

## O Dataverse e suas contribuições para a comunicação científica

**Pollianna Marys de Souza e Silva**

Universidade Federal da Paraíba/UFPB. Doutoranda em Ciência da Informação, Paraíba, Brasil  
<https://orcid.org/0000-0002-1134-6264>  
 pollianna\_marys@hotmail.com

**Levi Cadmiel Amaral da Costa**

Universidade Federal da Paraíba/UFPB. Mestrando em Ciência da Informação, Paraíba, Brasil  
<https://orcid.org/0000-0002-6937-0187>  
 levy.cadmiel@gmail.com

### Resumo

Os Repositórios de Dados (RD) garantem os princípios de transparência e oferecem um sistema de armazenamento seguro, além da possibilidade de se ter os dados de pesquisa disponíveis on-line, indexados, documentados, para serem acessados, baixados, visualizados e processados por pessoas ou por sistemas, estendendo-a uma comunidade mais ampla e conectada em rede. O objetivo dessa pesquisa é apresentar as contribuições proporcionadas pela utilização do RD Dataverse para promover eficiência na comunicação científica. Trata-se de uma pesquisa qualitativa, descritiva e bibliográfica na qual buscou-se identificar relações entre o conteúdo bibliográfico sobre comunicação científica, dados e repositórios de dados e conteúdos informativos sobre o RD Dataverse contidos nos sites “The Dataverse Project” – site de apresentação do Dataverse - e “Harvard Dataverse”. Os resultados demonstram as funcionalidades do Dataverse possui relações com aspectos que beneficiam a comunicação científica, dentre os quais podem ser mencionados: preservação de dados científicos, produção de metadados de qualidade, promoção de facilidade em processos de recuperação da informação científica, disponibilização de dados que podem ser citados, dentre outras questões. Portanto, conclui-se que o Dataverse proporciona benefícios que estão associados aos processos de comunicação entre pesquisadores científicos.

### Palavras-chave

Comunicação Científica. Eficiência na Comunicação Científica. Arquivos Científicos.

## Dataverse and its contributions to scientific communication

### Abstract

Data Repositories (RD) guarantee the principles of transparency and offer a secure storage system, in addition to the possibility of having research data available online, indexed, documented, to be accessed, downloaded, viewed and processed by people or by systems, extending it to a wider and networked community. The objective of this research is to present the contributions provided by the use of RD Dataverse to promote efficiency in scientific communication. This is a qualitative, descriptive and bibliographic research in which we sought to identify relationships between the bibliographic content on scientific communication, data and data repositories and informational content on the RD Dataverse contained in the sites “The Dataverse Project”- presentation site of the Dataverse - and “Harvard Dataverse”. The results demonstrate the functionality of Dataverse has relations with aspects that benefit scientific communication, among which can be mentioned: preservation of scientific data, production of quality metadata, promotion of ease in processes of retrieval of scientific information, availability of data that can be cited, among other issues. Therefore, it is concluded that the Dataverse provides benefits that are associated with the communication processes between scientific researchers.

### Keywords

Scientific Communication. Efficiency in Scientific Communication. Scientific Archives.



Licença de Atribuição BY do Creative Commons  
<https://creativecommons.org/licenses/by/4.0/>

Submetido em 19/02/2021  
 Aprovado em 03/05/2021  
 Publicado em 01/07/2021

## 1 INTRODUÇÃO

O presente artigo trata de comunicação científica no contexto dos Repositórios de Dados (RD), focando, mais especificamente, no Dataverse, o qual trata-se de um RD desenvolvido e mantido pela Universidade de Harvard. O interesse por esse tema surgiu pela necessidade de aprofundar conhecimentos sobre os RD de pesquisa, bem como compreender como o Dataverse pode se caracterizar como veículo de comunicação científica.

Os RD surgiram a partir das iniciativas do movimento de acesso aberto, cuja gênese foi a crise dos periódicos científicos na década de 70, porém, somente a partir da década de 90, as publicações científicas eletrônicas fizeram uma mudança radical na Comunidade Científica. Em 2001 foi realizada a reunião de Budapeste que definiu o 1º protocolo de interoperabilidade, o *Open Access Initiative – Protocol for Metadata Harvesting* (OAI – PMH) (BJORK, 2004), que constituem hoje uma inovação no sistema de comunicação científica. O 1º RD surgiu no início da década de 90, nos EUA, intitulado ArXiv5, o qual envolvia-se às seguintes áreas: ciências da computação, física, matemática e ciências não lineares (SHINTAKU; MEIRELES, 2010).

Para González de Gómez (2002), a governança informacional está baseada na convergência teórico-conceitual associada à governança, *accountability* (prestação de contas), transparência, bem como à garantia do direito ao acesso à informação pública e governamental.

Neste sentido, podemos considerar que os RDs podem ser associados à governança informacional, uma vez que, os dados armazenados em repositórios são resultantes de pesquisas que são aplicadas a partir dos investimentos públicos em instituições de ensino. Portanto, deve-se haver transparência e compartilhamento destes dados para que os mesmos sirvam de insumos para informações a serem disponibilizadas para a quem interessar.

Para apoiar a execução dos processos de gestão é necessário um arcabouço tecnológico e gerencial que compreenda todo o ciclo de vida dos dados. No centro desse arcabouço estão os RD de pesquisa que, por muitas razões e demandas, rapidamente se tornam parte essencial da infraestrutura mundial de pesquisa (FORMENTON, 2015). Para Sayão e Sales (2016) os RD garantem os princípios de transparência e oferecem um sistema de armazenamento seguro, além da possibilidade de se ter os dados de pesquisa disponíveis *on-line*, indexados, documentados, para serem acessados, baixados, visualizados e processados por pessoas ou por sistemas, estendendo-a uma comunidade mais ampla e conectada em rede.

Pela sua capacidade de agregação e organização e como instrumento de socialização de comunidades e pesquisadores, os RD têm sua relevância como recurso informacional e se tornam um dispositivo de troca de experiências e compartilhamento de dados, se configurando parte

essencial das pesquisas em escala global, portanto, podemos afirmar que este tipo de repositório contribui demasiadamente para colaboração e comunicação entre pesquisadores da comunidade científica.

Especificamente em relação ao Dataverse, este trata-se de um RD aberto, no qual há a viabilidade de compartilhar, citar, preservar, analisar e explorar dados oriundos de pesquisas científicas. Por meio do Dataverse, a comunidade científica colabora entre si, haja vista que é viável a elaboração de investigações com base em dados obtidos em pesquisas anteriores. Além disto, o RD em questão proporciona visibilidade na web e crédito acadêmico para investigadores, periódicos, editores, autores de dados, distribuidores de dados e instituições afiliadas (DATAVERSE, 2020).

Considerando estes termos, o objetivo da presente pesquisa é apresentar as contribuições proporcionadas pela utilização do Dataverse para promover eficiência na comunicação científica. Diante disto, como forma de fundamentar o objetivo da pesquisa e servir como questionamento para a busca de resultados, pretendemos responder à seguinte indagação: “de que forma o RD Dataverse contribui para eficiência na comunicação científica?”.

A realização deste trabalho de investigação se justifica ao levarmos em consideração as diversas pesquisas científicas que afirmam relevância dos repositórios de dados (BARBOSA, 2013; CURTY, 2017; PARANGABA, 2017), bem como as que ressaltam a importância da comunicação no âmbito científico (FERREIRA, 2017; TREVISOL NETO; CAFÉ; SILVA, 2017; PEREIRA, 2019). Portanto, este estudo dá visibilidade e poderá servir de subsídio para pesquisas posteriores que abordem estas duas temáticas de relevância científica.

## 2 DESENVOLVIMENTO

O desenvolvimento deste trabalho está dividido em três tópicos. No primeiro, apresenta-se o referencial teórico, no qual buscou-se identificar correlações entre comunicação científica, dados e repositórios de dados. No segundo, são abordados os procedimentos metodológicos utilizados para a realização da presente pesquisa. Por fim, no terceiro tópico, se realiza a análise e discussão dos dados, correlacionando estes ao referencial teórico.

## 2.1 Comunicação Científica e suas Correlações com Dados e Repositórios de Dados

De acordo com o que afirmam Cunha e Cavalcanti (2008), o conceito de comunicação científica foi elaborado em meados de 1930, no qual John Bernal se referiu a esta como um tipo de comunicação associada à produção, consumo e transferência de informações no âmbito científico.

Conforme mencionado por Branco e Gonzaga (2014 p. 147), a comunicação científica caracteriza-se “pela veiculação de informações científicas aos grupos de iguais – intrapares e grupos de afins – extrapares”. Como ressaltado pelos autores, neste tipo de comunicação, prevalece a linguagem científica como característica principal, haja vista que o público interessado neste tipo de comunicação é aquele caracterizado pela formação técnico-científica.

Neste sentido, podemos afirmar que a comunicação científica é viabilizada também através de repositórios digitais, haja vista que, de acordo com o que pontuam Freitas e Leite (2019), nestes repositórios, os investigadores disseminam informações por meio do compartilhamento de seus resultados de pesquisa. Divulgar resultados de investigação é fundamental para expandir a visibilidade e impacto da pesquisa científica.

Vale destacar que a comunicação científica é extremamente influenciada pelas tecnologias da informação e comunicação, bem como pelos esforços científicos em prol do acesso aberto, haja vista que estas questões impactaram nos meios de produção, divulgação e utilização do conhecimento no sistema de comunicação científica e colaboraram para maior rapidez no fazer científico, como também para expansão do conhecimento por meio de ferramentas que derrubaram barreiras de acesso ao conhecimento científico (SILVA; LAZZARIN, 2019).

Portanto, as tecnologias da informação e comunicação beneficiam as práticas associadas à comunicação científica. No contexto de acesso aberto, estas tecnologias são utilizadas, inclusive, para preservação de arquivos gerados a partir de produções científicas, os quais, quando compartilhados de forma aberta, podem ser reutilizados em pesquisas posteriores e, assim, contribuir para o desenvolvimento científico (MIRANDA; SHINTAKU, FIRME, 2017). Além disto, estas tecnologias proporcionam outras possibilidades que beneficiam a comunicação científica, tais quais: servidores em nuvem utilizados como recursos tecnológicos entre os pesquisadores para práticas de comunicação (SANTOS, 2016) e softwares abertos que ampliam as colaborações entre os pesquisadores através de funcionalidades tecnológicas usufruídas para formalizar a comunicação no âmbito científico (SILVA; SILVEIRA, 2019).

Uma vez que a comunicação científica adota, em seus processos, questões tecnológicas, esta possui relação com dados, pois, de acordo com o que menciona Souto (2007) todo e qualquer meio

de comunicação que usufrua de tecnologia deve, necessariamente, adotar medidas que garantam medidas de segurança no que diz respeito aos dados utilizados para a comunicação.

Compreendemos que a relação existente entre comunicação científica e dados ocorre atrelada aos interesses científicos e possibilidade de geração de novos saberes, visto que, conforme afirmado por Russo, Santos e Santos (2001 p. 01) “os cientistas necessitam divulgar suas pesquisas para que os dados manipulados se transformem em informação, que por sua vez vai gerar conhecimento.”.

Medeiros e Caregnato (2012) mencionam que o compartilhamento de dados no âmbito científico faz com que a colaboração entre pesquisadores seja estimulada por meio de ambientes nos quais a disponibilização de resultados de investigações promovam o avanço científico e tecnológico.

Discorrendo sobre comunicação científica e curadoria digital, Sales e Sayão (2012, p. 131) pontuam que

[...] a curadoria digital reduz o ciclo da comunicação científica quando oferece aos pesquisadores dados de pesquisa prontos e confiáveis para o reuso. Isso significa dizer que a curadoria disponibiliza dados tratados, acompanhados por metadados semânticos e estruturais – que assegurariam o seu significado e a reconstrução de sua apresentação correta - e metadados de preservação - que mantêm sua integridade, precisão e autenticidade desses dados.

Ao analisarmos os apontamentos realizados por Sales e Sayão (2012), percebemos que o compartilhamento de metadados de pesquisa também é outro fator colaborativo na comunicação científica. Sobre isto, Bentancourt e Rocha (2012) acrescentam que a transmissão de metadados em condições qualificáveis colabora para eficiência na comunicação científica. Desta forma, metadados de qualidade possuem as seguintes características: corretos, acessíveis, completos, formato adequado (BENTANCOURT; ROCHA, 2012).

Uma das tecnologias associadas à comunicação e dados científicos é o DDI Codebook. Este trata-se de padronizações que são capazes de tornar os dados interativos, os quais, por sua vez proporcionam benefícios relacionados à comunicação científica, através de leitura de informações, navegação em textos complementares, filtro de dados relevantes, dentre outras questões (PEDROZA; BEZERRA; NICOLAU, 2013; SANTOS FILHO, 2016; CURTY, 2019).

A relação que se dá entre a comunicação científica e os RD é algo que ocorre em consequência da forma em que os pesquisadores científicos estão realizando suas investigações na atualidade. Os RD proporcionam a criação de datasets e estes se caracterizam como beneficiadores da comunicação entre os investigadores, uma vez que, para realização de pesquisas científicas, os pesquisadores dependem cada vez mais de dados e tecnologias associadas a estes (FOX; HENDLER,

2011). Por meio destes datasets, a pesquisa científica e desenvolvimento da ciência ocorrem de forma mais eficiente, haja vista que nestes há o intercâmbio de informações associadas às produções científicas entre as instituições (CARVALHO, 2018).

O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT, 2018, p. [--]) define RD e classifica-o:

são bases de dados online que reúnem de maneira organizada a produção científica de uma instituição ou área temática [...] armazenam arquivos de diversos formatos [...] resultam em uma série de benefícios tanto para os pesquisadores quanto às instituições ou sociedades científicas, proporcionam maior visibilidade aos resultados de pesquisas e possibilitam a preservação da memória científica de sua instituição. Os RD podem ser institucionais ou temáticos. Os repositórios institucionais lidam com a produção científica de uma determinada instituição. Os repositórios temáticos com a produção científica de uma determinada área, sem limites institucionais.

## 2.2 Procedimentos Metodológicos

A pesquisa em questão é caracterizada conforme os objetivos em exploratória. De acordo com Gil (2017) essa forma de pesquisa permite maior afinidade com a questão problema, tornando-a mais explícita. Ela ainda aprimora ideias ou proporciona a descoberta de intuições.

Tendo em vista que a presente pesquisa busca descrever de que modo o Dataverse pode colaborar para eficiência na comunicação científica, este estudo classifica-se, ainda, como descritivo. De acordo com o que afirmam Farias e Arruda (2015), na pesquisa descritiva, o investigador limita-se a registrar e descrever fatos que foram observados, sem que haja interferência nestes, com o objetivo de descrever características de dado grupo ou fenômeno. Neste tipo de pesquisa, ainda conforme os autores, também são identificadas possíveis correlações entre variáveis.

Quanto aos procedimentos técnicos utilizados para coleta de dados, essa pesquisa é classificada como bibliográfica, desenvolvida a partir de materiais já elaborados envolvendo a temática comunicação científica para, posteriormente, correlacionar os dados bibliográficos obtidos com as características e funcionalidades do RD Dataverse. Conforme Marconi e Lakatos (2017, p. 236) “a pesquisa bibliográfica não é mera repetição do que já foi dito ou escrito sobre certo assunto, mas propicia o exame de um tema sob um novo enfoque ou abordagem, chegando a conclusões inovadoras”.

Foram utilizados estudos de caráter científico (livros, artigos científicos, trabalhos de conclusão de curso, dissertações, teses, trabalhos publicados em eventos) que abordassem questões referentes à comunicação científica, bem como a dados, repositórios e dados e as

contribuições dos mesmos para as práticas e processos relacionados à comunicação no âmbito científico.

Buscou-se correlacionar os dados e informações obtidas nos estudos utilizados com as informações contidas nos sites “The Dataverse Project” – site de apresentação do Dataverse - e “Harvard Dataverse” – RD propriamente dito, utilizado quando o usuário decide acessar os bancos de dados depositados e/ou criar seu próprio banco de dados – para que pudéssemos identificar os fatores que tornam o Dataverse um instrumento que pode beneficiar a comunicação no âmbito científico.

### 2.3 Resultados e Discussão

No Dataverse a preservação e acesso a longo prazo estão garantidos na identificação persistente, protegendo os documentos digitais com mecanismos que preveem a obsolescência dos dados - migração dos dados para um software mais recente e a prescrição que consiste em guardar o conjunto de bytes para serem consultados quando for necessário buscando (THE DATAVERSE PROJECT, 2018, tradução nossa).

Desta forma, percebe-se que uma das contribuições do Dataverse para a comunicação científica é justamente a preservação dos dados armazenados neste RD, visto que, de acordo com o que pontuam Sales e Sayão (2012), disponibilizar os dados científicos obtidos por meio de investigações é um dos principais pontos de relevância para promoção de comunicação científica.

O RD em questão armazena o conteúdo de dados brutos extraídos de arquivos em texto simples, arquivos delimitados por TAB. As informações de metadados que descrevem esse conteúdo são armazenadas separadamente, em um banco de dados relacional, para que possam ser acessadas de maneira eficiente pelo aplicativo (HARVARD DATAVERSE, 2018).

Sendo assim, a preocupação do Dataverse em relação à disponibilização de metadados eficientes pode ser considerada outra questão que colabora para eficiência na comunicação científica, haja vista que conforme os apontamentos de Bentancourt e Rocha (2012), a produção de metadados com qualidade é essencial para que a comunicação científica tenha maior visibilidade, uma vez que, atualmente, o modelo atual das revistas eletrônicas científicas possui foco na interoperabilidade dos metadados.

Para fins de preservação de arquivos, pode-se exportar, em arquivos XML de texto simples, usando um formato padronizado e aberto de DDI Codebook (HARVARD DATAVERSE, 2018, tradução nossa). Portanto, outra colaboração do Dataverse para comunicação científica é sua capacidade de

preservação de arquivos, visto que, como mencionado por Miranda, Shintaku e Firme (2017), as questões tecnológicas associadas às práticas de comunicações científicas visam, dentre outras questões, o acesso aberto aos arquivos referentes às publicações da comunidade científica.

O DDI Codebook ou Iniciativa de Documentação de Dados (também conhecida como DDI ou DDI Metadata – escrita XML) é um padrão internacional, livremente disponível, para descrever pesquisas, questionários, arquivos de dados estatísticos e informações em nível de estudo em ciências sociais, comportamentais, econômicas e de saúde. Este formato fornece um formato para conteúdo, troca e preservação de informações de questionário e arquivo de dados (DATA DOCUMENTATION INITIATIVE, 2018, tradução nossa).

Ainda no que diz respeito ao padrão DDI Codebook, Curty (2019) comenta que este é considerado como um padrão para elaborar um tipo de dicionário de dados interativos. Portanto, percebe-se que o DDI, visto que, a partir da visualização de dados interativos, os processos de comunicação são melhorados, uma vez que este tipo de dado viabiliza leitura de informação, filtragem de dados interessantes e navegação em textos complementares (PEDROZA, BEZERRA, NICOLAU, 2013; SANTOS FILHO, 2016).

Como forma de custodiar os dados, o Dataverse possui um esquema de metadados que facilita a recuperação da informação, além da existência de subconjuntos de dados denominados de datasets, que podem proporcionar uma veracidade aos dados que os pesquisadores estão buscando (THE DATAVERSE PROJECT, 2018, tradução nossa).

Um dos desafios associados à pesquisa científica moderna é o desenvolvimento de tecnologias para melhorias semânticas. Como as ciências modernas estão cada vez mais dependentes dos dados, há a necessidade de criação de datasets para a comunidade científica (FOX; HENDLER, 2011). Portanto, é notória a relevância dos datasets para a comunicação científica, uma vez que estes comprovam a veracidade de dados que podem ser reutilizados em pesquisas futuras. No Quadro 01, apresenta-se um exemplo de como o Dataverse é utilizado na Universidade Federal da Paraíba.

**Quadro 01** – Identificação do Dataverse UFPB.

|             |   |
|-------------|---|
| Dataverse   | Departamento de CI  |
| Identifier  | <a href="https://dataverse.ufpb.br/dataverse/dci">https://dataverse.ufpb.br/dataverse/dci</a>   |
| Category    | Department  |
| Affiliation | CCSA  |
| Description | Este repositório concentra dados de pesquisas do Departamento de Ciências da Informação (DCI) do Centro de Ciências Sociais Aplicadas (CCSA) da Universidade Federal da Paraíba (UFPB). |

**Fonte:** Barroso *et al.*, 2018, p. [--].



No campo “Dataverse” deve-se descrever o título do banco de dados, o departamento, a universidade ou a revista que conterá os dados. O “Identifier” refere-se a um nome curto a ser utilizado na URL (Uniform Resource Locator - Localizador Padrão de Recursos - formato de atribuição universal para designar um recurso na Internet). A “category” identifica o tipo de grupo da qual a base de dados pertence. O campo “Affiliation” deve compor algum nome ou sigla associado ao repositório criado. A descrição resumida do Dataverse deve ser feita no campo “Description” para exibição na página do repositório (THE DATAVERSE PROJECT, 2018, tradução nossa).

Portanto, a partir do exemplo apresentado no Quadro 01, percebe-se que o Dataverse proporciona comunicação científica entre departamentos, centros e instituições de ensino. Carvalho (2018) reafirma a importância do intercâmbio informacional entre instituições, ressaltando que a utilização de informações compartilhada facilita a comunicação científica entre as instituições de ensino e, conseqüentemente, colabora para que a ciência produza de maneira mais eficiente.

Os campos de metadados são escolhidos para uso em modelos de conjunto de dados a serem adicionados ao Dataverse criado. É possível, dentre as opções de padrões de metadados, a seleção de grupos específicos de domínio de acordo com as áreas do conhecimento. Quando os pesquisadores vão acessar um determinado banco de dados, também se faz necessário a criação de outros metadados – os metadados de citação - que deve conter informações relevantes para gerar uma citação dos dados como referência. Os principais campos a serem preenchidos são: título, autores, contato, descrição resumida da pesquisa, data da coleta dos dados no RD, tópicos da pesquisa, palavras-chave e outros (THE DATAVERSE PROJECT, 2018, tradução nossa).

Sendo assim, disponibilizar dados em condições de citação seria outro benefício proporcionado pelo Dataverse em relação à eficiência da comunicação científica, uma vez que, de acordo com o que afirmam Sayão e Sales (2012), compartilhar dados que possam ser citados e reutilizados são alguns dos fatores que adicionam velocidade ao ciclo de comunicação científica.

O acesso aos dados neste RD ocorre de quatro formas: aberto, fechado, gerenciado e moderado. Para o gerenciamento de acesso e segurança, buscam-se restringir o acesso aos dados para pesquisadores previamente cadastrados, cujas credenciais de acesso podem ser feitas a partir do cadastro individual do pesquisador no Dataverse utilizando uma conta do Gmail. O objetivo é manter um controle absoluto de acesso aos dados disponibilizados para estatísticas e elevar a segurança. Neste contexto, três importantes configurações garantem a transparência em relação à política de acesso aos dados: termos de uso, de acesso e Guestbooks (“livro de visitas”) (THE DATAVERSE PROJECT, 2018, tradução nossa).

Portanto, as medidas adotadas no Dataverse visando à segurança dos dados se caracterizam como outro fator que beneficia a comunicação científica, pois, segundo o que afirma Souto (2007), todos os processos de comunicações que envolvem questões tecnológicas devem, necessariamente, adotar servidores robustos e sistemas que sejam eficientes quanto à segurança de dados.

Para a preservação permanente dos dados depositados pelos pesquisadores no Dataverse existe a incorporação e integração com o DataCite. Quando os conjuntos de dados são publicados, os metadados são enviados para o DataCite. Além de um de suporte chamado O Protocolo da Open Archives Initiative for Metadata Harvesting (OAI-PMH), que facilita a coleta de conjuntos de dados de um sistema para outro (HARVARD DATAVERSE, 2018, tradução nossa).

O RD Dataverse fica disponível como um servidor em nuvem e software de código aberto (RICE; SOUTHAL, 2016). Estas são mais duas formas pelas quais o Dataverse contribui para eficiência na comunicação científica, visto que a computação em nuvem tem se tornado tendência para pesquisadores e cientistas aproveitarem recursos computacionais em práticas de comunicação científica (SANTOS, 2016) e, levando em consideração que os softwares abertos proporcionam maior ampliação na colaboração científica por meio de uma série de elementos tecnológicos para formalização da comunicação científica (SILVA; SILVEIRA, 2019).

O repositório em questão, de acordo com The Dataverse Project (2018, tradução nossa), tem se ampliado cada vez mais pela sua visibilidade na WEB, e o crédito acadêmico que proporciona o aumento de contagem das citações com referência a cada dado científico que este repositório custodia. No sítio eletrônico do Projeto Dataverse são caracterizadas quatro categorias de públicos: pesquisadores, revistas, instituições e desenvolvedores.

### 3 CONCLUSÕES

O avanço da tecnologia da informação e do processo de globalização possibilitaram a integração e o compartilhamento de informações de maneira instantânea, resultando em um aumento de publicações considerado acima do comum. Este crescimento foi auxiliado com o surgimento dos RD institucionais e de pesquisa, que armazenam os resultados de trabalhos de conclusão de curso, dissertações, teses, anais de encontros científicos, palestras e, no caso do Dataverse, dados brutos de pesquisas de todo o mundo.

Após a leitura das referências citadas fica evidente a relevância do Dataverse para a Comunicação Científica. Com aumento da produção de informações em detrimento do pouco

espaço de armazenamento (revistas, livros, eventos, etc), os RD surgem como uma possibilidade para armazenar esses dados “brutos”, descartados em média após cinco anos da data de início da coleta.

O Dataverse é capaz de proporcionar redução de custos em pesquisas e diminuição do tempo gasto para se obter novas informações. Através dos metadados e das estruturas de preservação digital, este RD garante a representação dos dados e o armazenamento seguro dos resultados das pesquisas disponibilizados por seus pesquisadores e/ou instituições.

Como sugestão para pesquisas posteriores, a comunidade científica poderia entrevistar pesquisadores que utilizam o Dataverse, de maneira a identificar como estes investigadores percebem o potencial deste RD para eficiência nos processos de comunicação envolvidos à ciência.

## REFERÊNCIAS

BARBOSA, J. P. M. **UPBox: Armazenamento na Nuvem para Dados de Investigação da U.Porto**. Dissertação (Mestrado Integrado em Engenharia Informática e Computação). Universidade do Porto, 2013.

BARROSO, P. A. L. *et al.* **Gestão de Dados Científicos**: relato de experiência da implementação do conjunto de dados das arboviroses no software Dataverse da Universidade Federal da Paraíba. Artigo Acadêmico – Programa de Pós-Graduação em Ciência da Informação – Universidade Federal da Paraíba. 2018.

BENTANCOURT, S. S. M. P.; ROCHA, R. P. Metadados de Qualidade e Visibilidade na Comunicação Científica. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. esp.2, p. 82-101, 2012.

BJORK, B. Open Access to Scientific Publications - an Analysis of the Barriers to Change? **Information Research**, v. 9, n. 2, January, 2004.

BRANCO, A. K. A. C.; GONZAGA, A. M. As Contribuições da Comunicação Científica para a Pesquisa Qualitativa em Educação: o caso do Projeto Observatório da Educação – POE/CAPES/UEA. **CIAIQ2014**, v. 1, 2014.

CARVALHO, L. S. **A Contribuição da Web Semântica para a Biblioteconomia**. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia). Universidade Federal da Paraíba. João Pessoa, 2018.

CUNHA, M. B.; CAVALCANTI, C. R. O. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008.

CURTY, R. Abordagens de Reuso e a Questão da Reusabilidade dos Dados Científicos. **Liinc em Revista**, v. 15, n. 2, p. 177-193, 2019.

CURTY, R. As Diferentes Dimensões do Reuso de Dados Científicos. **Anais do XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB)**, 2017.

DATAVERSE. **The Dataverse Project**. [S.l.], 2020. Disponível em: <https://dataverse.org>. Acesso em: 24 jun. 2020.

FARIAS, M.; ARRUDA, E. **Planejamento da Pesquisa Científica**. São Paulo: Atlas, 2015

FERREIRA, L. S. **O Uso do Software Zotero no Processo de Comunicação Científica**: um estudo de caso na ENSP/FIOCRUZ. Dissertação (Mestrado Profissional em Biblioteconomia). Universidade Federal do Estado do Rio de Janeiro. Rio de Janeiro, 2017.

FORMENTON, D. **Identificação de Padrões de Metadados para Preservação Digital**. São Carlos: UFSCar, 2015. 102 f. Dissertação de Mestrado - Universidade Federal de São Carlos.

FOX, P.; HENDLER, J. e-Science Semântica: o significado codificado na próxima geração de ciência digitalmente apropriada. In: HEY, T.; STEWARD, T.; TOLLE, K. (Org.). **O Quarto Paradigma**: descobertas científicas na era da e-Science. Tradução Leda Beck. São Paulo: Oficina de textos, 2011. p. 159-163.

FREITAS, M. A.; LEITE, F. C. L.; La Percepción de Investigadores Brasileños Sobre el Depósito de la Producción Científica en Repositorios Institucionales de Acceso Abierto. **Revista Interamericana de Bibliotecología**, Medellín, v. 42. n. 2, p. 159-173, 2019.

GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 6 ed. São Paulo: Atlas, 2017.

GONZÁLEZ DE GÓMEZ, M. N. Novos Cenários Políticos para a Informação. **Ciência da Informação, Brasília**, v.31, n.1, p. 27-40, jan./abr. 2002.

**HARVARD DATAVERSE**. Disponível em: <https://dataverse.harvard.edu>. Acesso em: agosto/2018.

IBICT. **Repositórios Digitais**. Disponível em: <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais/sobre-repositorios-digitais>. Acesso em: agosto/2018.

MARCONI, M. A; LAKATOS, E. M. **Metodologia Científica**. 7 ed. São Paulo: Atlas, 2017.

MEDEIROS, J. S.; CAREGNATO, S. E.; Compartilhamento de Dados e e-Science: explorando um novo conceito para a comunicação científica. **Liinc em Revista**, v. 8, n.2, p. 311-322, 2012.

MIRANDA, A. C. D.; SHINTAKU, M.; FIRME, S. M. Critérios para Avaliação de Ferramentas para Repositório Conforme a Ciência Aberta. **Tear: Revista de Educação, Ciência e Tecnologia**, v. 6, p. 01-15, 2017.

PARANGABA, C. R. Catálogo de dados dos trabalhos científicos de gestão ambiental e saúde da Escola Nacional de Saúde Pública Sérgio Arouca (ENSP/FIOCRUZ): Proposta. **Revista Informação na Sociedade Contemporânea**, v. 1, p. 1-19, 2017.

PEDROZA, N. P.; BEZERRA, E. P.; NICOLAU, M. Ferramentas para elaboração de infográficos: um estudo de caso. **Anais do Encontro Virtual de Documentação em Software Livre e Congresso Internacional de Linguagem e Tecnologia Online**, 2013.

PEREIRA, S. D. O. **A Leitura Infantil na Perspectiva da Ciência da Informação**. Trabalho de Conclusão de Curso. (Bacharelado em Biblioteconomia e Documentação). Universidade Federal de Sergipe. São Cristóvão, 2019.

RICE, R; SOUTHALL, J. **The Data Librarian's Handbook**, London: Facet, 2016. 169 p.

RUSSO, M.; SANTOS, E. T. G.; SANTOS, M. J. V.C. Produção Científica Brasileira: da comunicação à recuperação via WEB. **Fórum Nacional de Padronização e Divulgação da Produção Científica**, v.1, 2001.

SALES, L. F.; SAYÃO, L. F. O Impacto da Curadoria Digital dos Dados de Pesquisa na Comunicação Científica. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 17, n. 2, p. 118-135, 2012.

SANTOS, D. B. G. **O Conhecimento e a Pesquisa nas Nuvens**: uma pesquisa social sobre a aplicação das práticas de gestão do conhecimento associadas às tecnologias de computação em nuvem nos ambientes de pesquisa. Tese (Doutorado em Ciências). Universidade de São Paulo. São Paulo, 2016.

SANTOS FILHO, J. G. **Análise de Logs da Plataforma Moodle Utilizando Técnicas de Mineração de Dados**: um estudo de caso. Trabalho de Conclusão de Curso (Licenciatura em Ciências da Computação) Universidade Federal da Paraíba. Rio Tinto, 2016.

SAYÃO, L. F.; SALES, L. F. Curadoria Digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, v. 22, n. 3, p. 179-191, 2012.

SAYÃO, L. F.; SALES, L. F. Algumas Considerações Sobre os Repositórios de Dados de Pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p. 90-115, maio/agosto, 2016.

SHINTAKU, M.; MEIRELES, R. **Manual do DSPACE**: administração de repositórios. Salvador: EDUFBA, 2010.

SILVA, H. S.; LAZZARIN, F. A. Preprints na Comunicação Científica: proposta de modelo para aceleração do sistema de comunicação científica. **Informação@Profissões**, Londrina, v. 8, n. 2, p. 150-170, 2019.

SILVA, F. C. C.; SILVEIRA, L. O Ecossistema da Ciência Aberta. **Transinformação**, v. 31, p. 01-13, 2019.

SOUTO, P. N. E-publishing Development and Changes in the Scholarly Communication System. **Ciência da Informação**, Brasília, v. 36, n. 1, p. 158-166, 2007. Disponível em: <https://www.scielo.br/pdf/ci/v36n1/a12v36n1.pdf>. Acesso em: 25 jun.2020.

**THE DATAVERSE PROJECT**. Disponível em: <https://dataverse.org/>. Acesso em: junho/2018.

TREVISOL NETO, O.; CAFÉ, L. M. A.; SILVA, E. L.. As Instituições, Produtores e Produtos Científicos do Campo da Moda no Brasil. **Perspectivas em Ciência da Informação**, v. 22, n. 2, p. 152-169, 2017.