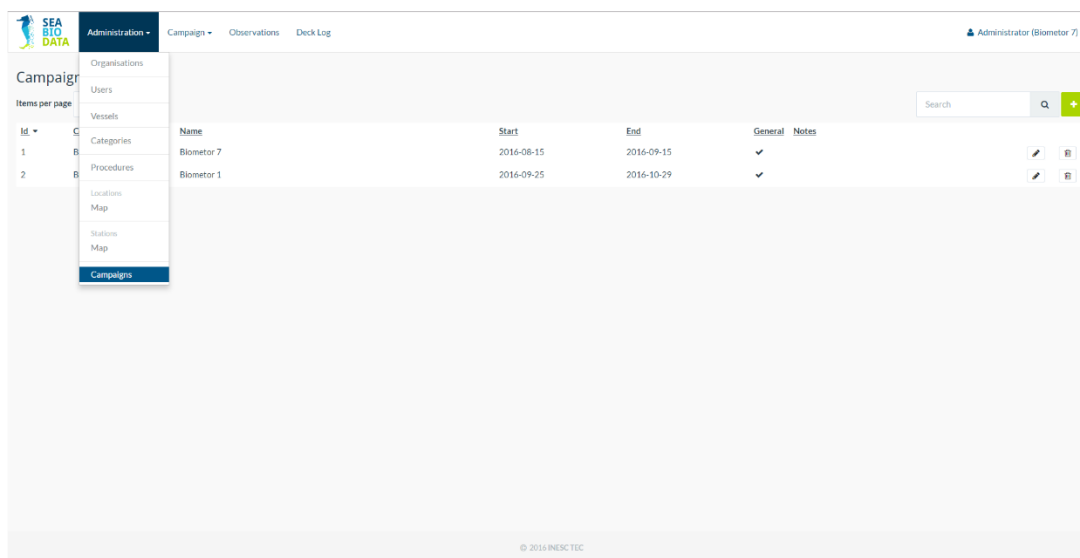
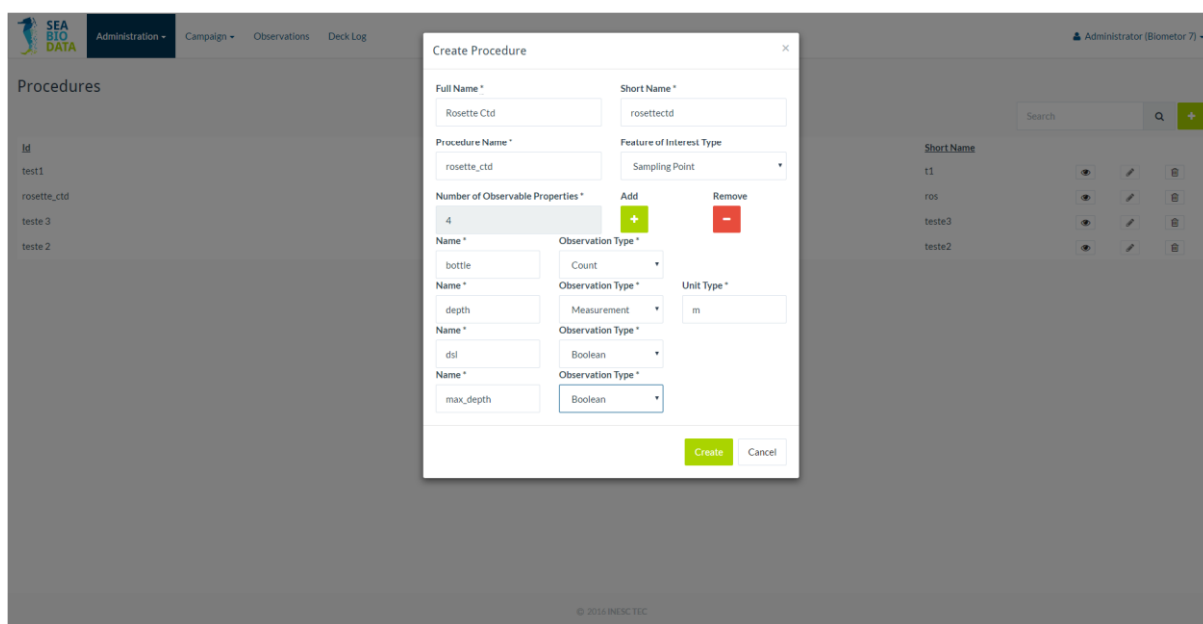


Figura 3: Interface de Administração

Aos formulários é necessário acrescentar visualizadores para os diversos tipos de dados e até para alguns dados processados. Serão também desenvolvidos serviços sobre os

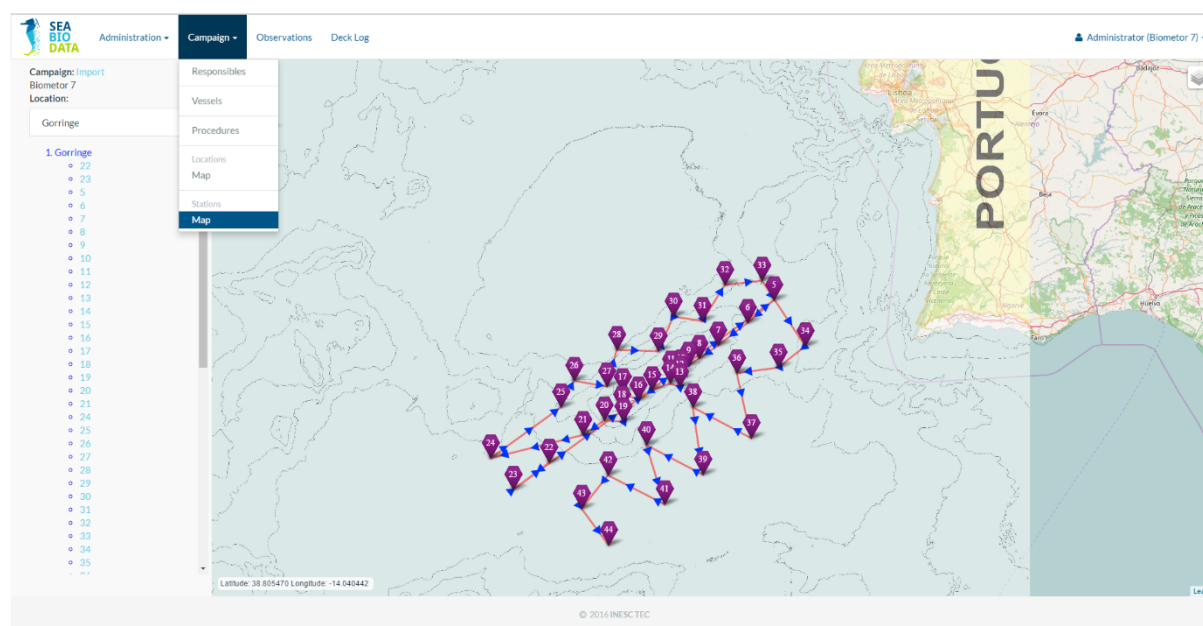
dados aptos a serem importados por plataformas de agregação de dados marinhos. Nas figuras 3, 4, 5 e 6 está apresentada parte da interface desenvolvida.

A componente de administração (figura 3) possibilita a gestão das organizações, utilizadores, navios, procedimentos, estações, etc., associados ao projeto. Foi implementada ainda a funcionalidade de criação de procedimentos (figura 4), de modo a que os investigadores consigam gerir todas as componentes dos projetos/campanhas



**Figura 4:** Interface de criação de procedimentos

A gestão da campanha (figura 5) permite associar utilizadores, navios, procedimentos, localizações e estações. Esta gestão deve ser feita antes do início de cada campanha, podendo ser atualizada conforme necessário durante a mesma.



**Figura 5:** Interface de gestão da campanha ativa, com a opção mapa das estações selecionada

De destacar a funcionalidade de criação/remoção de estações, capaz de acrescentar pontos diretamente num mapa (API do Google Maps) e definir depois a ordem de visita. Na visualização das observações introduzidas (figura 6) foi implementada uma vista em tabela, de modo a assemelhar-se com a interface de uma folha de cálculo, facilitando assim o reconhecimento e a adaptação dos investigadores à nova plataforma. Acrescentou-se também aqui a funcionalidade de poder associar ficheiros, como fotografias, a uma determinada observação.

Procedure	Station	Sample	Timestamp	Latitude	Longitude	bottle	depth [m]	dsl	max depth
rosette_ctd	7	ros 7 1	2016-09-01 20:52:54	41.178252	-8.594789	1	100		x
rosette_ctd	23	ros 23 2	2016-09-01 21:53:57	41.178252	-8.594789	2	80	x	
rosette_ctd	23	ros 23 3	2016-09-01 22:54:37	41.178252	-8.594789	3	60	x	

**Figura 6:** Interface de visualização das observações introduzidas para o procedimento selecionado

A interface de inserção de uma nova observação (figura 7) apresenta os campos que foram definidos como fazendo parte da observação durante a criação do respetivo procedimento, mais um cabeçalho com a estação da recolha, o código da amostra e ainda a posição (latitude e longitude) mais a data e hora da recolha.

**Figura 7:** Interface de introdução de uma observação

A interface apresentada foi desenvolvida com técnicas de adaptabilidade que a tornam adequada tanto para computadores como para tablets.

## Conclusão

Uma das tarefas prevista para o projeto era a realização de um primeiro teste em ambiente de campanha, que permitisse avaliar os pressupostos do ciclo de vida da produção de dados, especialmente nas primeiras fases de planeamento de campanhas e de utilização a bordo, e a usabilidade da interface de recolha de dados. Esse teste já teve lugar, no mês de setembro de 2016.

As reações obtidas já permitiram corrigir aspetos de usabilidade e reconhecer as vantagens do preenchimento por omissão de alguns valores, como as coordenadas GPS e o instante correntes, ou o procedimento, a estação e o operador do registo anterior, sem prejuízo de qualquer destes valores poder ser corrigido manualmente. Os tablets são ainda úteis para cronometrar e registar durações e tirar fotografias. Um segundo aspeto que ficou claro foi a necessidade de flexibilidade no planeamento, permitindo correções no percurso, com a inclusão de novas estações e a eliminação de estações ainda não utilizadas. Foi necessário até efetuar modificações nos procedimentos, de molde a permitir incluir propriedades não previstas. Para que tal seja possível, é necessário que nos resultados seja possível registar um valor nulo ou “não aplicável”, para preencher a nova propriedade nas observações anteriores à modificação.

O teste realizado não permite ainda estimar a percentagem de registos primários que continuarão a ser efetuados em formulários em papel, dado o carácter experimental do teste. É, no entanto, de crer, que essa percentagem continuará a ser não desprezável, o que coloca a questão da digitalização desses documentos e da sua associação às respetivas transcrições no repositório.

Os próximos passos do projeto são a realização de um teste em ambiente laboratorial, o carregamento sistemático dos dados das campanhas anteriores e a elaboração de serviços de visualização.

*Nota:* Este trabalho foi suportado pela EEA Grant PT02\_Aviso5\_0002 SeaBioData – Portuguese Seamounts Biodiversity Data Management.

## Referências bibliográficas

BRÖRING, A.; STASCH, C.; ECHTERHOFF, J. (2014) – *OGC Sensor Observation Service Interface Standard*, Tech. Rep. OGC and ISO 19156:2011(E), Open Geospatial Consortium (2014).

EUROPEAN COMMISSION (2007). *Council Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)*. OJ L108/1.

GARGANTA, Inês (2016) – *Exploring the Sea: Heterogeneous Geo-Referenced Data Repository*, dissertação de mestrado em Engenharia Informática e Computação, Faculdade de Engenharia da Universidade do Porto.

INSPIRE Working Group on Observations & Measurements (2011). *D2.9 Guidelines for the use of Observations & Measurements and Sensor Web Enablement – related standards in INSPIRE Annex II and III data specification development*, p16 [Em linha: [http://inspire.ec.europa.eu/documents/Data\\_Specifications/D2.9\\_O&M\\_Guidelines\\_v2.0rc3.pdf](http://inspire.ec.europa.eu/documents/Data_Specifications/D2.9_O&M_Guidelines_v2.0rc3.pdf)] [Consult. 2016-10-23].

IPMA (2014) – *Projeto EEA Grants PT02-0018 BIOMETORE – Biodiversity in seamounts: the Madeira-Tore and Great Meteor* [Em linha: <http://www.biometore.pt>] [Consult. 2016-10-23].

INESCTEC; IPMA (2015) – *Projeto EEA Grants, PT02\_Aviso5\_0002 SeaBioData – Portuguese Seamounts Biodiversity Data Management* [Em linha: <http://proj.inesctec.pt/seabiodata>] [Consult. 2016-10-23].