



**UNIVERSIDADE FEDERAL DA PARAÍBA– UFPB
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO**



**REPOSITÓRIOS DE DADOS DE PESQUISA: INVESTIGANDO SUA
ADOÇÃO EM INSTITUTOS BRASILEIROS DE PESQUISA**

Pesquisador PIBIC: *André Luiz de França Madeiro*

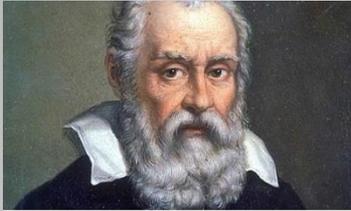
Orientador: *Prof. Dr. Guilherme Ataíde Dias*

JOÃO PESSOA

2018

INTRODUÇÃO

Galileu Galilei



- **Século XVII:** divulgação de atividades científicas (MASSARANI e MOREIRA, 2004)
- **Século XVIII:** *status* da ciência; criação da *Royal Institution* (ARAMBURÚ e PÉREZ, 2016)
- **Século XIX:** divulgação da ciência a nível mundial (MASSARANI e MOREIRA, 2004)
- **Século XX:** Albert Einstein e Marie Curie (MASSARANI e MOREIRA, 2004)
- **Final do século XX:** redes de comunicação entre computadores e gerenciamento de dados científicos coletados *online* (COSTA, 2017)

INTRODUÇÃO

- A ciência atual no contexto da *e-Science* (OLIVEIRA, 2016)
- Apresentação de um comparativo (Quadro 1) entre a *e-Science* e a pesquisa tradicional, conforme Medeiros e Caregnato (2012):

Características	<i>e-Science</i>	Pesquisa Tradicional
Participantes	Diversamente qualificados, equipe de pesquisa distribuída	Pesquisador individual ou pequena equipe local de pesquisa
Dados	Gerados, armazenados e acessíveis de localizações distribuídas	Gerados, armazenados e acessíveis apenas localmente
Computação e Instrumentação	Larga escala ou sob demanda ou acesso à informação compartilhada	Emprego da computação em lote ou emprego do computador ou instrumentos do próprio pesquisador
Rede	Confiança na Internet e em mediadores	Não confiante na Internet
Disseminação da Pesquisa	Via <i>websites</i> , portais especializados e repositórios	Via publicações impressas ou apresentações em conferências

FONTE: Adaptado de MEDEIROS e CAREGNATO (2012)

PROBLEMATIZAÇÃO

- A partir destas considerações iniciais, questionamos: **de que forma se configuram os repositórios de pesquisa das instituições brasileiras de pesquisa no que diz respeito ao suporte para a gestão de dados de pesquisa?**

OBJETIVOS

- OBJETIVO GERAL

- investigar de que maneira os repositórios de instituições de pesquisa se configuram no suporte para a gestão de dados de pesquisa.

OBJETIVOS

- OBJETIVOS ESPECÍFICOS:
 - Identificar os repositórios de dados de institutos brasileiros de pesquisa;
 - identificar a tipologia dos repositórios em institutos brasileiros de pesquisa;
 - Levantar características específicas dos repositórios de dados de institutos brasileiros de pesquisa;
 - Disponibilizar os dados coletados e material de apoio no repositório de dados no repositório *Dataverse* da UFPB.

A QUESTÃO DOS DADOS NO CONTEXTO DA CIÊNCIA CONTEMPORÂNEA

Dados	Dados de Pesquisa	Coleção de Dados
<ul style="list-style-type: none">(SETZER, 2015, <i>online</i>).	<ul style="list-style-type: none">Sayão e Sales (2015);Rodrigues et al. (2010).	<ul style="list-style-type: none">Simberloff et al (2005);Rodrigues et al. (2010).

A QUESTÃO DOS DADOS NO CONTEXTO DA CIÊNCIA CONTEMPORÂNEA

Classificação dos Dados

Quanto a origem	Quanto ao <i>status</i>
<p>Observacionais, Computacionais e Experimentais (SIMBERLOFF et al, 2005);</p> <p>Canônicos ou dados de referência e Dados de material de acompanhamento Green; Mcdonald e Rice (2009);</p>	<p>Brutos, Derivados ((GREEN; MCDONALD; RICE, 2009).</p>

A QUESTÃO DOS DADOS NO CONTEXTO DA CIÊNCIA CONTEMPORÂNEA

Repositórios Digitais	Classificação dos Repositórios Digitais	Classificação de Repositórios de Dados de Pesquisa
IBICT (2018, <i>online</i>); Camargo e Vidotti (2009).	institucionais - <i>E-LIS - EPrints in Library and Information Science</i> ; temáticos - <i>e-Prints Soton - repositório de Pesquisa da Universidade de Southampton</i> (LEITE et al., 2012, <i>online</i>).	institucionais - Edinburgh DataShare; disciplinares – <i>GenBank, PANGAEA</i> ; multidisciplinares – <i>Figshare</i> ; de projetos científicos - <i>Scientific Drilling Database (SDDB)</i> (PAMPEL et al., 2013, <i>online</i>).

- Definição de *e-Science* e *cyber infrastructure*. (OLIVEIRA, 2016)
- Etapas do Ciclo de Vida de dados: 1.Planejar; 2.Coletar; 3.Assegurar; 4.Descrever; 5.Preservar; 6.Descobrir; 7.Integrar; 8.Analisar (STRASSER et al., 2018, *online*).

PROCEDIMENTOS METODOLÓGICOS

- O percurso metodológico usado para condução desta investigação baseou-se numa abordagem bibliográfica, de cunho exploratório, quanti-qualitativo.
- Os descritores selecionados para a pesquisa foram: “repositórios brasileiros”; “repositórios de dados”; “data repositories” e “research repositories”, foram selecionados os sites: *R3DATA*, *ROAR* (*Registry of Open Access Repositories*) e *IBICT*. Os endereços Web (URL) dos sites escolhidos, respectivamente, foram:
<https://www.re3data.org/>; <http://roar.eprints.org/>;
<http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais/repositorios-brasileiros>
- Para auxiliar na identificação de repositórios de dados de pesquisa, foi utilizado os estudos de Leite et al. (2012), Pampel et al. (2013) e Sales (2014) como referências teóricas.

APRESENTAÇÃO DOS DADOS E DISCUSSÃO DOS RESULTADOS

- Foram encontrados três repositórios de dados em instituições brasileiras de pesquisa. Dentre os repositórios identificados como repositórios de dados, listamos os seguintes: *PPBiodata*, *IBICT Dataverse* e *carpeDIEN*.
- O repositório de dados científicos *IBICT Dataverse* dispõe de uma tipologia orientada para vários campos científicos. O mesmo usa o software *Dataverse*;
- *CarpeDIEN* é o nome do repositório do *IEN (Instituto de Energia Nuclear)*. Caracteriza-se por preservar dados científicos a respeito da energia nuclear” Este usa o software *Dspace*;
- *PPBiodata* é o nome do site do *PPBio*. Este repositório faz uso do software *metacat*.
- Conforme Pampel et al. (2013) e Sales (2014), respectivamente, os repositórios encontrados estão de acordo com a tipologia *repositórios de dados de pesquisa institucional* e *repositórios de dados*.

APRESENTAÇÃO DOS DADOS E DISCUSSÃO DOS RESULTADOS

- Além dos repositórios expostos, verificamos a existência de mais três, nos seguintes trabalhos científicos: Costa (2017) constatou Portal GEOINFO de infraestrutura de dados espaciais da EMBRAPA e o Portal da Biodiversidade. Peixoto (2012) encontrou o repositório SinBiota (Sistema de Informação Ambiental do Programa Biota/Fapesp).
- Foram encontrados um total de seis repositórios de dados de pesquisa de instituições brasileiras de pesquisa. Conforme Pampel et al. (2013) e Sales (2014), os repositórios encontrados estão de acordo com a tipologia repositórios de dados de pesquisa institucional e repositórios de dados. Há uma consonância entre a fundamentação teórica e a investigação realizada.

CONSIDERAÇÕES FINAIS

- Verifica-se uma heterogeneidade no que diz respeito aos software utilizados para implantar repositórios de dados utilizados, conforme identificada, anteriormente, foram encontrados: *DSpace*, *Metacat* e *Dataverse*.
- Repositórios de dados de instituições brasileiras científicas citados nos seguintes trabalhos científicos:
- A tese *Diretrizes para uma política de gestão de dados científicos no Brasil* de Costa (2017): a) Repositório de Dados do Programa de Pesquisa de Biodiversidade da Amazônia Ocidental (*PPBio*) b) Repositório de Dados do Programa de Pesquisas Ecológicas de Longa Duração (*PELD*) e Portal *GEOINFO* de infraestrutura de dados espaciais da *EMBRAPA* e Portal da Biodiversidade.
- O Trabalho de Conclusão de Curso denominado *Arquitetura de informação em repositório de dados científicos* de Peixoto (2012): *Species Link*, SinBiota (Sistema de Informação Ambiental do Programa Biota/Fapesp) e *PELD* (Programa de Pesquisas Ecológicas de Longa Duração) Sítio 1: Floresta Amazônica – Manaus.
- Avaliando estes dados na perspectiva de Pampel et al (2013) e Sales (2014), somente o termo *Species Link* não caracteriza um repositório de dados de pesquisa. O repositório *PELD* não foi encontrado na Internet.
- A partir da pesquisa realizada, concluímos que o número de repositórios ainda é insipiente.

REFERÊNCIAS

- ARAMBURU, M.R.; PÉREZ, R.M. **Unpaseo por la historia de la divulgación científica en España y en la comunidad de Madrid**: análisis de la situación actual de la divulgación medioambiental. In: **Tecnología y Desarrollo**: Revista de Ciencia, Tecnología y Medio Ambiente. Volumen XIV. AÑO 2016. Disponível em: <<https://www.uax.es/publicacion/un-paseo-por-la-historia-de-la-divulgacion-cientifica-en-espana-y-en-la.pdf>> Acesso em: 14 jun. 2018.
- CAMARGO, L.S. de A. de; VIDOTTI, S.A.B.G. **Arquitetura da informação para repositórios científicos digitais**. In: SAYÃO, L.F. et al. (Orgs). **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. Disponível em: <https://repositorio.ufba.br/ri/bitstream/ufba/473/3/implantacao_repositorio_web.pdf> Acesso em: 30 jul. 2018.
- CarpeDIEN. Disponível em: <<http://carpedien.ien.gov.br/>> Acesso em: 06 jun. 2018.
- COSTA, M.M. **Diretrizes para uma política de gestão de dados científicos no Brasil**. Brasília: 2017. 288p. Tese (Doutorado em Ciências da Informação). Faculdade de Ciência da Informação. Universidade de Brasília. Disponível em: <<http://repositorio.unb.br/handle/10482/24895>> Acesso em: 24 abr. 2018.
- DSpace. Disponível em: <<http://www.duraspace.org/dspace/about/>> Acesso em: 03 jun. 2018.
- Dataverse. Disponível em: <<https://dataverse.org/about>> Acesso em: 03 jun. 2018.
- FRICKÉ, M. **Big Data and Its Epistemology**. Journal of the Association for Information Science and Technology, 66 (4): 651 – 661, 2015.
- GREEN, Ann; MACDONALD, Stuart; RICE, Robin. **Policy-making for research data in Repositories**: a guide. Disponível em: <<http://www.disc-uk.org/docs/guide.pdf>> Acesso em: 08 maio. 2018.
- IBICT DATAVERSE. Disponível em: <<https://repositoriopesquisas.ibict.br/>> Acesso em: 06 jun. 2018.
- IBICT . **IBICT lança Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã**. Disponível em: <<http://www.ibict.br/Sala-de-Imprensa/noticias/2016/ibict-lanca-manifesto-de-acesso-aberto-a-dados-da-pesquisa-brasileira-para-ciencia-cidada>> Acesso em: 29 jun. 2018.

REFERÊNCIAS

IBICT. **Sobre repositórios digitais**. Disponível em:< <http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais>> Acesso em: 01 ago. 2018.

IERC (**European Research Cluster on the Internet of things**). Disponível em:<http://www.internet-of-things-research.eu/about_iot.htm> Acesso em: 16 abr. 2018.

KNB –**The Knowledge Network for Biocomplexity**. Disponível em:<<https://knb.ecoinformatics.org/knb/docs/intro.html#what-s-in-this-guide>> Acesso em: 03 jun. 2018.

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto / Fernando César Lima Leite. – Brasília: Ibict, 2009. Disponível em:<http://www.repositorio.unb.br/bitstream/10482/4841/1/LIVRO_ComoAmpliareGerenciar.pdf>Acesso em: 28 mar. 2018.

LEITE, F.et al. **Repositórios Institucionais**: boas práticas para a construção de repositórios institucionais da produção científica. Brasília: IBICT, 2012. Disponível em:<<http://livroaberto.ibict.br/bitstream/1/703/1/Boas%20pr%C3%A1ticas%20para%20a%20constru%C3%A7%C3%A3o%20de%20reposit%C3%B3rios%20institucionais%20da%20produ%C3%A7%C3%A3o%20cient%C3%ADfica.pdf>> Acesso em: 10 jun. 2018.

MASSARANI, L.; MOREIRA, I. de C. **Divulgación de la Ciencia: Perspectivas Históricas y Dilemas Permanentes**. In: **Revista Quark**, Barcelona: Observatorio de la Comunicación Científica (OCC) da Universitat Pompeu número 32, abril-junio 2004. Disponível em:< <http://quark.prbb.org/32/032030.pdf>> Acesso em: 29 abr. 2018.

MEDEIROS, J. da S.; CAREGNATO, S.E. **Compartilhamento de dados e e-Science**: explorando um novo conceito para a comunicação científica. In: Liinc em Revista, v.8, n.2, setembro, 2012, Rio de Janeiro, p. 311-322. Disponível em:< <http://revista.ibict.br/liinc/article/view/3368/2969> > Acesso em: 23 jul. 2018.

OLIVEIRA, A. C. S. **Desvendando a autoralidade colaborativa na e-Science sob a ótica dos direitos da propriedade intelectual**. João Pessoa: 2016. 297p. Tese (Doutorado em Ciências da Informação). Centro de Ciências Sociais Aplicadas. Universidade Federal da Paraíba. Disponível em:<<http://tede.biblioteca.ufpb.br/handle/tede/8849>> Acesso em: 7 jun 2018;

REFERÊNCIAS

PAMPEL, H. et al. **Making research data repositories visible**: the re3data.orgRegistry. PLoSOne, San Francisco, v. 8, n.11, 2013. Disponível em:<<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0078080>>. Acesso em: 19 maio. 2018.

PEIXOTO, N.M. **Arquitetura de informação em repositório de dados científicos**: análise da interface do repositório do PELD: Programa de Pesquisa Ecológicas de Longa Duração. 2012. Porto Alegre: 2012. 78p. Trabalho de Conclusão de Curso (Graduação) Universidade Federal do Rio Grande do Sul. Faculdade de Biblioteconomia e Comunicação. Curso de Biblioteconomia. Disponível em:< <https://lume.ufrgs.br/handle/10183/69727>> Acesso em: 29 jul. 2018.

RODRIGUES, Eloi et al. **Os repositórios de dados científicos: estado da arte**. 2010. (Relatório D-24: RCAAP). Disponível em: <<https://repositorio-aberto.up.pt/bitstream/10216/23806/2/44632.pdf>>Acesso em: 18 abr. 2018.

SALES, L.F. **Integração semântica de publicações científicas e dados de pesquisa**: proposta de modelo de publicação ampliada para a área de Ciências Nucleares. Rio de Janeiro: 2014. 265p. Tese apresentada ao Programa de Pós-graduação em Ciência da Informação da Universidade Federal do Rio de Janeiro/. Disponível em:<<http://repositorio.ibict.br/bitstream/123456789/874/1/LUANA%20SALES%20D.pdf>> Acesso em: 05 dez. 2017.

SAYÃO, L.F. **O Papel dos Repositórios Digitais na Curadoria de Dados de Pesquisa**. In: VECHIATO, F. et al. (Orgs) Repositórios Digitais: Teoria e Prática. Curitiba: EDUTFPR, 2017. Disponível em:<<https://repositorio.ufrn.br/jspui/handle/123456789/24189>> Acesso em: 16 abr. 2018.

SAYÃO, L.F.; SALES, L.F. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**: Rio de Janeiro: CNEN, 2015. Disponível em:< <http://www.aben.com.br/Arquivos/420/420.pdf> >Acesso em: 04 maio. 2018.

REFERÊNCIAS

SAYÃO, L.F.; SALES, L.F. **Algumas considerações sobre os repositórios digitais de dados de pesquisa.** *Inf. Inf.*, Londrina, v. 21, n. 2, p. 90 – 115, maio/ago 2016. Disponível em:<<http://www.uel.br/revistas/uel/index.php/informacao/article/view/27939/20122>>Acesso em: 18 abr. 2018.

SETZER, V.W. **Dado, Informação, Conhecimento e Competência.** Versão de 25/05/2015. Disponível em:<<https://www.ime.usp.br/~vwsetzer/dado-info.html>>Acesso em: 05 dez. 2017.

SIMBERLOFF, D. et al. **National Science Board. Long-Lived Digital Data Collections:** Enabling Research and Education in the 21st Century. National Science Board, National Science Foundation, 2005. Disponível em:<<https://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf>>Acesso em: 15 abr. 2018.

STRASSER, Carly et al. **Primer on Data Management:** What you always wanted to know. California: CDL, 2012 In: SAYÃO, L.F.; SALES, L.F. Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores: Rio de Janeiro: CNEN, 2015. Disponível em:< https://www.dataone.org/sites/all/documents/DataONE_BP_Primer_020212.pdf >Acesso em: 12 maio. 2018.

OBRIQADO!

andreluizfmadeiro@gmail.com