

UNIVERSIDADE FEDERAL DO RIO GRANDE
CENTRO DE CIÊNCIAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO
CURSO DE MESTRADO EM ENGENHARIA DE COMPUTAÇÃO

Dissertação de Mestrado

Proposta de inspeção de usabilidade de um repositório de dados de pesquisa brasileiro

Luis Alberto Barbosa Azambuja

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal do Rio Grande, como requisito parcial para a obtenção do grau de Mestre em Engenharia de Computação

Orientador: Prof. Dr. Eduardo Nunes Borges

Rio Grande, 2019

Ficha catalográfica

A991p Azambuja, Luis Alberto Barbosa.
Proposta de inspeção de usabilidade de um repositório de dados de pesquisa brasileiro / Luis Alberto Barbosa Azambuja . – 2019.
103 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Computação, Rio Grande/RS, 2019.

Orientador: Dr. Eduardo Nunes Borges.

1. Dados Abertos 2. Acesso Aberto a Dados de Pesquisa
3. Repositório de Dados 4. Usabilidade I. Borges, Eduardo Nunes
II. Título.

CDU 02:004

Banca examinadora:

Prof. Dr. Eduardo Nunes Borges (presidente)
Universidade Federal do Rio Grande

Prof^a. Dr^a. Graçaliz Dimuro
Universidade Federal do Rio Grande

Prof. Dr. Daniel Lichtnow
Universidade Federal de Santa Maria

RESUMO

AZAMBUJA, Luis Alberto Barbosa. **Proposta de inspeção de usabilidade de um repositório de dados de pesquisa brasileiro**. 2019. 103 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal do Rio Grande, Rio Grande.

Nos últimos anos, a evolução dos sistemas *Web* proporcionaram o surgimento de iniciativas de promoção do Acesso Aberto a Dados de Pesquisa em todo o mundo. Dentre as principais motivações para que os pesquisadores disponibilizem dados resultantes de projetos de pesquisa em plataformas de acesso aberto destacam-se aumentar a visibilidade dos grupos de pesquisa, facilitar o intercâmbio de informações, validar resultados e a transparência na prestação de contas para as agências de fomento públicas ou privadas. A escolha de uma ferramenta para esse fim assim como a definição de metodologias e políticas de uso desses repositórios se tornou imprescindível para o desenvolvimento da pesquisa no país. Este trabalho apresenta um estudo sobre iniciativas de acesso aberto a dados de pesquisa e soluções tecnológicas para repositórios de dados de pesquisa, mostra como essas soluções foram selecionadas com base em uma série de critérios de comparação. A seguir é descrito o processo de implantação da solução e um conjunto de testes realizados. Por fim, técnicas de avaliação de usabilidade são apresentadas e aplicadas sobre uma instância do sistema de repositório de dados em diferentes experimentos, um empírico e um analítico, resultando em um conjunto de recomendações de correções para os problemas encontrados.

Palavras-chave: Dados abertos, acesso aberto a dados de pesquisa, repositório de dados, usabilidade.

ABSTRACT

AZAMBUJA, Luis Alberto Barbosa. **Proposal for usability inspection of a Brazilian research data repository.** 2019. 103 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal do Rio Grande, Rio Grande.

In recent years, the evolution of Web systems has led to the emergence of initiatives to promote Open Access to Research Data around the world. Among the primary motivations for researchers to make data available from research projects on open access platforms stand out is to increase research group visibility, facilitate information exchange, validate results, and accountability transparency for research agencies - public or private development. The choice of a tool for this purpose as well as the definition of methodologies and policies for the use of these repositories became essential for the development of research in the country. This work presents a study of open data access initiatives and technological solutions for research data repositories, showing how these solutions were selected based on a series of comparison criteria. The following describes the solution deployment process and a set of tests performed. Finally, usability assessment techniques are presented and applied to an instance of the data repository system in different experiments, an empirical and other analytical, resulting in a set of recommendations for corrections to the problems founded.

Keywords: open data, open access to research data, data repository, usability.

LISTA DE FIGURAS

Figura 1	Ciclo de vida dos dados	18
Figura 2	Benefícios do Acesso Aberto	23
Figura 3	Pirâmide de Dados segundo a The Royal Society	28
Figura 4	Modelo de Referência OAIS	31
Figura 5	Relação entre número de avaliadores e problemas encontrados	41
Figura 6	Estrutura do Dspace	49
Figura 7	Estrutura do Dataverse	50
Figura 8	Tela com resultado da busca por “peixes de aquário”	62
Figura 9	Tela com ajuda de campo da busca avançada	62
Figura 10	Tela de resultados com arquivo indisponível	68
Figura 11	Tela de busca no conjunto de dados com ajuda não traduzida	69

LISTA DE TABELAS

Tabela 1	Funções de um sistema de gestão de dados de pesquisa	34
Tabela 2	Lista de heurísticas de Nielsen	51
Tabela 3	Níveis de severidade	51
Tabela 4	Comparativo dos principais requisitos desejados para o repositório nacional de dados de pesquisa.	54
Tabela 5	Indicadores de nível de atendimento dos requisitos	55
Tabela 6	Comparativo completo dos requisitos levantados pelo GT-RDP Brasil	56
Tabela 7	Lista de problemas e heurísticas afetadas - Catastrófico	60
Tabela 8	Lista de problemas e heurísticas afetadas - Grave	61
Tabela 9	Lista de problemas e heurísticas afetadas - Simples	63
Tabela 10	Lista de problemas e heurísticas afetadas - Cosmético	64
Tabela 11	Lista de problemas e heurísticas afetadas - Sem importância	65
Tabela 12	Notas das questões gerais sobre a satisfação dos especialistas	65
Tabela 13	Notas do questionário de satisfação dos usuários	69

LISTA DE ABREVIATURAS E SIGLAS

- AADP - Acesso Aberto a Dados de Pesquisa
- ABNT - Associação Brasileira de Normas Técnicas
- BDTD - Biblioteca Digital de Teses e Dissertações
- BSD - *Berkeley Software Distribution*
- C3 - Centro de Ciências Computacionais (FURG)
- CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- CCSDS - *Consultive Committee for Space Data Systems*
- CEDAP - Centro de Documentação e Acervo Digital de Pesquisa (UFRGS)
- CONARQ Conselho Nacional de Arquivos
- CPD - Centro de Processamento de Dados (UFRGS)
- DCC - Digital Curation Centre
- ETD - *Electronic Theses and Dissertations*
- FABICO - Faculdade de Biblioteconomia e Comunicação (UFRGS)
- FAIR - *Findable, Accessible, Interoperable and Re-usable*
- FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo
- Fiocruz - Fundação Oswaldo Cruz
- FURG - Universidade Federal do Rio Grande
- GT-RDP - Grupo de Trabalho em Rede de Dados de Pesquisa Brasileira
- HTTP - *Hypertext Transport Protocol*
- IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia
- ICHI - Instituto de Ciências Humanas e da Informação
- ISO - *International Organization For Standardization*
- MCTIC - Ministério da Ciência, Tecnologia, Inovações e Comunicações
- MD - Ministério da Defesa
- MEC - Ministério da Educação
- MinC - Ministério da Cultura

MS - Ministério da Saúde
NIST - *National Institute of Standards and Technology*
NSF - *National Science Foundation*
OAI - *Open Archives Initiative*
OAIS - *Open Archival Information System*
OECD - *Organisation for Economic Co-operation and Development*
PGD - Plano de Gestão de Dados
RDA - *Research Data Alliance*
RNP - Rede Nacional de Ensino e Pesquisa
TIC - Tecnologia da Informação e Comunicação
UC3 - *University of California Curation Center*
UFBA - Universidade Federal da Bahia
UFRN - Universidade Federal do Rio Grande do Norte
UnB - Universidade de Brasília
UFRGS - Universidade Federal do Rio Grande do Sul
VM - *Virtual Machine*
XML - *Extensible Markup Language*

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Motivação	13
1.2	Objetivos	13
1.3	Organização do texto	14
2	DADOS DE PESQUISA	16
2.1	Definição	16
2.2	Ciclo de vida dos dados	17
2.3	Acesso Aberto a Dados de Pesquisa	18
2.4	Benefícios dos dados abertos	22
2.5	Plano de Gestão de Dados	25
2.6	Repositórios de Dados de Pesquisa	26
2.6.1	Definição e conceitos envolvidos	26
2.6.2	Padronização e Confiabilidade	28
2.6.3	Sistemas de gestão de dados de pesquisa	32
2.6.4	Soluções tecnológicas	33
3	USABILIDADE	35
3.1	Definição	36
3.2	Métodos de avaliação	37
3.2.1	Métodos empíricos	38
3.2.2	Métodos analíticos	41
4	TRABALHOS RELACIONADOS	44
4.1	Comparativos entre <i>softwares</i> para repositório de dados	44
4.2	Avaliação da usabilidade de sistemas de repositórios	45
5	METODOLOGIA PROPOSTA	47
5.1	Ferramentas avaliadas	48
5.2	Avaliação da usabilidade	49
6	RESULTADOS	53
6.1	Análise das soluções tecnológicas e prototipação do repositório	53
6.1.1	Comparativo das principais características desejadas	53
6.1.2	Implantação do protótipo	57
6.2	Análise dos resultados da avaliação de usabilidade	59
6.2.1	Avaliação heurística	60
6.2.2	Avaliação empírica	64

6.2.3	Síntese das sugestões de melhoria	70
7	CONCLUSÃO	72
	REFERÊNCIAS	74
	APÊNDICE A DOCUMENTOS PARA AVALIAÇÃO HEURÍSTICA	82
A.1	Descrição da metodologia	83
A.2	Termo de consentimento e formulário	86
	APÊNDICE B DOCUMENTOS PARA AVALIAÇÃO EMPÍRICA	94
B.1	Descrição da metodologia	95
B.2	Roteiro do experimento	96
B.3	Ficha de acompanhamento	97
B.4	Termo de consentimento	100
B.5	Lista de tarefas	102
B.6	Formulário	103

1 INTRODUÇÃO

Cada vez mais a comunidade científica, os órgãos de fomento, as editoras e a população em geral estão atentas à produção científica nacional, principalmente a que é feita com recursos públicos, no tocante a divulgação dos dados da pesquisa e a sua consequente verificabilidade e reprodutibilidade.

O compartilhamento de dados de pesquisa traz benefícios a todos os atores envolvidos e está se tornando cada vez mais uma prática adotada pelos pesquisadores, seja por iniciativa própria ou não, pode ser incentivada e até mesmo uma etapa obrigatória na conclusão do processo de publicação de resultados.

Alguns editores e órgãos de fomento estão exigindo a disponibilização desses dados para acesso aberto e contínuo (SAYÃO; SALES, 2016), de modo que permitam sua verificabilidade. Esta prática está se tornando cada vez mais comum e a perspectiva é que se torne um item obrigatório para todos trabalhos publicados em revistas conceituadas e também como uma forma de comprovação de resultados aos financiadores da pesquisa.

Nos últimos anos houve um crescimento nas iniciativas de promoção do Acesso Aberto a Dados de Pesquisa (AADP) em todo o mundo. Dentre as principais motivações para que os pesquisadores disponibilizem os dados utilizados em seus projetos de pesquisa em plataformas de acesso aberto destacam-se aumentar a visibilidade dos grupos de pesquisa, facilitar o intercâmbio de informações, validar resultados obtidos e a transparência na prestação de contas para as agências de fomento públicas ou privadas.

Implantar um serviço de submissão, armazenamento e recuperação de dados de pesquisa envolve a correta escolha dessa ferramenta e o incentivo ao seu uso. Não basta convencer a comunidade produtora sobre os benefícios desse compartilhamento, é necessário oferecer um sistema eficiente e agradável de utilizar. Segundo SANTOS; MOURA; PAVÃO (2014), avaliar a usabilidade da interface de um repositório é tão importante quanto manter a qualidade dos metadados e dos objetos digitais depositados.

1.1 Motivação

Este trabalho pretende avaliar a usabilidade de um repositório de dados de pesquisa que surge da necessidade de atender uma demanda específica do Grupo de Trabalho em Rede de Dados de Pesquisa Brasileira (GT-RDP Brasil). Ele é uma iniciativa da Rede Nacional de Ensino e Pesquisa (RNP)¹ e do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)², que em março de 2017, após o lançamento de uma chamada pública sobre AADP³, selecionaram a proposta intitulada Rede de Dados de Pesquisa Brasileira⁴.

Este Grupo de Trabalho é formado por uma parceria entre a Universidade Federal do Rio Grande do Sul (UFRGS) e a Universidade Federal do Rio Grande (FURG), sob a coordenação do Prof. Dr. Rafael Port da Rocha, da UFRGS. Ele envolve o Centro de Documentação e Acervo Digital de Pesquisa (CEDAP), órgão auxiliar da Faculdade de Biblioteconomia e Comunicação (FABICO) da UFRGS, o Centro de Processamento de Dados (CPD) da UFRGS, e o Grupo de Pesquisa em Gerenciamento de Informações do Centro de Ciências Computacionais (C3) da FURG.

A atividade fundamental do GT-RDP Brasil é identificar as principais práticas de AADP no país e implementar um protótipo de serviço computacional para a comunidade científica.

1.2 Objetivos

O objetivo geral desta dissertação de mestrado é avaliar a usabilidade de uma solução tecnológica para um serviço de repositório de dados de pesquisa com acesso aberto para a comunidade científica brasileira. Para alcançar esse objetivo foi efetuado o estudo e a aplicação de metodologias para avaliação de usabilidade, tanto baseadas em análise de especialistas quanto avaliação por usuários, os quais julgam subjetivamente as interfaces de interação, medindo sua satisfação na utilização dos repositórios.

Os principais objetivos específicos desta pesquisa são:

- Estudar a literatura científica a respeito do ciclo de vida dos dados de pesquisa, destacando como os repositórios contribuem na solução de compartilhamento em acesso aberto e na gestão desses dados.

¹Entidade vinculada ao Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) e mantida em conjunto com os ministérios da Educação (MEC), Cultura (MinC), Saúde (MS) e Defesa (MD), que provê a integração global e a colaboração apoiada em tecnologias de informação e comunicação para a geração do conhecimento e a excelência da educação e da pesquisa. Mais informações estão disponíveis em <<https://www.rnp.br/institucional/quem-somos>>.

²Órgão vinculado ao MCTIC que tem como uma das atribuições a promoção da popularização da informação científica e tecnológica. Hoje, o IBICT é referência em projetos voltados ao movimento do acesso livre ao conhecimento. Mais informações estão disponíveis em <<https://www.ibict.br>>.

³<https://www.rnp.br/noticias/rnp-e-ibict-lancam-chamada-promover-acesso-aberto-dados-pesquisa>

⁴<https://www.rnp.br/noticias/rnp-divulga-proposta-selecionada-pela-chamada-dados-abertos-pesquisa>

- Identificar e analisar soluções tecnológicas para repositórios de dados mais usados no mundo. Esta análise inclui a comparação das soluções de acordo com uma série de requisitos funcionais e não funcionais levantados junto ao GT-RDP para atender o ecossistema da pesquisa nacional.
- Disponibilizar tanto um protótipo quanto um piloto de repositório de dados de pesquisa com acesso aberto. O primeiro tem como público alvo usuários do setor de infraestrutura da RNP, um grupo *beta-tester* da comunidade científica brasileira de Redes de Computadores e o próprio GT-RDP. Já o piloto inclui dados reais de diferentes instituições de ensino e pesquisa.
- Avaliar a usabilidade de uma instância do repositório através de experimentos utilizando diferentes metodologias, visando identificar melhorias e ajustes necessários para implantação da solução tecnológica para todas as instituições do país.

1.3 Organização do texto

O restante do texto está organizado da seguinte forma.

O Capítulo 2 relaciona uma série de conceitos essenciais para o entendimento do trabalho proposto. Primeiramente são definidos dados de pesquisa e ilustradas as etapas do ciclo de vida dos dados. Após, são diferenciados Acesso Aberto e Acesso Aberto a Dados de Pesquisa, onde os principais benefícios desta abertura são discutidos. Também é enfatizada a importância de se estabelecer um Plano de Gestão de Dados em organizações de ensino e pesquisa. Ainda são definidos e exemplificados repositórios de dados de pesquisa, destacando padrões e as principais tecnologias envolvidas na sua construção e implantação.

Conceitos e metodologias de avaliação de usabilidade são abordados no Capítulo 3.

Logo após, no Capítulo 4 são apresentados uma ligeira revisão bibliográfica e um conjunto de trabalhos relacionados. Estes estudos comparam diferentes *softwares* para repositórios de dados, a partir de múltiplos critérios, ou utilizam as principais metodologias de avaliação de usabilidade estudadas sobre sistemas de repositório.

A metodologia proposta, tanto para comparar e escolher uma solução tecnológica para repositório de dados de pesquisa, quanto para avaliar a interface deste *software* são descritas no Capítulo 5.

O Capítulo 6 apresenta os resultados da pesquisa relatada nesta dissertação. Primeiramente é descrita a avaliação dos requisitos levantados comparando as diferentes soluções tecnológicas estudadas. Após, é relatada a implantação do protótipo do Repositório Brasileiro de Dados de Pesquisa, destacando a configuração das tecnologias. Ainda são analisados os resultados de avaliações de usabilidade heurísticas e empíricas, realizadas por usuários e especialistas.

Por fim, no Capítulo 7 são apresentadas as conclusões, principais contribuições, publicações relacionadas e direções futuras.

2 DADOS DE PESQUISA

Este capítulo identifica, define e classifica dados de pesquisa. Os benefícios da abertura dos dados de pesquisa são evidenciados para cada ator envolvido nas distintas fases de seu ciclo de vida. Por fim, conceitos, padrões, tecnologias e ferramentas que envolvam repositórios de dados de pesquisa são abordados.

2.1 Definição

Por dados de pesquisa, entende-se os dados provenientes da investigação científica. Os mesmos não devem ser confundidos com a produção acadêmico-científica tradicional (resumos, artigos, relatórios, capítulos de livros, livros, etc.), nem com dados governamentais, sejam eles administrativos, censitários ou de outra natureza. Dados são componentes centrais do processo de pesquisa. Eles são registros científicos que embasam os resultados de pesquisa publicados na forma de dissertações, teses, artigos, patentes e trabalhos científicos diversos.

Dados de pesquisa referem-se a informações, em particular fatos ou números, coletadas para serem examinadas e consideradas como base para fundamentação, discussão ou cálculo (ECReI, European Commission Research & Innovation, 2017). Em um contexto de pesquisa, exemplos de dados incluem estatísticas, resultados de experimentos, medições, observações resultantes de trabalho de campo, resultados de pesquisas, registros de entrevistas e imagens. O foco está em dados de pesquisa que estão disponíveis em formato digital. Portanto, os usuários deveriam poder acessar, minerar, explorar, reproduzir e divulgar dados de pesquisa gratuitamente.

A Organisation for Economic Co-operation & Development no seu *OECD Principles and Guidelines for Access to Research Data from Public Funding* (OECD, Organisation for Economic Co-operation & Development, 2007), descreve dados de pesquisa como registros factuais (apontamentos numéricos, registros textuais, imagens e sons) usados como fontes primárias para a pesquisa científica e que são comumente aceitos na comunidade científica como necessários para validar os resultados da pesquisa.

Todo pesquisador responsável deve gerenciar os dados de sua pesquisa. Isso significa

pensar em que dados coletar, como descrevê-los, o que armazenar, por quanto tempo, e como torná-los acessíveis a longo prazo. É recomendado elaborar um Plano de Gestão de Dados de Pesquisa e escolher um repositório para manter estes dados.

A National Science Foundation (NSF)¹, agência federal independente dos EUA, classifica os dados de acordo com sua origem (NSB, National Science Board, 2005):

- Dados observacionais - obtidos através de observações diretas e associados a uma data e local específicos. Como não podem ser recoletados, precisam ser arquivados indefinidamente.
- Dados computacionais - obtidos como resultado da execução de um modelo ou simulação computacional. Se informações suficientes sobre o modelo estiverem disponíveis, não será necessário preservar os dados pois eles poderão ser reproduzidos desde que o modelo e um conjunto robusto de metadados sejam arquivados.
- Dados experimentais - podem não ser passíveis de reprodução precisa dadas algumas condições e variáveis não serem conhecidas ou facilmente reproduzidas, ou ainda os custos de reprodução dos experimentos são proibitivos. Nesse caso, a preservação a longo prazo dos dados deve ser garantida.

Não importando a origem, o formato ou que sejam dados brutos ou processados, o que importa é que precisam ser preservados adequadamente, com metadados ricos e informações suficientes sobre sua origem.

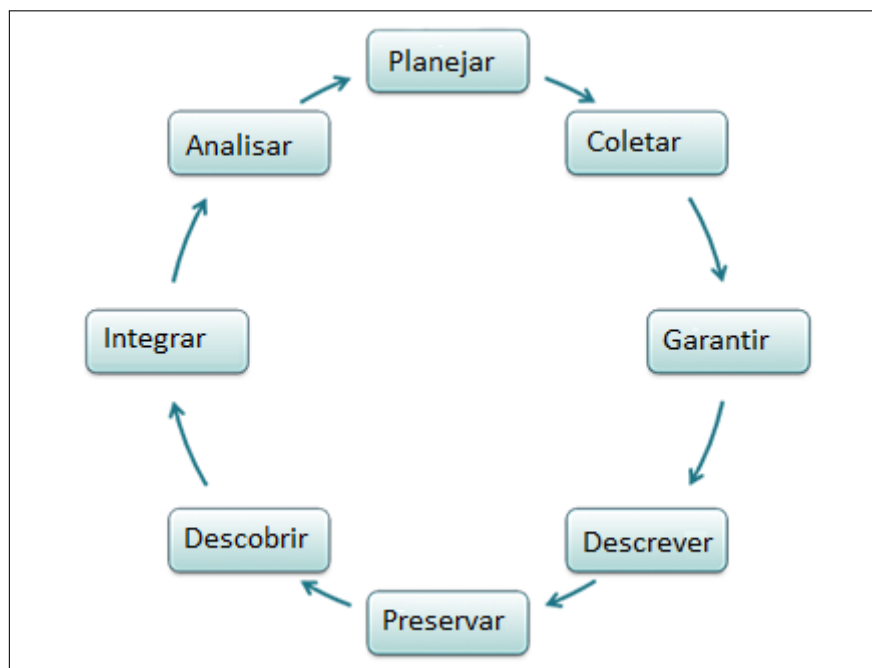
2.2 Ciclo de vida dos dados

Segundo DataONE, Data Observation Network for Earth (2012), do ponto de vista do pesquisador, a gestão do ciclo de vida dos dados de pesquisa envolve oito componentes, ilustrados na Figura 1:

- Planejar - descrição dos dados que serão compilados e como eles serão gerenciados e disponibilizados ao longo de sua vida útil;
- Coletar - as observações são realizadas manualmente, por sensores ou outros instrumentos e os dados são colocados em uma forma digital;
- Garantir - aprimorar a qualidade dos dados através de controles e inspeções;
- Descrever - os dados são descritos com precisão e detalhamento usando padrões de metadados apropriados;
- Preservar - os dados são enviados para repositórios adequados para preservação a longo prazo;

¹<http://www.nsf.gov>

Figura 1: Ciclo de vida dos dados



Fonte: Adaptado de (DataONE, Data Observation Network for Earth, 2012)

- Descobrir - possibilitar dados potencialmente úteis sejam descobertos e obtidos, juntamente com informações relevantes sobre os dados (metadados);
- Integrar - dados de fontes diferentes são combinados para formar um conjunto de dados homogêneo que pode ser prontamente analisado;
- Analisar - os dados são analisados para reiniciar o ciclo.

A fase *Descobrir* é especialmente importante para esta proposta de dissertação porque o Repositório Brasileiro proposto disponibilizará os dados de pesquisa em acesso aberto, possibilitando que qualquer pessoa ou *software* possa recuperar os conjuntos de dados facilmente. As fases anteriores, *Descrever* e *Preservar*, são fundamentais nesse processo. Os conceitos acerca do acesso aberto são apresentados a seguir.

2.3 Acesso Aberto a Dados de Pesquisa

Segundo ECRéI, European Commission Research & Innovation (2017), o acesso aberto refere-se à prática de fornecer acesso *online* a informações científicas que são gratuitas para o usuário final e reutilizáveis. No contexto da pesquisa e da inovação, informação científica pode significar:

- artigos de pesquisa científica revisados por pares, publicados em anais de eventos científicos ou revistas científicas;

- dados de pesquisa subjacentes a publicações, dados curados ou ainda dados brutos.

Segundo COSTA (2006), “... acesso aberto é tema indiscutivelmente relevante, envolvente e abrangente. Relevante, porque põe em cheque paradigmas da ciência e, como tal, pode contribuir para o surgimento de novos paradigmas”. Este autor cita ainda o movimento Filosofia Aberta, que abrange, entre outras questões:

- *software* aberto (ou livre), para o desenvolvimento de aplicações em computador;
- arquivos abertos, para interoperabilidade em nível global;
- acesso aberto – questão mais polêmica – para a disseminação ampla e irrestrita de resultados da pesquisa científica.

De acordo com HARNAD et al. (2008) existem duas estratégias básicas para publicações com acesso aberto:

- Via Dourada (*Golden Road*), que consiste na publicação de um artigo em um periódico com licença de acesso aberto;
- Via Verde (*Green Road*), que compreende o auto-arquivamento do artigo científico, já publicado ou aceito para publicação em um periódico referendado, que tenha recebido permissão (sinal verde) dos editores, em um servidor de arquivos abertos (repositório), com acesso aberto.

Já o Acesso Aberto a Dados de Pesquisa (AADP) remete à disponibilização de dados de pesquisa em plataformas digitais de livre acesso, respeitando os direitos autorais, períodos de embargo, questões de confidencialidade, privacidade, especificidades das áreas científicas, entre outros aspectos.

Normalmente os pesquisadores focam na produção e publicação dos resultados e das análises científicas, sendo que apenas os dados apresentados nestas publicações ficam disponíveis e acessíveis. Porém esses dados estão disponíveis apenas parcialmente pois para produzir a pesquisa, muitos outros dados foram gerados, coletados, tabulados, transcritos e registrados. Ocorre que eles acabam ficando aos cuidados do pesquisador que os coletou e, muitas vezes, são perdidos pela fragilidade das mídias e pela obsolescência tecnológica. Os dados, que quase nunca são analisados em sua totalidade pela pesquisa que os ensejou, poderiam gerar novas análises científicas e a proposição de novos projetos de estudo (ROCHA et al., 2017).

Desde o final da década de 1990 que a *Web* e a evolução na área de Tecnologia da Informação e Comunicação (TIC) possibilitaram um maior alcance da produção científica, mas com isso muitas editoras técnico-científicas restringiram e cobraram o acesso às suas publicações, principalmente na forma de assinaturas RODRIGUES; OLIVEIRA

(2012). Essa decisão fez com que grande parte desse conhecimento, em sua grande maioria produzido com recurso público, ficasse restrito a quem pudesse pagar ou ter acesso através de alguma instituição pagante, tal como a CAPES no Brasil.

A disseminação desse conhecimento se torna então imprescindível e várias iniciativas nesse sentido começam surgir. Entre essas iniciativas podemos citar: *Budapest Open Access Initiative* (2002)², *Bethesda Statement on Open Access Publishing* (2003)³ e *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* (2003)⁴, que expandiu o conceito de acesso aberto a dados de pesquisa (SANTOS, 2017). Todas elas promovem a disseminação da filosofia do acesso aberto e surgem como consequência da *Open Archives Initiative - (OAI)*⁵, - instituição criada em 1999 que desenvolve e promove padrões de interoperabilidade que visam facilitar a disseminação eficiente de conteúdo.

COSTA (2006) cita um relatório da *Association of Learner and Professional Society Publishers* de 2005 (Kaufman-Wills Group, 2005):

“As discussões sobre acesso aberto (*open access*) tendem a ser fortes em retórica, mas pequenas em fatos.”

Felizmente desde essa época muita coisa evoluiu e ações concretas de várias nações e organizações em todo o mundo estão tornando o AADP uma realidade.

No Brasil é importante citar a iniciativa do IBICT (2005) em seu “Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica”, inclusive quando declara que todo esse conhecimento científico produzido é um patrimônio cultural. Em 2016, o IBICT lança o “Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã” (IBICT, 2016) reafirmando a importância do acesso aberto e extendendo-a aos dados de pesquisa.

Também no Brasil, um exemplo a ser seguido é o da Scielo, um programa da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) que desde 1998, anos antes da iniciativa Budapeste, promove o acesso aberto (PACKER et al., 2014).

Mais recentemente vários empreendimentos de acesso aberto surgiram e incluem o AADP. Entre eles se destaca o programa “*Horizon 2020*”⁶, lançado em 2011, que tem como base o acesso aberto a publicações científicas e seus dados de pesquisa. A Comissão Europeia pretende fazer com que todos os resultados de pesquisa, financiados pelos seus bilhões de euros⁷ do programa de pesquisa “*Horizon 2020*” (2014-2020), sejam acessíveis a todos (ECReI, European Commission Research & Innovation, 2014).

Segundo a Fundação Oswaldo Cruz (Fiocruz), o programa visa⁸:

²<http://www.budapestopenaccessinitiative.org/read>

³<http://legacy.earlham.edu/~peters/fos/bethesda.htm>

⁴<https://openaccess.mpg.de/Berlin-Declaration>

⁵<https://www.openarchives.org/organization/>

⁶<https://ec.europa.eu/programmes/horizon2020/en/history-horizon-2020>

⁷http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/281113_Horizon_2020_standard_presentation.pdf

⁸<http://www.ensp.fiocruz.br/portal-ensp/informe/site/materia/detalhe/30777>

- definir o acesso aberto a publicações sujeitas à apreciação prévia de avaliadores, como princípio geral no Horizonte 2020, através da edição aberta (acesso aberto “dourado”) ou do autoarquivamento (acesso aberto “verde”);
- promover o acesso aberto a dados de investigação (resultados e observações experimentais, informação gerada por computador, etc.) e definir um quadro-piloto no Horizonte 2020, tendo em conta preocupações legítimas em matéria de privacidade, interesses comerciais e questões associadas a grandes volumes de dados;
- desenvolver e apoiar infraestruturas eletrônicas para alojar e partilhar informação científica (publicações e dados) que sejam interoperáveis a nível europeu e mundial;
- ajudar os investigadores a cumprirem as obrigações relativas ao acesso aberto e a promoverem uma cultura de partilha.

Um importante marco em relação ao AADP é que, desde maio de 2013, o Grupo Editorial *Nature*⁹, visando uma maior transparência e uma melhor reprodutibilidade dos resultados publicados, passou a exigir que para todos artigos submetidos às suas revistas fossem disponibilizados os dados em repositórios de dados de pesquisa, assim como os métodos e outros detalhes que frequentemente são mal ou parcialmente descritos. Desde então, outras editoras, tais como Elsevier e Wiley, passaram a fazer a mesma exigência (DUDZIAK, 2016; MELERO; HERNÁNDEZ-SAN-MIGUEL, 2014).

Dentre as iniciativas globais, destaca-se a *Research Data Alliance* (RDA)¹⁰. A RDA foi criada em 2013 com financiamento da Comissão Europeia¹¹, da *National Science Foundation* (NSF)¹², do *National Institute of Standards and Technology* (NIST)¹³ e do *Australian Government’s Department of Innovation*, com o objetivo de construir uma infraestrutura social e técnica para permitir o compartilhamento aberto de dados científicos nas mais diversas áreas do conhecimento.

Segundo a própria RDA, fica evidente o seu comprometimento com a ciência aberta, respeitando ao mesmo tempo, a pluraridade e a interdisciplinaridade tão comuns e necessárias ao meio científico:

“Com mais de 6.000 membros de 130 países (setembro de 2017), a RDA fornece um espaço neutro onde seus membros podem unir-se através de grupos de trabalho e interesse globais para desenvolver e adotar infraestruturas que promovam compartilhamento de dados e pesquisa orientada a dados, além de acelerar o crescimento de uma comunidade de dados

⁹<http://www.nature.com/authors/policies/availability.html>

¹⁰<https://www.rd-alliance.org>

¹¹https://ec.europa.eu/commission/index_pt

¹²U.S. independent federal agency - <https://www.nsf.gov>

¹³U.S. Department of Commerce - <https://www.nist.gov>

coesa que integra membros através das fronteiras nacionais, de domínio, da pesquisa, geográficas e geracionais.”

LEITE (2009) cita a *Budapest Open Access Initiative*¹⁴ (2001):

“A única restrição à reprodução e distribuição e a única função do copyright neste contexto devem ser o controle dos autores sobre a integridade de sua obra e o direito de serem adequadamente reconhecidos e citados.”

2.4 Benefícios dos dados abertos

Esta seção apresenta o conjunto de atores envolvidos em todas as fases da pesquisa científica, inclusive os que não tem participação direta mas que de uma maneira ou de outra usufruem dos benefícios da ciência aberta, mais especificamente dos benefícios do acesso aberto a dados de pesquisa.

Segundo COSTA (2006), os principais atores envolvidos no processo de comunicação científica são universidades e seus pesquisadores, editores científicos e agências de fomento. Outros atores citados como governos, bibliotecas e museus não foram detalhados pois não foram considerados alvos primários do trabalho do GT-RDP Brasil. Já FINCH (2013), considera os “interesses e aspirações dos diversos participantes do sistema de publicação acadêmica: pesquisadores, universidades, financiadores, sociedades científicas e editores”.

Para este trabalho foi considerado o seguinte conjunto de atores: pesquisadores, instituições de pesquisa, editores, órgãos de fomento, comunidade científica representada pelos pesquisadores que se beneficiam do acesso aberto e, talvez o mais importante beneficiário, a sociedade.

Em 22 de maio de 2012, na *University of North Texas*, um grupo multidisciplinar se reuniu para elaborar a “Declaração de Denton: Um Manifesto de Dados Abertos” (DENTON, 2012). Nela estão dispostos importantes princípios que devem ser observados na adoção de uma política de dados abertos e como consequência natural benefícios a toda sociedade.

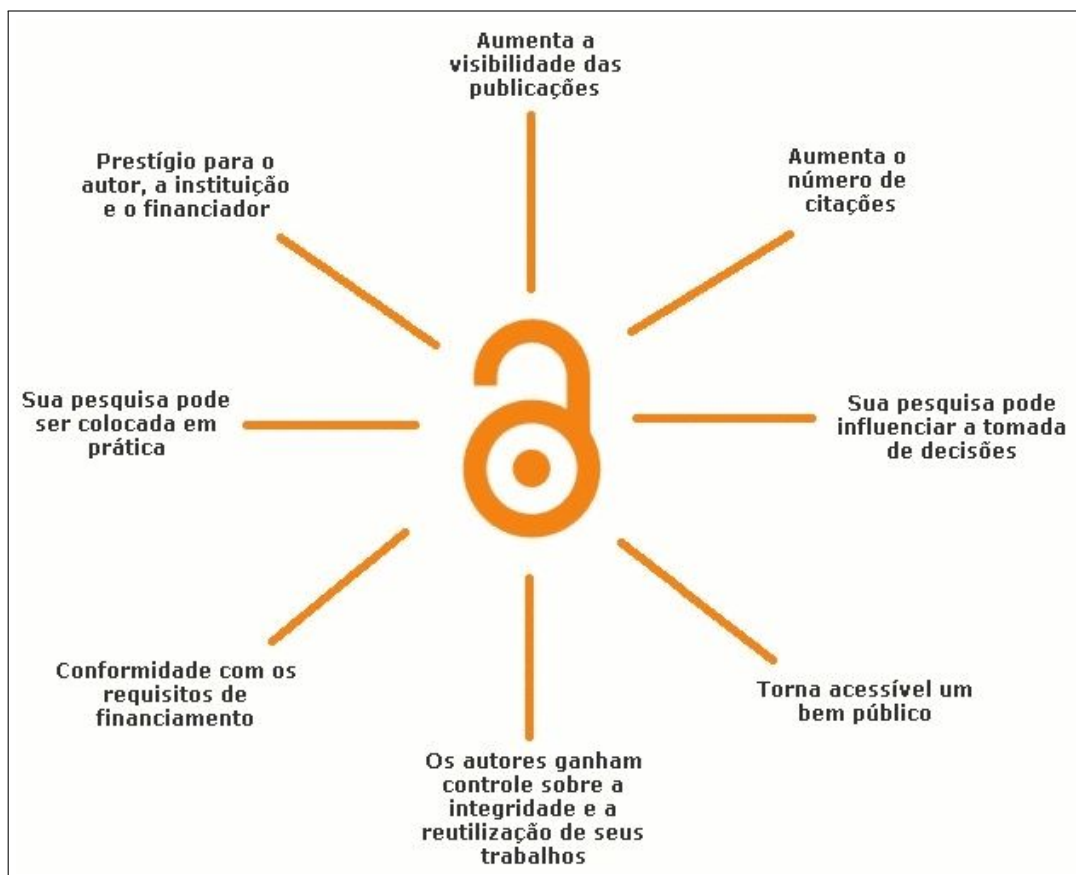
Assim como o acesso aberto às publicações traz vantagens a todos os envolvidos, o acesso aberto a dados de pesquisa também pode trazer diversos benefícios. A Figura 2 reúne as principais vantagens e benefícios do acesso aberto a publicações científicas, os quais podem ser usados para traçar um paralelo com AADP.

Segundo LAWRENCE (2001), o acesso aberto aumenta a visibilidade das publicações:

“Se assumimos que os artigos publicados em um mesmo local são de qualidade similar, a análise por local sugere que os artigos on-line são mais citados devido à sua disponibilidade mais fácil.”

¹⁴<http://www.budapestopenaccessinitiative.org/>

Figura 2: Benefícios do Acesso Aberto



Fonte: Adaptado da ilustração “Benefícios do Acesso Aberto”, CC Danny Kingsley e Sarah Brown

Isso também é válido com relação à publicação dos dados de pesquisa (MELERO; HERNÁNDEZ-SAN-MIGUEL, 2014), pois seus resultados podem ser validados e ganham credibilidade. Este princípio também vale no momento inverso. Baseando-se em duas publicações com algum resultado conflitante entre elas, mas uma delas apresenta os dados de pesquisa e pode ser comprovada, qual você citaria em seus artigos?

Seguindo a lógica anterior, fica claro que com esse aumento de credibilidade consequentemente haverá um aumento no número de citações, o que causará uma maior visibilidade (HARNAD et al., 2008). Esse aumento na visibilidade, na credibilidade e no número de citações traz benefício direto a todos envolvidos: cresce o prestígio do autor, da instituição, do financiador e da editora. Consequentemente todos serão mais visados pelos pesquisadores em seus novos trabalhos. Essa visibilidade também diminui as possibilidades de plágios (LEITE, 2009), pois a publicação em repositórios confiáveis registra no tempo a autoria de artigos e dados.

O depósito em repositórios confiáveis oferece um ambiente seguro para que os tra-

balhos sejam permanentemente armazenados (LEITE, 2009) já que um dos grandes problemas enfrentados pelos pesquisadores é quanto à preservação de seus dados. É comum esquecer deles em mídias pouco confiáveis e mal armazenadas, sujeitas a intempéries e ao esquecimento. A preservação consciente em bases de dados confiáveis e públicas ajuda muito a garantir o controle sobre a integridade e reutilização de seus dados a longo prazo (SAYÃO, 2010).

Assim, o AADP é fundamental para o avanço da ciência (DENTON, 2012), permite a continuidade da pesquisa, que ela seja colocada em prática e pode auxiliar na tomada de decisões de políticas públicas.

Todo investimento com recursos públicos deve ter sua transparência garantida e o AADP torna acessível um bem público e todo e qualquer resultado posterior advindo do uso destes dados deverá ser benéfico para a sociedade que a possibilitou. Vários órgãos de fomento, como já ocorre na Comunidade Europeia (ECReI, European Commission Research & Innovation, 2017), exigem o depósito dos dados de pesquisa em repositórios públicos em conformidade com os requisitos do financiamento. Além disso, a reutilização desses dados “contribui de forma significativa para que a ciência avance e maximize os investimentos aplicados em programas de pesquisa” (SAYÃO; SALES, 2015).

Os pesquisadores devem ter a garantia de prazos para se beneficiarem de seus próprios trabalhos, mas que também permita a verificação dos resultados (DENTON, 2012), respeitando os direitos autorais, períodos de embargo, questões de confidencialidade, privacidade, especificidades das áreas científicas, entre outros aspectos. Para atender estes requisitos, os dados podem ser armazenados no repositório de forma privada com um período de embargo, que respeite o tempo necessário para publicação dos resultados ou para a publicação das patentes submetidas. Esse tempo pode ser configurado pelo pesquisador ou exigido pelo edital que financia o projeto.

As instituições de pesquisa conseguem acesso centralizado à produção de seus pesquisadores com relatórios das atividades, citações e contribuições (LEITE, 2009). Isso ajuda no planejamento estratégico, no compartilhamento e reuso das informações produzidas. Como consequência, o AADP proporciona aumento na visibilidade, reputação e prestígio da instituição.

Como citado anteriormente, as editoras estão atentando para o depósito dos dados de pesquisa, tanto para uma melhor conferência do resultado da publicação quanto para uma maior transparência. Essa transparência gera uma maior confiabilidade na publicação e consequentemente na editora o que gera uma maior procura pela comunidade acadêmica, tanto para novos trabalhos quanto para busca por referências.

Já para os órgãos de fomento, o principal benefício é que novos investimentos podem ser planejados levando em conta autores, áreas temáticas e instituições que comprovam seus trabalhos quando abrem suas pesquisas e dados.

A maior contribuição seja a facilitação resultante, a nível global, na busca e

colaboração de pesquisas. Os dados produzidos por um podem ser reutilizados por vários, o que diminui custos (financeiros, tempo, etc.) ou, no mínimo, permite redirecioná-los. O que não pode deixar de ser citado é o benefício à sociedade, já que muitas vezes é ela que provê os recursos necessários à pesquisa. Nada mais justo que os dados coletados e/ou gerados sejam disponibilizados à toda comunidade.

Em suma, o AADP torna possível atestar veracidade e utilidade de resultados publicados, reduzindo custos e redirecionando recursos para novas pesquisas.

2.5 Plano de Gestão de Dados

Um Plano de Gestão de Dados (PGD) é um documento curto, geralmente com no máximo duas páginas, que descreve que dados serão coletados ou gerados, quais metodologias e padrões serão utilizados, como serão compartilhados, curados e preservados. Portanto, um PGD é imprescindível no contexto de um projeto de AADP, já que trata diretamente da preservação, compartilhamento e reuso dos dados e está sendo cada vez mais exigido pelas agências de fomento como item obrigatório para garantir o financiamento (SAYÃO; SALES, 2015).

Segundo a FAPESP¹⁵, um PGD deve responder a duas questões básicas:

1. Quais dados serão gerados pelo projeto?
2. Como serão preservados e disponibilizados, considerando questões éticas, legais, de confidencialidade e outras?

Ele deve conter:

- descrição dos dados e metadados produzidos pelo projeto;
- restrições legais ou éticas para compartilhamento, políticas de privacidade, confidencialidade, segurança, propriedade intelectual e outros;
- política de preservação e compartilhamento, contendo períodos de carência e até quando serão preservados e disponibilizados;
- descrição de mecanismos, formatos e padrões para armazená-los e disponibilizá-los.

Atualmente estão disponíveis algumas ferramentas para auxiliar a elaboração de um PGD como, por exemplo, a DPMonline¹⁶. Desenvolvida pela *Digital Curation Centre* (DCC) e pela *University of California Curation Center* (UC3), é uma ferramenta *online* de código aberto que possui inclusive versões adaptadas e traduzidas em outras línguas.

¹⁵<http://www.fapesp.br/gestaodedados>

¹⁶<https://dmponline.dcc.ac.uk>

Esta ferramenta possui diversos modelos de PGD baseados nos requisitos específicos listados nos documentos de políticas dos financiadores, inclusive o exigido pela Comissão Europeia no programa “*Horizon 2020*”. Apesar disso, DPMonline lembra que os pesquisadores devem sempre consultar as diretrizes do financiador diretamente para obter informações autoritativas¹⁷.

O PGD tem importância fundamental no ciclo de vida dos dados e além de ser uma exigência de financiadores e editoras, é um recurso inestimável para a continuidade de projetos de pesquisa.

2.6 Repositórios de Dados de Pesquisa

Para que os dados de pesquisa possam ser reaproveitados eles precisam ser preservados e passíveis de localização e recuperação facilitadas. Segundo HEDSTROM (1997), a preservação digital é um processo que envolve o “planejamento, alocação de recursos e aplicação de métodos e tecnologias de preservação necessários para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável”. Portanto, ela é um conjunto de atividades necessárias para assegurar que os objetos digitais possam ser localizados, reproduzidos, utilizados e compreendidos no futuro.

2.6.1 Definição e conceitos envolvidos

Os repositórios digitais de dados de pesquisa surgem como ferramentas de apoio à execução dos processos necessários a gestão desses dados (SAYÃO; SALES, 2016). Os repositórios possibilitam a armazenagem em sistemas seguros com a devida preservação a longo prazo e recuperação dos dados de pesquisa.

Segundo LEITE (2009), repositórios digitais são “vários tipos de aplicações de provedores de dados que são destinados ao gerenciamento de informação científica, constituindo-se, necessariamente, em vias alternativas de comunicação científica” e podem ser classificados em três tipos:

- Repositórios institucionais - voltados à produção intelectual de uma instituição, especialmente universidades e institutos de pesquisa. Exemplo: e-Prints Soton — repositório de Pesquisa da Universidade de Southampton¹⁸, LUME da UFRGS¹⁹ e RI FURG²⁰;
- Repositórios temáticos ou disciplinares - voltados a comunidades específicas. Tratam, portanto, da produção intelectual de áreas do conhecimento em particular.

¹⁷https://dmponline.dcc.ac.uk/public_templates

¹⁸<http://eprints.soton.ac.uk>

¹⁹<https://lume.ufrgs.br>

²⁰<http://repositorio.furg.br>

Exemplo: *E-LIS — EPrints in Library and Information Science*²¹ e arXiv.org²²;

- Repositórios de teses e dissertações (*Electronic Theses and Dissertation — ETDs*) - repositórios que lidam exclusivamente com teses e dissertações. Muitas vezes a coleta das muitas ETDs é centralizada por um agregador. Exemplo: Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade de Brasília²³, do IBICT²⁴ e da FURG²⁵.

Considerando repositórios de dados de pesquisa, SAYÃO; SALES (2016) afirmam que há um certo consenso entre os autores classificando-os em quatro tipos:

- Repositórios institucionais de dados de pesquisa - são gerenciados e funcionam no âmbito de uma instituição acadêmica ou instituto de pesquisa e geralmente armazenam somente a própria produção científica. Pelas suas características são também na maioria das vezes multidisciplinares. Exemplo: Edinburgh Data Share²⁶;
- Repositórios disciplinares, ou temáticos, de dados de pesquisa - são dedicados ao armazenamento de domínios específicos de pesquisa e normalmente mantidos por grupos ou entidades voltadas a um campo restrito do conhecimento científico. Exemplo: *BioModels Database*²⁷, GenBank²⁸ e PANGAEA²⁹;
- Repositórios multidisciplinares de dados de pesquisa - aceitam dados provenientes de diversas áreas do conhecimento e de diversas fontes, ou seja, de pesquisadores de diversas instituições. Exemplo: Dryad³⁰ e Figshare³¹;
- Repositórios de dados de pesquisa orientados por projetos, ou de um projeto específico - são voltados para o arquivamento da produção de um grupo e/ou projeto específico. Exemplo: *The Scientific Drilling Database*³².

Como os autores sugerem, essa classificação é útil para distinguir os tipos de repositórios de dados em relação a quem pode submeter e o tipo de material armazenado, mas não consegue espelhar os níveis de complexidade das infraestruturas em relação a permanência, acesso e alcance desses dados.

²¹<http://eprints.rclis.org>

²²<http://arxiv.org>

²³<http://bdtd.bce.unb.br/tesesimplificado>

²⁴<http://bdtd.ibict.br>

²⁵<https://argo.furg.br/?BDTD>

²⁶<https://datashare.is.ed.ac.uk>

²⁷<https://www.ebi.ac.uk/biomodels>

²⁸<https://www.ncbi.nlm.nih.gov/genbank>

²⁹<https://www.pangaea.de>

³⁰<http://datadryad.org>

³¹<https://figshare.com>

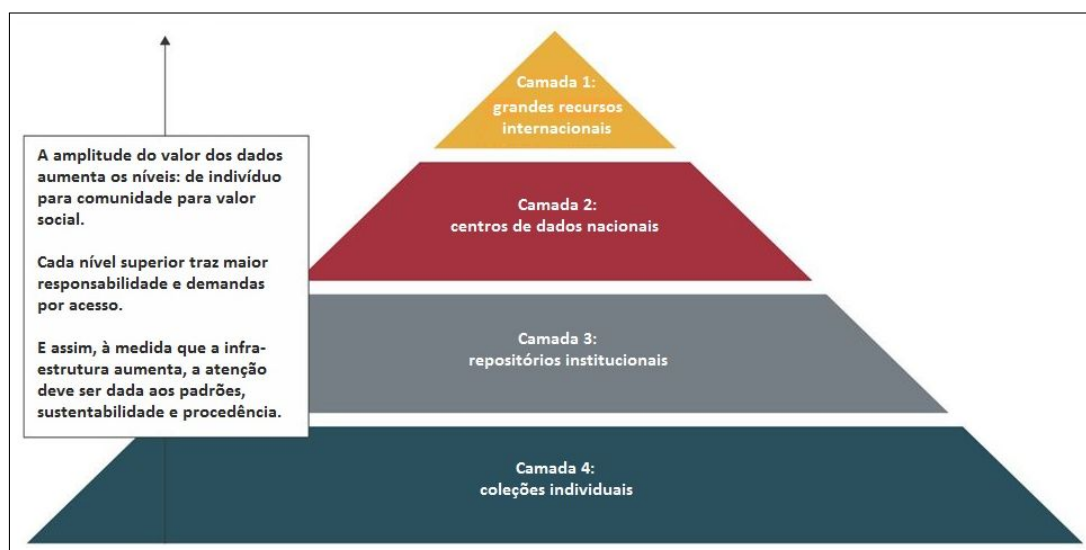
³²<http://gfzpublic.gfz-potsdam.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:236508>

Já a THE ROYAL SOCIETY (2012) classifica os repositórios através de uma pirâmide hierárquica de acordo com o alcance que esses dados têm e sua correspondente importância percebida e o conseqüente custo para mantê-los em quatro camadas, citadas a partir do topo: internacional, nacional, institucional e comunitário ou individual. A Figura 3 ilustra essa pirâmide.

O relatório da THE ROYAL SOCIETY (2012) também explica que compartilhar dados de pesquisa pode ser complexo e caro, portanto é preciso manter estimativas realistas da demanda por esses dados. Outro dado citado é que quanto mais próximo da base da pirâmide mais dificuldades e menos recursos têm os gestores de dados. Além disso se a importância dos dados disponibilizados cresce ou é absorvida por outras bases de dados ocorre uma migração natural para a camada superior.

Devido a abrangência do projeto e a diversidade da comunidade científica nacional, que precisa armazenar uma grande variedade de tipos e formatos de dados que caracterizam suas áreas de pesquisa, se faz necessário encontrar uma solução que atenda esse tipo de repositório, conhecido como “cauda longa”. Esse tipo de solução foi pesquisada entre as disponíveis no mercado (*software* de prateleira) e são apresentadas na Seção 2.6.4.

Figura 3: Pirâmide de Dados segundo a The Royal Society



Fonte: Science as an open enterprise – THE ROYAL SOCIETY (2012, pag 60)

2.6.2 Padronização e Confiabilidade

Na busca por critérios para comparar as soluções disponíveis, foram encontrados alguns modelos e padrões amplamente utilizados para avaliar e referenciar repositórios de dados. Os três principais já usados como referencial para orientar a implantação de um repositório confiável são:

- Princípios FAIR;
- Modelo de Referência OAIS;
- Certificação *CoreTrustSeal*.

2.6.2.1 Princípios FAIR

Conforme introduzido na Seção 2.2, um bom repositório de dados de pesquisa deve ser capaz de suprir as necessidades de descrever, preservar e permitir a descoberta dos dados. Para auxiliar nessa tarefa, em 2014 foi proposto o termo FAIR. Dois anos após os princípios resultantes foram publicados (FORCE11, 2016), indicando que os dados devem ser localizáveis (*Findable*), acessíveis (*Accessible*), interoperáveis (*Interoperable*) e reutilizáveis (*Re-usable*). Este é um dos requisitos altamente desejáveis pelo projeto do GT-RDP Brasil.

Segundo FORCE11 (2016), os princípios orientadores FAIR são os seguintes.

Para o dado ser localizável:

- F1 - os (meta)dados estão atribuídos a um identificador globalmente exclusivo e eternamente persistente;
- F2 - os dados estão descritos com metadados ricos;
- F3 - os (meta)dados estão registrados ou indexados em um recurso pesquisável;
- F4 - os metadados especificam o identificador de dados.

Para o dado ser acessível:

- A1 - os (meta)dados são recuperáveis pelo seu identificador usando um protocolo de comunicação padronizado;
 - A1.1 - o protocolo é aberto, gratuito e universalmente implementável;
 - A1.2 - o protocolo permite um procedimento de autenticação e autorização, quando necessário;
- A2 - os metadados estão acessíveis, mesmo quando os dados não estão mais disponíveis.

Para o dado ser interoperável:

- I1 - os (meta) dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento;
- I2 - os (meta)dados usam vocabulários que seguem os princípios FAIR;
- I3 - os (meta)dados incluem referências qualificadas a outros (meta)dados.

Para o dado ser reutilizável:

- R1 - os meta(dados) tem uma pluralidade de atributos precisos e relevantes;
 - R1.1 - os (meta)dados são liberados com uma licença de uso de dados clara e acessível;
 - R1.2 - os (meta)dados estão associados à sua proveniência;
 - R1.3 - os (meta)dados atendem aos padrões da comunidade relevantes ao domínio.

2.6.2.2 *Modelo de Referência OAIS*

Atualmente, existem diversos modelos de referência e princípios que regulam e orientam a implantação de um repositório. Um dos mais importantes é o Modelo de Referência OAIS - *Open Archival Information System* (CCSDS, The Consultative Committee for Space Data Systems, 2012), um modelo conceitual desenvolvido pelo *Consultive Committee for Space Data Systems* – CCSDS ³³, já normatizado na ISO 1472:2003³⁴, que disciplina e orienta um repositório na preservação e manutenção do acesso a informações digitais por longo prazo.

Para um repositório seguir a norma OAIS precisa ser composto por pessoas e sistemas com a responsabilidade de preservar a informação e torná-la disponível (CONARQ, Câmara Técnica de Documentos Eletrônicos - Conselho Nacional de Arquivos, 2015). O modelo, representado na Figura 4, especifica que o ambiente do repositório deve ser formado pelo produtor, pelo consumidor e pelo administrador (gestão), que precisam ser identificados e compreendidos.

SOUZA et al. (2012) descreve como ocorre o fluxo de pacotes de informação no modelo OAIS, onde cada entidade é responsável por parte do processo:

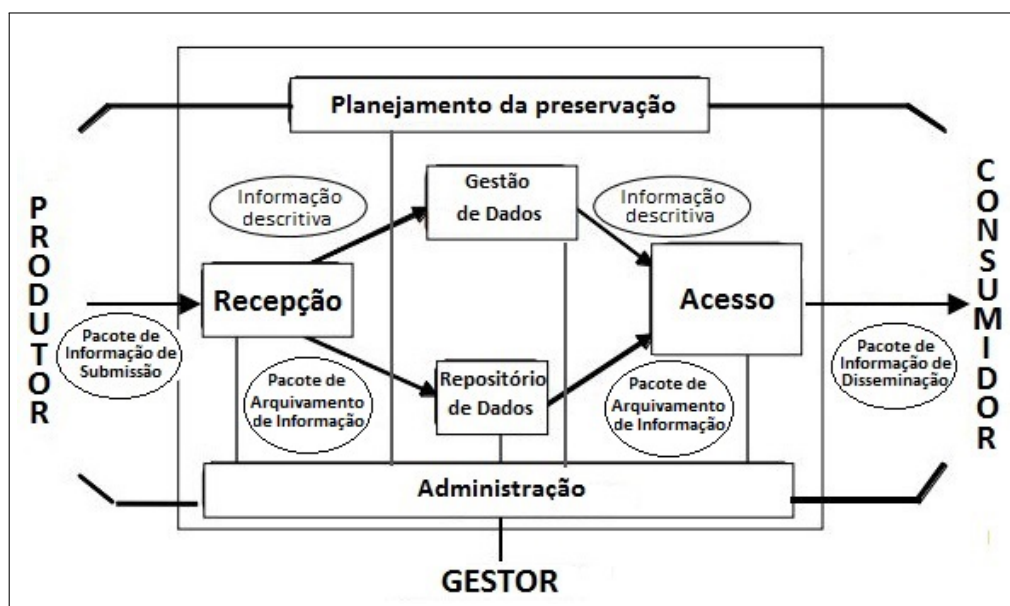
- o “produtor”, que pode ser uma pessoa ou sistema, fornece a informação, descrita como um SIP (Pacote de Submissão de Informação);
- a “recepção” recebe o SIP e transfere as informações descritivas para a “gestão de dados” e gera um ou mais AIP (Pacote de Arquivamento de Informação) para ser armazenado e preservado;
- quando o “consumidor” solicita uma informação o “acesso” é responsável por auxiliá-lo para determinar a existência, descrição, localização e disponibilidade da informação armazenada;

³³Comitê formado pelas maiores agências espaciais do mundo, com o objetivo de oferecer um fórum para discussão de problemas comuns sobre o desenvolvimento e a operação de sistemas de dados espaciais.

³⁴<https://www.iso.org/standard/57284.html>

- se a informação existir e puder ser fornecida o o DIP (Pacote de Informação de Disseminação) é formado por um ou mais AIP junto com suas informações descritivas que então é entregue ao consumidor;
- a entidade “administração” é responsável por gerenciar toda rotina operacional do sistema.

Figura 4: Modelo de Referência OAIS



Fonte: CCSDS, The Consultative Committee for Space Data Systems (2012)

É importante salientar que na pesquisa compreendida nesta dissertação, apesar de os desenvolvedores de cada ferramenta afirmarem que sim, não foi encontrada nenhuma publicação que ateste que quaisquer das soluções avaliadas nesse trabalho cumpram ou não todos os requisitos do modelo OAIS. Isso decorre inclusive de que muito do que o modelo preconiza é responsabilidade do processo de gestão do repositório e das políticas aplicadas, portanto a solução tecnológica deve prover meios, sendo que o produtor e o gestor devem efetivamente aplicá-los.

2.6.2.3 CoreTrustSeal

A certificação *CoreTrustSeal* fornece uma certificação de nível básico a repositórios de dados. A *World Data System of the International Science Council* (WDS) e a *Data Seal of Approval* (DSA) se uniram para substituir suas respectivas certificações por uma única, refletindo as principais características dos repositórios de dados confiáveis, tudo sob a égide da *Research Data Alliance* (RDA).

Segundo seu sítio³⁵, *CoreTrustSeal* é uma organização internacional, baseada em comunidades, não governamental e sem fins lucrativos que promove infraestruturas de dados sustentáveis e confiáveis.

A certificação é baseada em 16 requisitos que refletem características próprias de repositórios confiáveis, sendo todos eles obrigatórios, independentes e possuindo pesos iguais. Na página *Web* da certificadora³⁶ foram encontradas as principais soluções tecnológicas para repositórios estudadas nesta dissertação: *Dataverse*³⁷ e *DSpace*³⁸, confirmando que as ferramentas cumprem os requisitos exigidos pela certificadora.

2.6.3 Sistemas de gestão de dados de pesquisa

Entre as várias tecnologias disponíveis que sustentam e apoiam as boas práticas e as diretrizes do modelo de referência OAIS e os princípios FAIR, algumas estão se tornando padrões e sendo adotadas pelos principais *softwares* de repositórios atuais. Dois deles, relativos à submissão e recuperação automatizadas, são os protocolos SWORD e OAI-PMH.

SWORD (*Simple Web-service Offering Repository Deposit*) é um protocolo leve para depositar conteúdo de um local para outro. Ele possibilita a integração entre sistemas que publicam informações e repositórios, permitindo a um *software* externo verificar, adicionar, alterar, excluir e publicar arquivos e metadados no repositório. O protocolo consiste em dois estágios: primeiro é feita a autenticação e o retorno dos detalhes do que o usuário pode submeter, a seguir vem o depósito do arquivo da coleção escolhida. Após as devidas verificações dessas duas etapas, por exemplo, conferência das credenciais, compatibilidade de formatos, verificação de integridade (MD5), etc., o repositório enviará uma resposta indicando o sucesso ou não do procedimento.

O protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) é um protocolo desenvolvido pela *Open Archives Initiative*, baseado nas normas abertas *Hypertext Transport Protocol* (HTTP) e *Extensible Markup Language* (XML), que define um mecanismo para coleta de registros de metadados em repositórios. Este protocolo estabelece dois diferentes tipos de agentes:

- provedores de dados (*data providers*) - quem, de fato, contém o objeto digital que será disponibilizado na rede;
- provedores de serviço (*service providers*) - quem realiza a coleta da descrição dos documentos. Esta operação é denominada *harvesting* dos metadados.

As funcionalidades esperadas em um repositório de dados de pesquisa vão muito além

³⁵<https://www.coretrustseal.org/about/>

³⁶<https://www.coretrustseal.org/why-certification/certified-repositories/>

³⁷<https://www.tilburguniversity.edu/about/university-library/about-university-library/research-support/dataverse-nl>

³⁸<https://dspace-clarin-it.ilc.cnr.it/>

de simplesmente armazenar e permitir a recuperação destes dados. A Tabela 1 relaciona as funções de um sistema de gestão de dados de pesquisa (SAYÃO; SALES, 2016).

A questão que fica é a dissociação entre a riqueza dos metadados e a interoperabilidade requerida pelo FAIR. A relação entre quantos metadados serão necessários para descrever uma pluralidade de dados e a capacidade desses (meta)dados serem interoperáveis é inversamente proporcional na atualidade. Isso poderá ser minimizado com a extensão e padronização dos metadados necessários mas mesmo assim pode ser uma meta difícil de se alcançar (ERP et al., 2018).

2.6.4 Soluções tecnológicas

Com a conclusão da primeira etapa entrega do projeto do GT-RDP Brasil, foi entregue um relatório (PAVÃO et al., 2018) que identificou repositórios brasileiros de dados de pesquisa e ao mesmo tempo procurou-se identificar características quanto aos seguintes itens:

- âmbito (nacional ou cooperação internacional);
- áreas temáticas de abrangência;
- número de instituições envolvidas na coleta e disponibilização de dados;
- políticas do repositório (institucional, direitos autorais, licenças de uso);
- formato dos dados;
- interoperabilidade e padrões de metadados;
- *software* utilizado, aplicações e interfaces;
- existência de informações quanto às orientações para elaboração do plano de gestão de dados e ao ciclo de vida da pesquisa.

Em alguns casos não foi possível identificar algumas características acima, pois não foram encontradas tais informações nas páginas desses repositórios.

Dos 15 repositórios analisados, 4 são de abrangência internacional, envolvendo várias instituições, 7 são de abrangência nacional e 5 são de abrangência multi-institucional. Dentre os *softwares* utilizados, verificou-se que apenas 8 citam a solução adotada, nos outros 7 não há menção de que ferramentas foram usadas ou implementadas.

Esse levantamento revelou que existem poucas iniciativas nacionais de compartilhamento de dados de pesquisa e que era urgente tratar esse assunto de forma mais sistemática.

Tabela 1: Funções de um sistema de gestão de dados de pesquisa

CAPTURA DE DADOS	
Funções	<ul style="list-style-type: none"> • Seleção dos dados passíveis de serem arquivados; • Verificação do enquadramento no escopo do repositório; • Verificação dos formatos de arquivos aceitáveis para submissão; • Verificação dos direitos associados às coleções (copyright e licenças); • Verificação de dados sensíveis (dados não anonimizados, confidenciais, pessoais); • Verificação do volume e quantidade de arquivos; • Verificação dos metadados gerais e disciplinares que acompanham os dados; • Normalização para elenco de formatos padronizados aceitos para arquivamento e disseminação; • Controle de qualidade dos dados; • Definição de tempo de embargo.
CATALOGAÇÃO DAS COLEÇÕES DE DADOS	
Funções	<ul style="list-style-type: none"> • Atribuição de: metadados descritivos, estruturais, administrativos, técnicos (que inclui os relativos às dependências técnicas dos objetos digitais); • Atribuição de metadados de preservação, que assegurem a proveniência, autenticidade e integridade dos dados ao longo do tempo; • Uso de taxonomias especializadas e disciplinares; • Atribuição de identificador persistente (DOI, Handles, UNF, URN, etc.) que permita que os dados possam ser localizados de forma persistente e citados como as publicações acadêmicas; • Identificação do autor (ORCID ID, Scopus Author ID, ResearcherID, etc.); • Inclusão de documentação sobre os dados, incluindo descrição do projeto, dos arquivos e dos parâmetros; cadernos de laboratório e de campo, protocolos de pesquisa ou metodologia, etc.; • Vinculação (por links) a publicações e a dados relacionados internos e externos ao repositório.
ARQUIVAMENTO E PRESERVAÇÃO	
Funções	<ul style="list-style-type: none"> • Armazenamento em sistemas seguros; • Gestão da preservação de curto prazo (backups, backups redundantes offsite; checagem de integridade, armazenamento seguro, criptografia, compressão); • Gestão de longo prazo (migração, emulação, reformatação para formatos padronizados, aplicação de normas pertinentes (OAIS, TRAC), informação de fixidade voltada para validar a autenticidade e integridade de um objeto digital (checksums, assinatura digital); • Implementação de trilhas de auditoria.
INTEROPERABILIDADE	
Funções	<ul style="list-style-type: none"> • Disponibilização de metadados segundo o protocolo OAI-PMH; • Agregação para formação de publicações ampliadas segundo o padrão OAI-ORE; • Uso dos padrões, <i>web service</i>, <i>linked data</i> e outros • Empacotamento de metadados para intercâmbio segundo o padrão METS;
RECUPERAÇÃO, ACESSO E REUSO	
Funções	<ul style="list-style-type: none"> • Disponibilização de interfaces <i>Web</i> para recuperação, acesso e downloading; • Oferta de aplicações e serviços sobre as coleções;

Fonte: SAYÃO; SALES (2016)

3 USABILIDADE

Este capítulo introduz e conceitua usabilidade, descreve e categoriza as metodologias e técnicas mais conhecidas. Ao final deste capítulo é descrita e explicada a escolha das técnicas de avaliação utilizadas neste estudo.

Como visto na Seção 2.4, a clara visão de todos os envolvidos sobre os benefícios do acesso aberto é um dos fatores determinantes para o sucesso de um repositório de dados de pesquisa. Mas além disso, a solução tecnológica adotada precisa ser realmente aceita pela comunidade. Portanto, para o pleno uso dos repositórios de dados de pesquisa é necessário avaliar a satisfação do usuário, que está diretamente ligada à usabilidade do sistema (WINCKLER; PIMENTA, 2002).

Idealmente, segundo DIX et al. (2004), a avaliação deve ocorrer durante todo o ciclo de vida do projeto, com modificações sendo feitas de acordo com os resultados dessas avaliações. Ele enfatiza as feitas nos estágios iniciais de desenvolvimento onde é muito mais fácil alterar um projeto e antes de dispendir esforços e recursos em implementações com problemas.

Quando os primeiros programas interativos surgiram, eles eram confeccionados para uso quase exclusivo dos seus próprios desenvolvedores. Mais tarde outros usuários passaram a dispor de soluções informáticas e tinham acesso a treinamento intensivo na ferramenta que normalmente eram mais focadas na solução do problema do que na interface. Quando finalmente o grande público teve acesso aos programas de computador, começou-se a discutir problemas de usabilidade (CYBIS; BETIOL; FAUST, 2007).

Até hoje muitas pessoas associam a incapacidade de entender e utilizar um programa com sua própria incapacidade de entender a tecnologia. Nos ambientes corporativos onde existe a obrigatoriedade do uso de um *software* específico, imposto pela chefia, a frustração e o estresse pode culminar em doenças físicas e psicológicas. Este fato gera prejuízo para as empresas, que perde produtividade. Esses problemas geram alta rotatividade de funcionários e custos adicionais de treinamento. Portanto, o baixo investimento no desenvolvimento das interfaces de usuário tem consequências diretas. Segundo CYBIS; BETIOL; FAUST (2007), “o barato terá custado caro”.

DIX et al. (2004) afirma que as pessoas utilizam um computador para realizar um

trabalho e isso descreve as três questões que mais importam: as pessoas, os computadores e as tarefas. Os autores lembram então que o sistema deve suportar a tarefa do usuário surgindo assim um quarto foco: a usabilidade. Portanto, se o usuário for forçado a adotar um modo de trabalho inaceitável, ele não o utilizará.

A interface de um *software* é muito mais que um aspecto visual agradável. Para BON-SIEPE (1997), a interface revela o caráter de ferramenta dos objetos e o conteúdo comunicativo das informações. Além disso, ela transforma objetos em produtos e sinais em informação interpretável, modificando a simples presença física em disponibilidade.

3.1 Definição

A definição dada pela norma ISO 9241 (International Organization for Standardization, 1998), tanto pela versão original quanto sua revisão (International Organization for Standardization, 2018), assim como pela versão brasileira NBR 9241 (Associação Brasileira de Normas Técnicas, 2011, p. 11), é de que usabilidade é a “medida na qual um produto pode ser usado por usuários específicos para alcançar objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso”. Essa definição enfatiza que a usabilidade de um produto é afetada não apenas pelas características do próprio produto, mas também pelas circunstâncias específicas em que um produto é usado (MAGUIRE, 2001).

Segundo NIELSEN (1993), a usabilidade não é uma propriedade única e unidimensional de uma interface de usuário. Ela tem múltiplos componentes e está associada a cinco atributos distintos:

- Aprendizagem – o sistema deve ser de fácil aprendizado para que o usuário possa começar a utilizá-lo rapidamente.
- Eficiência – o sistema deve ser eficiente para que o usuário possa alcançar um alto nível de produtividade assim que aprenda a utilizá-lo.
- Memorização – o sistema deve ser fácil de lembrar, para que um usuário casual possa utilizá-lo novamente sem ter que aprender tudo novamente após um tempo sem usá-lo.
- Erros – a taxa de erros deve ser baixa, erros catastróficos não devem ocorrer. Mesmo que um usuário cometa algum erro, ele deve poder recuperar-se facilmente.
- Satisfação – o sistema deve ser agradável de usar, para que os usuários fiquem satisfeitos ao usá-lo.

Segundo LAVERY; COCKTON; ATKINSON (1997), um problema de usabilidade é um aspecto do sistema ou uma demanda imposta ao usuário que torna desagradável, ine-

ficiente, oneroso ou impossível para o usuário atingir os seus objetivos em uma situação normal de uso.

Segundo CYBIS; BETIOL; FAUST (2007), a usabilidade “é a qualidade que caracteriza o uso dos programas e aplicações”, portanto ela não é intrínseca ao sistema. A usabilidade varia de acordo com a expertise do usuário, a frequência de uso e até mesmo com a capacidade do computador em que está sendo operado. Os autores ainda afirmam que “a essência da usabilidade é o acordo entre interface, usuário, tarefa e ambiente” e por isso essa interação é uma composição flexível entre aspectos objetivos, relacionados à produtividade, e subjetivos, relacionados ao prazer do usuário.

A usabilidade deveria ser uma qualidade intrínseca a todo sistema pois “se a informação existe para servir ao seu público-alvo, seria primordial aumentar cada vez mais a usabilidade das interfaces das ferramentas tecnológicas, permitindo que os usuários recuperem a informação de maneira eficaz, eficiente e satisfatória” (KAFURE; CUNHA, 2006).

O especialista em usabilidade Steve Krug responde uma frequente pergunta sobre qual é a coisa mais importante para assegurar a usabilidade de um site ou aplicativo, com a exclamação, e título de seu livro, “Não me faça pensar!”. Ele ainda declara que essa é a sua primeira lei de usabilidade (KRUG, 2014, p. 9).

3.2 Métodos de avaliação

Métodos de avaliação de usabilidade possuem características próprias, podendo ser classificados de acordo com estas características. Eles podem ser aplicados em diversas etapas do processo de desenvolvimento do sistema e podem envolver ou não usuários e especialistas (PEREIRA, 2011).

Essa diversidade serve para constatar que não existe um único método capaz de identificar todos os problemas de usabilidade possíveis em uma interface, e sim que mais de um método deve ser utilizado para cobrir um espectro maior de problemas, além disso, devem ser aplicados regularmente devido as mudanças na interface e nas características dos usuários (WINCKLER; PIMENTA, 2002).

Apesar de encontramos várias classificações e agrupamentos desses métodos na literatura, este trabalho considera essas técnicas agrupadas de acordo com a participação ou não dos usuários na avaliação de usabilidade, classificação seguida por vários autores, entre eles DIX et al. (2004), PEREIRA (2011), GRAY; SALZMAN (1998) e JORDAN (1998), sendo que o último declara, em sua experiência, que não há substituto quanto a ver pessoas tentando usar um produto.

Considerando esta característica, os métodos podem ser classificados em: empíricos, quando há a participação dos usuários; e analíticos, quando não há.

Segundo NASCIMENTO; AMARAL (2010), além da grande variedade de métodos e

técnicas, e os autores citam alguns exemplos encontrados na literatura que suportam a premissa de que não existe padronização na terminologia adotada. Além dessa diversidade de nomenclatura, alguns autores chamam a atenção para as variações desses métodos e os nomes diversos para técnicas parecidas ou adaptadas (WINCKLER; PIMENTA, 2002; CYBIS; BETIOL; FAUST, 2007; PEREIRA, 2011).

3.2.1 Métodos empíricos

As técnicas objetivas ou empíricas, segundo PEREIRA (2011), exigem versão já implementada e funcional do sistema, pois dependem da interação do usuário e/ou da opinião dele.

Alguns métodos podem envolver equipamentos especiais e até mesmo salas específicas para esse fim, configuradas de acordo com cada teste efetuado (CYBIS; BETIOL; FAUST, 2007). Os autores também descrevem laboratórios de usabilidade que são dotados de salas especiais equipadas com vidros espelhados, microfones, câmeras que possam filmar vários ângulos da interação e *softwares* específicos para registro, marcação e análise dos acontecimentos. Isso diminui o constrangimento do participante, pois apesar de saber que está sendo observado, ele não vê os avaliadores olhando para ele. Um facilitador pode participar do processo. Ele pode ser alguém próximo do usuário para auxiliar o desenvolvimento do trabalho em vários níveis ou ainda dar suporte ao usuário por meio de um canal de comunicação como telefone ou *chat*.

Entre essas técnicas empíricas, as mais conhecidas são a observação direta e a observação indireta, sendo usadas como base para vários outros métodos. No livro de ROCHA; BARANAUSKAS (2003) é feita uma comparação entre os dois métodos de observação. Segundo as autoras, na observação direta, que é considerado o método mais invasivo, o avaliador permanece ao lado do usuário e faz anotações em tempo real da interação deste com o sistema, podendo fazer outros registros, como o tempo de realização de certas tarefas. Essa observação pode constranger o usuário e prejudicar a avaliação, principalmente por ele estar consciente de estar sendo observado e monitorado. Outra consequência negativa é a dificuldade de coleta de dados já que é manual e em tempo real. Por outro lado, elas citam o “efeito Hawthorne”, que causaria melhora no desempenho do avaliado, consequência dele se sentir valorizado por receber atenção de terceiros, no caso o avaliador. Em relação à observação indireta, as autoras a consideram menos invasiva, mesmo que o usuário esteja consciente o tempo todo da observação. Uma gravação em vídeo em conjunto com outros registros da interação do usuário com o sistema, tais como *logs*, formarão um conjunto de dados bastante rico que poderá ser amplamente analisado mais tarde. A principal desvantagem dessa técnica é que é necessário um laboratório de alto custo, plenamente equipado.

NIELSEN (1994) afirma que “métodos empíricos são a principal maneira de avaliar as interfaces do usuário”. Já STANTON et al. (2005), por sua vez, complementa dizendo

que “a observação das atividades, combinadas com entrevistas com as pessoas que executam essas atividades, são as principais abordagens no kit de ferramentas do ergonômista”. Segundo este autor, as entrevistas podem ir desde completamente não estruturadas (conversa informal), passando por semi-estruturadas (seguindo agenda de perguntas, mas com liberdade para tópicos não programados), indo até completamente estruturadas (seguindo rigorosamente um questionário preconcebido). A abordagem semi-estruturada é recomendada para avaliação de um produto, mas a completamente estruturada permite tabular melhor as respostas já que todos os usuários terão opiniões sobre as mesmas questões.

Variações dos métodos observacionais e de entrevista são comuns, alguns são largamente mencionados em trabalhos da área (CYBIS; BETIOL; FAUST, 2007; PEREIRA, 2011; MACHADO; FERREIRA; VERGARA, 2014).

Uma técnica de entrevista que também serve de base para outras variações é conhecida como *thinking aloud protocol* ou protocolo verbal. O usuário é solicitado para interagir com o sistema segundo um roteiro predefinido e então verbalizar seus pensamentos, sejam opiniões ou meros sentimentos. Para muitas pessoas essa não é uma atitude muito natural então os usuários precisam ser estimulados pelo avaliador (WINCKLER; PIMENTA, 2002). No livro de CYBIS; BETIOL; FAUST (2007), onde nomeiam essa técnica como “verbalização simultânea”, é sugerido que o avaliador pergunte ao usuário, no decorrer do ensaio, questões tais como:

- Poderia me dizer o que está pensando?
- O que está tentando fazer?
- Por que fez isso?

Esta técnica necessita de meios eficientes de registro das interações e verbalizações. STANTON et al. (2005) recomenda a gravação de áudio e vídeo para uma melhor análise dos dados.

Já a técnica de codescoberta, que é similar ao protocolo verbal, é realizada por dois participantes que interagem e realizam tarefas segundo a orientação de um avaliador. A troca de opiniões e pensamentos e a superação de dificuldades em conjunto permitem a codescoberta de problemas na interface do sistema (PEREIRA, 2011 apud DIAS, 2007).

CYBIS; BETIOL; FAUST (2007) descrevem outra técnica alternativa ao protocolo verbal, que chamam de *verbalização consecutiva*. Nesta técnica, os usuários são questionados após a execução das tarefas, ao mesmo tempo em que visualiza o vídeo de sua interação. Desta forma, o usuário percebe o que fez, como fez e o que esperava fazer.

O método de medida de desempenho não envolve diretamente uma entrevista com o usuário. Em sua forma simplificada, mede apenas o tempo gasto na realização de tarefas (eficiência) e se foram realizadas corretamente (eficácia). Podem ser adotadas várias

métricas: rapidez na correta execução, erros, frustrações, desistências, etc. (PEREIRA, 2011 apud DIAS, 2007).

Alguns métodos empíricos podem não envolver diretamente a observação da interação do usuário com o sistema, mas se baseiam nesta observação. Exemplos dessas técnicas são o *focus group* e o *card sorting*.

O *focus group*, grupo focal ou grupo de foco, consiste em uma conversa entre usuários, geralmente variando entre 6 e 12 (CYBIS; BETIOL; FAUST, 2007; STANTON et al., 2005), sendo preferível convidar participantes com origens e perfis diferentes para colher opiniões diversificadas. O grupo é guiado por um entrevistador ou moderador seguindo um roteiro, e ele deve manter o foco da discussão garantindo a independência de opiniões e a contribuição de todos. Por se tratar de um técnica subjetiva onde as opiniões são expressadas livremente fica mais difícil mensurar os dados resultantes, mas a vantagem é a descoberta de características atrativas e desagradáveis muitas vezes não percebidas e nem imaginadas pelo responsável pela montagem do roteiro de entrevista. Tanto CYBIS; BETIOL; FAUST (2007) quanto KRUG (2014) sugerem que essa técnica não é apropriada para avaliações, já que não vemos os usuários usando sistemas ou produtos, apenas os ouvimos falando sobre eles. Entretanto, os autores ressaltam que a técnica é melhor aproveitada nas etapas de planejamento.

Já o *card sorting* é um método que é útil para descobrir e entender o modelo mental do usuário. Nele os usuários recebem um conjunto de cartões, cada um representando uma parte do sistema, e pede-se que eles os organizem de modo que sejam úteis para eles, possibilitando a criação de estruturas de informação adequadas aos seus modelos mentais (NASCIMENTO; AMARAL, 2010). Modelos mentais podem ser definidos como a representação do conhecimento dos usuários sobre como funciona e de como usar um sistema (CAFÉ; MUÑOZ, 2016). Os modelos mentais são “individuais, incompletos e dinâmicos no tempo pois representam apenas as partes de um sistema com as quais usuários específicos tiveram contato” (CYBIS; BETIOL; FAUST, 2007). Cada usuário vai conceber um modelo mental inicial baseado na primeira impressão do sistema e em suas experiências e interações anteriores. Quanto mais usar a interface, mais aprimorado vai ficando esse relacionamento e, por consequência, melhor o modelo mental.

Como já citado na introdução desta seção, é recomendada a aplicação de mais de uma técnica para alcançar melhores resultados na avaliação. Por exemplo, NASCIMENTO; AMARAL (2010) sugere a utilização de um método observacional após a utilização da técnica de *card sorting* para mapear caminhos e tarefas do sistema.

Todo teste com usuários envolve questões éticas. Quem está sendo testado é a interface, e não o usuário. Portanto, ele deve ser informado claramente dos procedimentos. Como estes métodos são estressantes por natureza, o avaliador deve criar um clima agradável para deixá-lo à vontade. Além disso, o usuário deve poder abandonar o teste a qualquer momento (WINCKLER; PIMENTA, 2002).

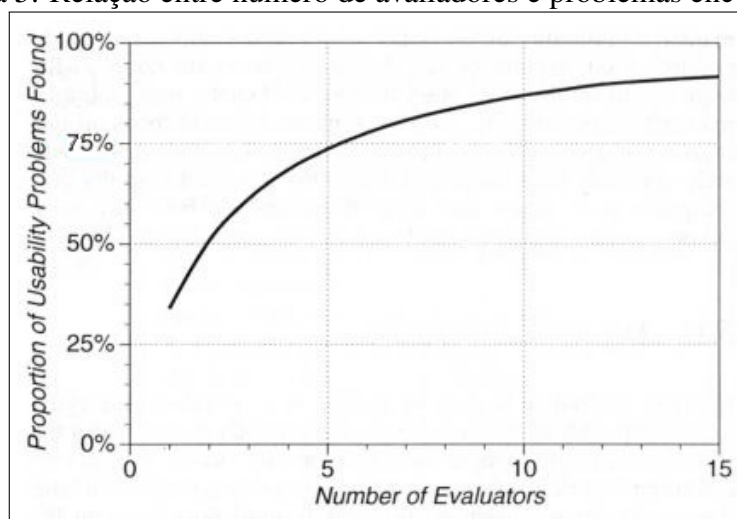
3.2.2 Métodos analíticos

As técnicas analíticas, que não envolvem usuários, são realizadas diretamente pelos avaliadores, normalmente projetistas e especialistas na área de domínio da aplicação. Elas podem ser aplicadas em qualquer etapa do desenvolvimento do *software*, inclusive quando já em produção. MACHADO; FERREIRA; VERGARA (2014) citam a avaliação heurística e o percurso cognitivo como os métodos analíticos mais utilizados na análise de interfaces.

Segundo NIELSEN; MOLICH (1990), a inspeção heurística é realizada ao se tentar chegar a uma opinião sobre uma interface a partir apenas da observação. O ideal seria usar regras listadas em documentos padrões de diretrizes de usabilidade, mas como algumas podem ter mais de mil regras podem ser intimidadoras para a maioria dos desenvolvedores. Sendo assim, a maioria das avaliações heurísticas são feitas com base na experiência e intuição dos avaliadores (NIELSEN, 1993).

Cada avaliador inspeciona a interface individualmente. Após todas as avaliações serem concluídas, os avaliadores podem se comunicar, desta forma garantindo independência e imparcialidade. NIELSEN (1993) sugere que sejam utilizados de 3 a 5 avaliadores e cita que em média apenas 35% dos problemas de usabilidade são descobertos por apenas um avaliador, esse percentual sobe para 75% quando são usados 5 avaliadores. O aumento significativo do número de avaliadores não reflete proporcionalmente na detecção de problemas de usabilidade, conforme demonstrado na Figura 5.

Figura 5: Relação entre número de avaliadores e problemas encontrados



Fonte: (NIELSEN, 1993).

O método de percurso cognitivo ou inspeção cognitiva, proposto em 1992 por POLSON et al. (1992), é uma técnica indicada para identificar problemas decorrentes de lógicas inadequadas (CYBIS; BETIOL; FAUST, 2007). Esse método vai além de apenas verificar a conformidade ergonômica e tenta avaliar o processo de aprendizagem (MA-

CHADO; FERREIRA; VERGARA, 2014). Assim como outras técnicas analíticas, ela se baseia em uma lista de verificações que levam os avaliadores a analisar as condições oferecidas pelo sistema para a realização das tarefas propostas (CYBIS; BETIOL; FAUST, 2007). DIX et al. (2004) lembra que é importante uma indicação da gravidade de cada problema encontrado, ajudando os projetistas a priorizarem a correção destes problemas. Apesar de ser um método analítico, o percurso cognitivo pode ser aplicado também em usuários sem experiência em usabilidade, como foi o caso de SILVA (2017). Em suas conclusões, o autor considerou o método bastante eficiente, permitindo uma análise mais completa do site avaliado.

SMITH; MOSIER (1986) propuseram um dos mais antigos métodos de avaliação de usabilidade desenvolvido para sistemas computacionais. Os autores definiram um conjunto de 944 diretrizes (*Guidelines*) para o projeto de *software* de interface de usuário, distribuídas em seis áreas funcionais de interação entre usuário e sistema. Apesar de criadas para serem usadas como guias durante as fases de projeto e desenvolvimento do *software*, elas são aplicadas em avaliações de usabilidade em técnicas de revisões de *guidelines* (MARTINS NETTO, 2002; NASCIMENTO; AMARAL, 2010). Quase 20 anos depois, NIELSEN (2005) faz uma revisão do trabalho citado e analisa 10 diretrizes de cada área, como resultado afirma que 90% daquelas diretrizes propostas ainda eram válidas, apesar de identificar que parte delas se referem a elementos de interface não mais utilizados. Entretanto, este conjunto é pequeno quando comparado as 3.507 diretrizes identificadas pelo *Nielsen Norman Group*.

A técnica de inspeção por listas de verificação, segundo CYBIS; BETIOL; FAUST (2007), tem a vantagem de permitir que profissionais, não necessariamente especialistas em ergonomia, identifiquem problemas menores e repetitivos das interfaces. Ainda, segundo os autores, esse tipo de técnica se baseia mais nas qualidades explícitas da ferramenta ou sistema e bem menos nos conhecimentos implícitos dos avaliadores, como na técnica heurística. Por exemplo, o projeto ErgoList¹, coordenado por Walter Cybis, disponibiliza *checklists* que possibilitam a descoberta de falhas ergonômicas mais flagrantes em uma interface. Entretanto, como o próprio sítio avisa, o uso dos módulos disponibilizados é recomendado apenas para uma aplicação didática e sugere o uso de listas baseadas na norma ISO 9241.

A escolha de um ou mais métodos de avaliação se dá em função de vários fatores, entre os quais se destacam: a fase do desenvolvimento do sistema; a disponibilidade de avaliadores, especialistas ou não; e outros recursos como financeiro, tempo e equipamentos.

Seguindo principalmente os métodos focados no usuário abordados por KRUG (2014) e NIELSEN; MOLICH (1990), que propõe avaliações com especialistas utilizando suas 10 heurísticas (CORRÊA; MIRANDA, 2012), o estudo apresentado nesta dissertação uti-

¹<http://www.labiutil.inf.ufsc.br/ergolist/>

liza duas técnicas de avaliação de usabilidade: uma heurística, conduzida por pesquisadores especialistas em Ciência da Informação da FURG, e outra empírica, com usuários alunos e pesquisadores da mesma universidade. A escolha dos especialistas para a avaliação analítica foi motivada principalmente pela questão do custo associado à contratação de uma empresa especializada em usabilidade ou à compra de equipamentos para montar um laboratório específico para este fim. KRUG (2014, p. 113-114) afirma que, a partir da década de 90, os custos caíram de 20 a 50 mil dólares para de 5 a 10 mil dólares. Mesmo assim, o valor é proibitivo, portanto o autor sugere usar “testes de usabilidade do tipo faça-você-mesmo”. Apesar de não serem profissionais da área de usabilidade, a escolha dos mesmos deve-se ao grande interesse pela aplicação e ao conhecimento adquirido por esses pesquisadores ao longo de sua carreira na área fim do sistema a ser avaliado.

4 TRABALHOS RELACIONADOS

Este capítulo apresenta um conjunto de trabalhos relacionados ao tema e objetivos desta dissertação, os quais estão divididos em duas seções. A primeira apresenta comparativos entre *softwares* para repositórios de dados. A segunda seção discute trabalhos relacionados à avaliação da usabilidade de sistemas para repositórios.

4.1 Comparativos entre *softwares* para repositório de dados

Por ser um assunto novo, são encontrados poucos trabalhos sobre comparação de repositórios de dados de pesquisa. As buscas realizadas na BDTD do IBICT e no portal de periódicos da CAPES¹ não retornaram nenhum documento relevante sobre comparação de repositórios de dados de pesquisa que levassem em conta os *softwares* DSpace² e Dataverse³. Já na ferramenta Google Acadêmico⁴ foram encontrados alguns poucos trabalhos discutidos a seguir.

O trabalho de AMORIM et al. (2015) traça um comparativo entre DSpace, CKAN⁵, Zenodo⁶ e Figshare, mas não cita o Dataverse. Os mesmos autores, mais recentemente publicaram outro artigo (AMORIM et al., 2017) comparando 15 repositórios para dados de pesquisa, incluindo CKAN, DSpace e Dataverse, mas restringindo os critérios de comparação a poucos itens. Já SHANKAR MAHATO; GAJBE (2018) comparam somente Dataverse e CKAN.

Nos sítios dos desenvolvedores são disponibilizados comparativos entre essas soluções. Como estratégia de *marketing*, cada desenvolvedor estabelece seus próprios critérios e parâmetros, enaltecendo seus próprios recursos e funcionalidades frente aos sistemas concorrentes.

O Dataverse⁷ publica uma revisão comparativa entre oito softwares para repositório de

¹<http://www.periodicos.capes.gov.br/>

²<https://duraspace.org/dspace/>

³<https://dataverse.org/>

⁴<https://scholar.google.com>

⁵<https://ckan.org/>

⁶<https://zenodo.org/>

⁷<https://dataverse.org/blog/comparative-review-various-data-repositories>

dados para ajudar o grupo de desenvolvedores a descobrir tendências e evoluir a própria ferramenta. Está disponível um *link* para uma planilha eletrônica *online* com esse comparativo, onde é permitido que qualquer pessoa possa deixar comentários sobre imprecisões, problemas ou sugestões.

A página do CKAN⁸ descreve seus diversos recursos comparando-os ao Dataverse e Nesstar⁹, plataforma proprietária para dados de pesquisa, mas ressalta que embora ambos os concorrentes sejam plataformas bem estabelecidas, dedicadas especialmente a dados de pesquisa (o que o CKAN não é), nenhuma delas atende a maioria dos critérios de comparação com CKAN. Nesstar não é software livre, dependendo do sistema operacional Windows para submissão dos dados. Já o Dataverse não suportaria submissão via API. Entretanto, esse recurso já está implementado na versão atual.

4.2 Avaliação da usabilidade de sistemas de repositórios

Apesar do estudo sobre ergonomia e usabilidade ser bem difundido e encontrarmos vários artigos sobre esses temas, a usabilidade de repositórios ainda é incipiente. São poucos os trabalhos que tratam de repositórios institucionais e não fomos capazes de encontrar estudo sobre usabilidade de repositórios de dados de pesquisa, o que caracteriza uma das contribuições desta dissertação.

Uma busca no portal de periódicos da CAPES com as palavras chaves “usabilidade” e “repositório” em qualquer campo de busca recuperou 93 documentos. Já com os mesmos termos em inglês, foram retornados mais de 18 mil. Após verificar os artigos retornados percebeu-se que a maior parte dos documentos trazia pouca ou nenhuma informação relevante sobre usabilidade de repositórios. Alterando a estratégia de busca para “usabilidade” e “repositório” apenas no título, somente 2 artigos foram retornados. Com os mesmos termos em inglês, após filtrar duplicatas e documentos irrelevantes, somente 8 documentos foram encontrados. Por fim, na BDTD apenas 4 teses ou dissertações foram encontradas.

A maioria dos trabalhos relacionados avaliam repositórios institucionais de publicações utilizando métodos já consolidados. O repositório institucional da Universidade do Minho (UMinho), em Portugal, é avaliado na dissertação de SILVA (2014). O autor destaca que esta foi a primeira instalação a nível mundial de um repositório científico de língua portuguesa a entrar em funcionamento com base na plataforma DSpace. CAFÉ; MUÑOZ (2016) avaliam a usabilidade do repositório institucional da Universidade de Brasília (UnB). SANTOS; PAVÃO; MOURA (2016) avaliam o LUME, repositório institucional da UFRGS. As autoras relatam que foram encontrados poucos problemas na interface do repositório. KOSHIYAMA (2014) avalia o repositório institucional da Uni-

⁸<https://ckan.org/2013/09/25/edawax/>

⁹<http://www.nesstar.com/>

versidade Federal do Rio Grande do Norte (UFRN) onde encontra vários problemas de usabilidade e confirma a hipótese gerada no início da sua pesquisa de que a arquitetura da informação do sistema desfavorece a sua usabilidade. SEABRA (2018) avalia a usabilidade do repositório institucional da Universidade Federal da Bahia (UFBA) baseando-se nas heurísticas de Nielsen e também encontra diversos problemas na interface do sistema mas ressalta que testou uma versão antiga do *software* e muitos deles já foram corrigidos nas versões atuais. Todos os repositórios destas universidades são baseados no *software* DSpace.

Outros trabalhos propõem métodos de avaliação de usabilidade a os aplicam em repositórios científicos. Este é o caso de PALMA et al. (2018), que elaboram uma proposta de avaliação para repositórios digitais e a aplica no Repositório Institucional Digital do IBICT (RIDI¹⁰). Os sul-coreanos HEE KIM; HO KIM (2008) também propõem uma metodologia e a aplicam no *software* para repositórios *dCollection*¹¹, um projeto do Serviço de Informação de Educação e Pesquisa da Coreia do Sul (KERIS). Como o estudo foi financiado pelo KERIS, as sugestões apresentadas já começaram a ser implementadas e novos testes de usabilidade foram previstos.

Trabalhos mais específicos ainda foram encontrados na pesquisa relatada nesta dissertação. FURTADO (2016) avalia a usabilidade de um repositório da UFRN, baseado em DSpace, voltado para informação acessível. Foi considerada na avaliação a opinião de usuários com condições de baixa visão e cegueira. Outro estudo, conduzido por KÖRBER; SULEMAN (2008), aborda a usabilidade dos processos de instalação e configuração do DSpace. Os autores destacam problemas significativos e sugerem como podem ser resolvidos.

O artigo de VEIGA et al. (2014) faz uma revisão bibliográfica sobre usabilidade do DSpace através do método de avaliação cooperativa. Após as buscas iniciais e descartes de documentos não relevantes, restringe a análise para 14 estudos. Apesar da extensa busca inicial, as autoras não encontraram trabalhos sobre a aplicação do método escolhido em repositórios institucionais.

¹⁰<http://repositorio.ibict.br>

¹¹<http://www.dcollection.net>

5 METODOLOGIA PROPOSTA

Este capítulo descreve como foi feita a escolha da solução tecnológica para repositório de dados de pesquisa, quais ferramentas foram analisadas, e os métodos de avaliação da usabilidade utilizados e propostos. O projeto do GT-RDP Brasil adotou os seguintes pressupostos que também foram adotados nesse estudo:

- a solução tecnológica adotada deve ser capaz de implementar e configurar políticas de submissão, acesso e gestão de conjuntos de dados, estudos e grupos, sejam essas formais ou informais, inclusive representações organizacionais;
- os conjuntos de dados podem ser “dados da pesquisa”, produzidos durante a pesquisa ou “dados para a pesquisa”, disponibilizados por organizações para serem usados em pesquisa futura;
- a solução também deve respeitar ao máximo os princípios FAIR e OAIS.

Na busca por soluções de repositórios foi feita uma pesquisa no *Research Data Repositories Information (Re3Data)* ¹, que é um registro global de repositórios de dados de pesquisa de diferentes disciplinas acadêmicas e fornece informações sobre o armazenamento e acesso de conjunto de dados a pesquisadores, órgãos financiadores, editores e instituições. O registro foi lançado em 2012 e é financiado pela Fundação Alemã de Pesquisa (DFG) ². Ele é frequentemente usado como referência para identificação e escolha de um repositório apropriado, inclusive é recomendado pelo programa *Horizon 2020* (DATACITE, 2018).

O diretório Re3Data lista o DSpace e o Dataverse como dois dos *softwares* para repositórios mais utilizados. Um terceiro *software* bastante utilizado, CKAN, não foi considerado nesse trabalho, apesar de o repositório respeitar de maneira significativa a maioria dos princípios FAIR, por se tratar de um *software* mais apropriado à publicação de dados governamentais e organizacionais e não propriamente dados de pesquisa (ERP et al., 2018). Além disso, ele não implementa uma série de funcionalidades de gestão de dados

¹<http://www.re3data.org/>

²German Research Foundation - <http://www.dfg.de>

para preservação a longo prazo presentes no DSpace e Dataverse, tornando-se então uma solução mais adequada apenas como camada de publicação dos dados, aliada a outras soluções de repositório.

O primeiro, e talvez mais importante requisito para a escolha da solução tecnológica de repositório adotada no âmbito do projeto é que ela fosse gratuita, já que será indicado e suportado pela RNP e IBICT e instalado e mantido pelas instituições de ensino e/ou pesquisa, as quais não podem ficar reféns de políticas orçamentárias que não permitam a continuidade do projeto a longo prazo. Por isso, as soluções avaliadas não puderam conter nenhum componente proprietário pago. Mesmo que gratuito, qualquer componente de *software* deveria ser um projeto de *software* livre, para não ser dependente de um empresa desenvolvedora continuar distribuindo ou não o *software*. Também não poderiam ser serviços Web hospedados por terceiros, principalmente em uma nação estrangeira, sujeita a políticas e leis específicas.

5.1 Ferramentas avaliadas

O Dspace³ foi originalmente desenvolvido em um projeto colaborativo entre *Massachusetts Institute of Technology (MIT) Libraries* e *Hewlett-Packard Labs* sendo transferidos os direitos autorais para a *DSpace Foundation* em 2007 e posteriormente para a Duraspace, uma organização sem fins lucrativos independente. Na página oficial do Duraspace⁴ ainda é possível ver um aviso de que alguns códigos-fonte ainda podem conter uma declaração de direitos ao MIT e HP.

A Figura 6 exemplifica como podem estar representadas as estruturas funcionais de comunidades e coleções dentro do DSpace. Uma comunidade pode conter outras comunidades que por sua vez contêm coleções. Uma coleção é um conjunto de itens, que compreendem arquivos e seus metadados.

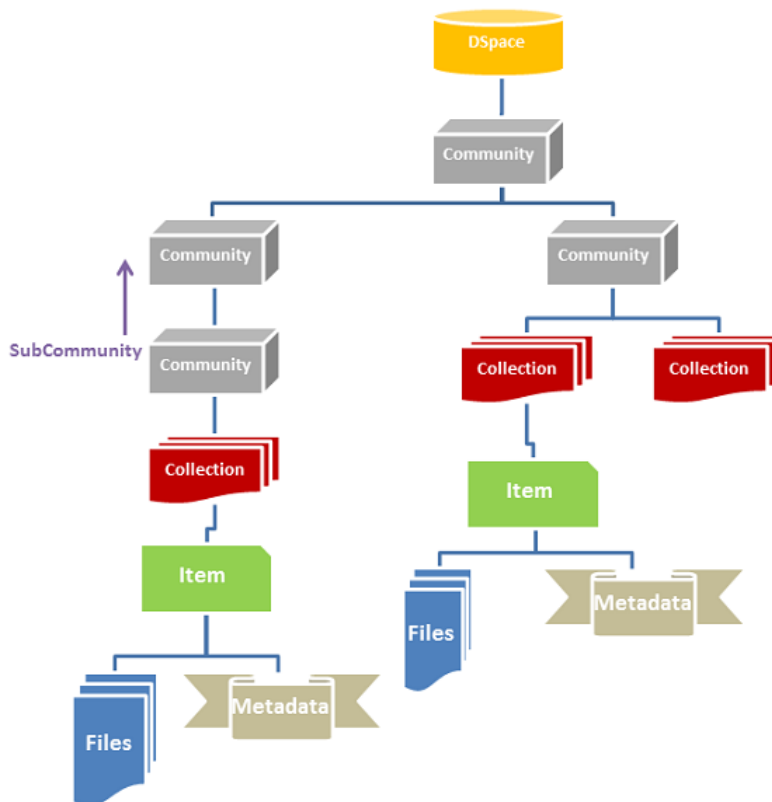
O Dataverse⁵ foi desenvolvido especificamente para compartilhar, preservar, citar, explorar e analisar dados de pesquisa. Ele é desenvolvido e mantido pelo *Institute for Quantitative Social Science (IQSS)* da Universidade de Harvard que lidera um grupo internacional de colaboradores e contribuidores. A Figura 7 mostra uma estrutura básica de um repositório Dataverse. Existe o ambiente *dataverse* principal e aninhado nele podem ser criados outros *dataverses* hierárquicos, tantos níveis quanto necessário (*subdataverses*). Cada *dataverse* possui uma administração e configuração própria independente dos demais. Eles são administrados por um pesquisador, grupo de trabalho ou comunidade com interesses mútuos que submetem conjuntos de dados (*datasets*). Cada conjunto de dados contém: os dados propriamente ditos, os descritores (metadados), metadados de citação, termos de uso e *guestbook*.

³<https://duraspace.org/dspace/resources/technical-specifications>

⁴<https://duraspace.org/dspace/license/>

⁵<https://dataverse.org/about>

Figura 6: Estrutura do Dspace



Fonte: <https://wiki.duraspace.org>

Tanto o DSpace quanto o Dataverse são bem parecidos em relação a estrutura e finalidade, sendo que o DSpace foi desenvolvido originalmente para o armazenamento, preservação e disponibilização de publicações textuais (documentos digitais) e o Dataverse para o de dados de pesquisa.

5.2 Avaliação da usabilidade

Para avaliar a usabilidade da solução tecnológica escolhida foi proposto o uso de dois métodos distintos: uma avaliação analítica, seguindo as heurísticas propostas por NIELSEN (1994), e outra empírica, com participação de pesquisadores.

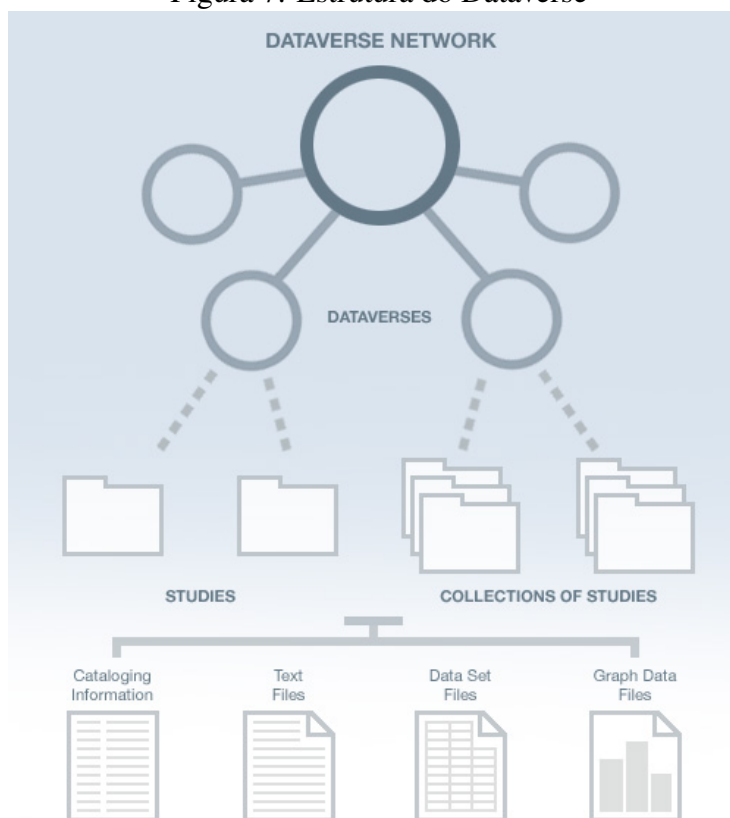
Por falta de heurísticas específicas para avaliação de repositórios de dados, foram usadas as dez heurísticas propostas por NIELSEN (1993) e atualizadas por ele em seu sítio⁶, mostradas na tabela 2, que são genéricas e totalmente aplicáveis ao *software* Dataverse. A escala de severidade, usada para classificar a gravidade dos problemas encontrados, mostrada na tabela 3, também foi a proposta por Nielsen⁷. As duas tabelas são apresentadas na seção 6.2.

Nielsen é reconhecido mundialmente por seu trabalho no campo da usabilidade, in-

⁶<https://www.nngroup.com/articles/ten-usability-heuristics>

⁷<https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems>

Figura 7: Estrutura do Dataverse



Fonte: Institute for Quantitative Social Science

clusivo eleito pelo *New York Times*⁸ como o guru da usabilidade de páginas *Web*. Junto com Don Norman, que cunhou o termo “*User eXperience*” (UX), fundaram uma empresa de consultoria e pesquisa sobre “Experiência do Usuário” e mantêm no sítio do *Nielsen Norman Group*⁹ versões atualizadas de vários artigos e resumos sobre seus trabalhos, inclusive um descritivo de como realizar uma avaliação heurística¹⁰.

Os avaliadores foram escolhidos por seu interesse e sua expertise no campo da usabilidade, interfaces e repositórios, sendo 3 da área de ciência da informação e 3 da área de tecnologia da informação. Desta forma, foi garantido o balanceamento entre o número de respostas de cada área.

- Profa. Dra. do Instituto de Ciências Humanas e da Informação (ICHI) da FURG, coordenadora do Portal de Periódicos eletrônicos da FURG¹¹ e do Repositório Institucional¹², com interesse em serviço de referência, gestão de bibliotecas, editoração eletrônica e repositório institucional;
- Profa. Dra. do ICHI da FURG, tem experiência na área de comunicação científica,

⁸<https://archive.nytimes.com/www.nytimes.com/library/tech/98/07/cyber/articles/13usability.html>

⁹<https://www.nngroup.com/>

¹⁰<https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>

¹¹<http://www.seer.furg.br>

¹²repositorio.furg.br

Tabela 2: Lista de heurísticas de Nielsen

1	Compatibilidade entre o sistema e o mundo real;
2	Consistência e uso de padrões;
3	Design estético e minimalista;
4	Reconhecimento ao invés de memorização;
5	Visibilidade do status do sistema;
6	Flexibilidade e eficiência de uso;
7	Controle e liberdade para o usuário;
8	Ajude os usuários a reconhecerem, diagnosticarem e recuperarem-se de erros;
9	Prevenção de erros;
10	Ajuda e documentação.

Fonte: NIELSEN (1993)

Tabela 3: Níveis de severidade

0	Sem importância
1	Cosmético
2	Simples
3	Grave
4	Catastrófico

Fonte: NIELSEN (1993)

processamento e busca de informações, estudos métricos em ciência da informação;

- Bibliotecária na FURG, Dra. em Educação em Ciências, com interesse em serviço de referência, gestão de bibliotecas, editoração eletrônica, repositório institucional, comunicação científica, métricas da informação, produção e avaliação da ciência;
- Analista de Tecnologia da Informação da FURG, Mestre em Modelagem Computacional, responsável pelo desenvolvimento e implantação do *framework* de desenvolvimento de *software* institucional da FURG e de vários sítios da Instituição, com larga experiência em Ciência da Computação e interesse em inteligência artificial, projeto de *software* e interfaces;
- Técnico de Tecnologia da Informação da FURG, integrante do Grupo de Trabalho de projetou e desenvolveu o primeiro sistema computacional de gestão de bibliotecas da Instituição, posteriormente tornando-se o seu coordenador responsável, com larga experiência em ciência da computação e interesse em projeto de *software*, interfaces e gestão eletrônica de documentos;
- Técnico de Tecnologia da Informação da FURG, integrante do Grupo de Trabalho que projetou e desenvolveu o primeiro sistema computacional de gestão de bibliotecas da Instituição, com larga experiência em Ciência da Computação e interesse

em projeto de *software*, interfaces e usabilidade.

Eles foram convidados a utilizar e avaliar um repositório de dados de pesquisa de acordo com a documentação constante no apêndice A, e a descrever os problemas de usabilidade encontrados, conflitantes com as heurísticas propostas. Também foi solicitado aos avaliadores que para cada problema encontrado fosse apresentada uma sugestão de correção.

A avaliação empírica, não importando a técnica utilizada, é importante para verificar os pontos onde os usuários comuns, não especialistas, enfrentam problemas ao utilizar uma interface. As falhas no *design*, os erros na interação e qualquer outro problema de usabilidade afetam principalmente aqueles usuários que não estão acostumados a interagir com aquela interface. Esse primeiro contato é importante para estabelecer a dificuldade de adaptação e aprendizado, isso é, o quão rápido e fácil esse novo usuário vai conseguir usar o sistema com a eficácia e eficiência necessária.

O experimento foi realizado segundo a metodologia proposta e documentada no apêndice B. Os usuários selecionados foram acompanhados individualmente na execução das tarefas propostas e incentivados pelos avaliadores a verbalizar suas opiniões, intenções e sugestões.

Não foi possível fazer um teste extenso no período compreendido por um mestrado, mas os resultados alcançados fornecem subsídios suficientes para um trabalho mais aprofundado no futuro. KRUG (2014, p. 112) afirma que testar um usuário é 100% melhor do que não testar nenhum e NIELSEN (1993) que, quanto mais usuários e avaliadores disponíveis, melhor. O mesmo vale para a variedade de testes aplicados. Portanto o objetivo proposto para este estudo foi alcançado e plenamente satisfeito com as avaliações efetuadas.

6 RESULTADOS

Este capítulo relata os principais resultados da pesquisa apresentada nesta dissertação. Inicialmente é apresentado o comparativo entre as soluções tecnológicas escolhidas de acordo com uma série de critérios estabelecidos. O processo de implantação e testes do protótipo são relatados. Por fim, os resultados da avaliação de usabilidade são analisados e discutidos.

6.1 Análise das soluções tecnológicas e prototipação do repositório

6.1.1 Comparativo das principais características desejadas

Para o estudo e análise das soluções tecnológicas foram levantados os principais requisitos desejados. Foram identificados os princípios FAIR, OAIS e os critérios do modelo de certificação *Core Trust Seal*. A Tabela 4 apresenta o comparativo destes requisitos feito para este estudo.

Já a Tabela 6 apresenta a relação completa de requisitos levantados pelo GT-RDP Brasil em seu relatório de Soluções Tecnológicas (ROCHA et al., 2018). Estes requisitos foram estruturados e categorizados com base no modelo OAIS, e organizados nas seguintes categorias:

- Representação do ambiente do repositório (AMB1-6);
- Representação dos conjuntos de dados (PAC1-6);
- Descrição e documentação dos conjuntos de dados (DOC1-11);
- Produção dos conjuntos de dados (SUB1-7);
- Armazenamento a longo prazo e planejamento da preservação (PD1-5);
- Acesso e uso dos conjuntos de dados (AC1-15).




A essas foi acrescentada mais uma categoria com características relativas ao uso, desenvolvimento e manutenção do *software* (SW1-6). Os indicadores de nível de atendimento dos requisitos é apresentada na Tabela 5.

Tabela 4: Comparativo dos principais requisitos desejados para o repositório nacional de dados de pesquisa.

	DSpace	Dataverse
Comunidades e grupos	Sim	Sim
Controle de usuário e permissão	Sim	Sim
Fluxo de submissão	Sim	Sim
Permite auto arquivamento	Sim	Sim
Aceita qualquer tipo de material digital	Sim	Sim
Customização da interface	Sim	Sim
Internacionalização (pt-br)	Sim	Sim
Otimizado para dispositivos móveis	Sim	Sim
Citação	Item	Sim
Embargo	Sim	Futuro(2020)
Versionamento	Sim	Sim
Versionamento de dados de pesquisa	Não	Sim
Permite esquemas de metadados prontos	Sim	Sim
Permite criar/alterar esquemas de metadados	Sim	Sim
Busca e interoperabilidade semântica	Sim	Sim
Protocolo SWORD	Sim	Sim
Protocolo OAI-PMH	Sim	Sim
Identificador Persistente (DOI/Handle)	Sim	Sim
Princípios FAIR	Sim	Sim
Princípios OAIS	Em parte	Em parte
Certificação <i>CoreTrustSeal</i>	Sim	Sim
Licença de código livre	BSD	Apache 2
Tecnologias utilizadas	Livres	Livres
Desempenho	Suficiente	Suficiente
Comunidade de usuários brasileiros	Grande	Pequena

Fonte: Autor

Tabela 5: Indicadores de nível de atendimento dos requisitos

Indicador	Cor
Plenamente ou Suficiente	
Suficiente com restrições	
Parcialmente ou muito parcialmente	
Não atende ou Insuficiente	

Fonte: ROCHA et al. (2018)

Os dois sistemas permitem publicação, compartilhamento e armazenamento de conteúdo digital e possuem recursos bem similares. Eles permitem uma estrutura hierárquica representativa de comunidades diversas, entidades autônomas com customizações próprias, chegando à personalização do acesso ao nível do documento ou arquivo digital individual, inclusive licenças de uso e identificadores únicos universais e persistentes: ambos aceitam DOI e *Handle System*.

Apesar do DSpace ser originalmente destinado à publicação de documentos e o Dataverse à disponibilização de dados de pesquisa, ambos permitem armazenar e gerenciar praticamente qualquer tipo de material digital. O PostgreSQL é a opção de código aberto para o Sistema de Gerenciamento de Banco de Dados em ambos os *softwares*. Ainda com relação ao armazenamento, tanto DSpace quanto Dataverse suportam a submissão por máquinas através do protocolo SWORD e a coleta de registros de metadados através do protocolo OAI-PMH.

As configurações de submissão também são similares, oferecendo desde autoarquivamento até níveis suplementares de conferência e aprovação, e permitindo que cada comunidade estabeleça suas regras e fluxos. Os sistemas de permissões são bastante flexíveis. Os usuários recebem autorizações que os habilitam a realizar certas funções dentro dos ambientes. Essas permissões são geralmente agrupadas em papéis (*roles*) que identificam o nível de autoridade do indivíduo ou grupo de usuários. É possível ainda, se assim desejado, liberar acesso de leitura e *download* para usuários sem cadastro (anônimo). Um recurso imprescindível ao projeto foi a integração com a comunidade acadêmica federada (CAFe) da RNP, usando as credenciais do usuário na instituição de origem. Essa importante funcionalidade é possível em ambos os sistemas de repositório através da configuração do sistema Shibboleth¹.

A interface gráfica é formada por temas configuráveis podendo ser diferentes para cada comunidade. Também é possível alterar a língua padrão de mensagens do sistema, sendo essa tarefa bem mais fácil na interface XMLUI do DSpace, na qual basta ajustar um parâmetro nos arquivos de configuração e o arquivo de tradução incluído na distribuição ficará disponível no ambiente *Web*. Já o Dataverse oferece um arquivo de mensagens que deve ser traduzido na língua desejada. Algumas traduções para determinados idiomas

¹<https://www.shibboleth.net>

Tabela 6: Comparativo completo dos requisitos levantados pelo GT-RDP Brasil

	Requisito	DSpace	Dataverse
AMB1	Representação do ambiente		
AMB2	Políticas de funcionamento do ambiente		
AMB3	Políticas descentralizadas		
AMB4	Sistemas de Informação de Pesquisa		
AMB5	Dados Abertos e Ligados		
AMB6	Transparência e feedback		
PAC1	Natureza dos conjuntos de dados		
PAC2	Representar os conjuntos de dados em pacotes		
PAC3	Formatos de arquivos aceitos		
PAC4	Versionamento de conjuntos de dados		
PAC5	Uso de padrões		
PAC6	Dados 5 estrelas		
DOC1	Informação de proveniência		
DOC2	Metadados administrativos		
DOC3	Metadados descritivos		
DOC4	Metadados estruturais		
DOC5	Metadados técnicos		
DOC6	Vocabulários controlados		
DOC7	Definir e estender esquemas de metadados		
DOC8	Recursos para realizar mapeamentos (crosswalks)		
DOC9	Descrever em Linked Data		
DOC10	Descrever integrado com Sist. de Informação da Pesquisa		
DOC11	Documentação dos dados		
SUB1	Gerenciamento de Acordos de Submissão e Licença		
SUB2	Fluxos de submissão		
SUB3	Transferência legal da custódia dos dados ao repositório		
SUB4	Validação e verificação de pacotes de submissão		
SUB5	Extração, verificação, produção e registro de metadados		
SUB6	Submissão em lote por máquina		
SUB7	Uso de padrões (para pacotes e protocolos de submissão)		
PD1	Serviços/microserviços para garantir acesso a longo prazo		
PD2	Planejamento e ações de preservação digital		
PD3	Integração com serviços de preservação digital		
PD4	Exclusão de dados		
PD5	Usado em Repositório Digital Confiável certificado		
AC1	Recuperação de informação		
AC2	Informações sobre direitos de uso, licenças		
AC3	Informações de proveniência e para uso dos dados		
AC4	Informações de citação		
AC5	Identificadores globais e persistentes		
AC6	Identificad. globais e persistentes para recursos relacionados		
AC7	Acesso aos (meta)dados via identificador e protocolo aberto		
AC8	Restrições de acesso, embargos		
AC9	Gerenciamento e autenticação de usuários		
AC10	Entrega de dados ao consumidor		
AC11	Entrega dos metadados ao consumidor		
AC12	Ferramentas para visualização e análise de dados		
AC13	Estatísticas e relatórios de uso		
AC14	Acesso às descrições em formatos para Linked Data		
AC15	Recuperação em ambiente Linked Data		
SW1	Tecnologia e Plataforma		
SW2	Distribuição e Versionamento		
SW3	Estratégia de Desenvolvimento do Software		
SW4	Licença de Uso		
SW5	Desempenho e Escalabilidade		
SW6	Presença – Usuários, Uso no Brasil		

Fonte: ROCHA et al. (2018)

são disponibilizadas pela comunidade de usuários, mas funcionam apenas em versões específicas do sistema.

Apesar de ambos gerarem citações, a citação de dados não é tão simples. Ela pode incluir, por exemplo, subconjuntos de observações e/ou variáveis. O Dspace não está preparado para isso até a sua versão atual (6.3). O mesmo se aplica no caso do versionamento. Já no quesito embargo é o contrário, o Dataverse não provê uma funcionalidade automatizada para isso, mas é uma meta estratégica a disponibilização desse recurso para o ano de 2020².

Os dois softwares suportam Dublin Core para descrição de metadados, sendo possível utilizar outros padrões e até estender padrões existentes. Apesar de ambos gerarem citações, estas podem incluir subconjuntos de observações e/ou variáveis específicas. O Dspace não oferece esse suporte até a versão analisada neste comparativo. Também não oferece versionamento automático dos conjuntos de dados publicados de forma incremental. No requisito embargo de dados é o contrário, o Dataverse ainda não provê uma funcionalidade automatizada, mas informa essa funcionalidade como meta estratégica para o ano de 2020.

Dataverse e DSpace suportam o sistema *Lots of Copies Keep Stuff Safe* (LOCKSS³) para distribuir cópias de conteúdo de forma compartilhada com outros repositórios, aumentando assim a segurança na preservação destes materiais. No quesito recuperação de informações, ambos utilizam o *software* Solr⁴, uma popular plataforma de código aberto de indexação e busca, que é baseada no Apache Lucene⁵.

6.1.2 Implantação do protótipo

Após a definição das ferramentas que se adequaram ao escopo do projeto do GT-RDP Brasil, DSpace e Dataverse, foi decidido que ambas as soluções deveriam ser implantadas e avaliadas quanto à usabilidade, visto que os recursos disponibilizados são muito parecidos e o atendimento aos critérios de avaliação descritos na seção anterior é equivalente. Portanto, esta seção descreve os procedimentos realizados para a implantação (instalação e configuração e testes) das soluções tecnológicas para repositórios de dados de pesquisa.

As soluções foram instaladas e configuradas em *virtual machines* – VMs distintas, usando o VirtualBox⁶ da Oracle. Esse ambiente foi escolhido pois, além de ser um dos oferecidos pela RNP, também era de conhecimento do autor e dos demais membros da equipe técnica do GT-RDP. O equipamento base utilizado foi um computador pessoal com processador Core I5 3.3 GHz de 4 núcleos, 12 GB de memória RAM, rodando sistema operacional Windows 7 64 bits. As máquinas virtuais foram criadas com um processador

²<https://www.iq.harvard.edu/roadmap-dataverse-project>

³<https://www.lockss.org>

⁴<http://lucene.apache.org/solr>

⁵<http://lucene.apache.org>

⁶<https://www.virtualbox.org>

e 4GB de memória cada, rodando sistema operacional Linux Ubuntu Server 16.04.4 LTS.

Apesar de exigirem conhecimentos técnicos especializados, os roteiros de instalação disponibilizados pelos desenvolvedores das soluções de repositório foram suficientes para instalar e configurar cada uma delas. Na VM destinada para o DSpace foram instalados e configurados os seguintes pacotes de *software*: Java 8, Maven 3.3.9, Ant 1.9.6, PostgreSQL 9.5, Apache Tomcat 7 e DSpace 6.2. Na segunda VM foram: Java 8, Glassfish 4.1, Postgres 9.5, Solr 4.6.0 e Dataverse 4.8.6.

Uma bateria de testes foi executada visando confirmar a análise dos requisitos levantados e dos recursos oferecidos. Por fim, decidiu-se pela adoção do software Dataverse como base do protótipo de repositório de dados de pesquisa com acesso aberto da RNP, por este atender a maior parte dos requisitos identificados no planejamento do repositório.

Estas VMs foram exportadas e replicadas na FABICO da UFRGS, para verificação dos testes pelos outros membros do GT. Após a validação dos roteiros de instalação e da adequação das funcionalidades às necessidades, foi solicitado à RNP que fornecesse os recursos necessários para a replicação em seu *data center*. O processo de instalação nas VMs da RNP foi parecido com o original, apenas se optou por instalar as últimas versões dos componentes que tinham atualizações disponibilizadas na época: Maven 3.5.4, Ant 1.10.5, DSpace 6.3, Solr 7.3.0 e Dataverse 4.9.1. Foi configurado o “Handle” como identificador persistente global, principalmente devido ao baixo custo quando comparado ao DOI.

Após a criação de usuários para os participantes do GT foram feitas submissões e pesquisas, testando várias configurações e ajustes para verificar os recursos disponibilizados pelos repositórios. Não foram feitas customizações para que a comparação de recursos e utilização fosse em relação à configuração básica de cada *software*. Para estes testes foram convidados pesquisadores e técnicos que submeteram dados da Rede Ipê, os quais apresentam medições de qualidade de 2012 a 2019 realizadas nos 27 Pontos de Presença (PoPs) da RNP destinados a rede acadêmica brasileira. Esta rede interliga 1522 campi e unidades de instituições de ensino, pesquisa e saúde em todo o país.

Esse envolvimento com usuários e suas pesquisas, fornecendo dados reais, serviu para efetivamente experimentar os recursos e capacidades das soluções tecnológicas testadas, e determinar a escolha do Dataverse como *software* base na implantação do repositório de dados de pesquisa pelo GT-RDP Brasil. A RNP e o IBICT apoiaram a decisão e já começaram a incentivar a sua disseminação como padrão nacional.

Atualmente, o repositório está implantado em versão beta⁷. Foram prototipados Dataverses para armazenar dados de pesquisa produzidos pelo próprio GT-RDP durante a execução do projeto, dados da Rede Ipê da RNP e dados da comunidade de pesquisadores em Redes de Computadores. Atualmente, estão sendo implementados pilotos de

⁷<https://dadosabertos.rnp.br>

repositórios de dados de pesquisa para as seguintes organizações: a RedeRio⁸, rede que interconecta instituições acadêmicas, centros de ensino e pesquisa e órgãos públicos fluminenses; e a Fundação Oswaldo Cruz (FIOCRUZ⁹), instituição de ciência e tecnologia em saúde, vinculada ao Ministério da Saúde.

6.2 Análise dos resultados da avaliação de usabilidade

A metodologia de avaliação da usabilidade proposta foi aplicada a uma instância de repositório de dados de pesquisa baseado na solução tecnológica Dataverse. Como alvo da avaliação foi escolhido o repositório de dados de pesquisa do IBICT¹⁰, que junto da RNP coordena as atividades do GT-RDP. Esta escolha deve-se ao fato deste repositório já estar mais maduro do que o recém implantado, contendo mais *dataverses* e *datasets*, além de já estar traduzido para a língua portuguesa.

NIELSEN (1993) alerta que “usabilidade aplica-se a todos os aspectos de um sistema com o qual um humano pode interagir, incluindo procedimentos de instalação e manutenção”. Isso ficou evidente em relação aos produtos instalados e configurados. Tanto DSpace quanto Dataverse têm manuais com um detalhamento suficiente para que técnicos experientes consigam segui-los, mas confusos o suficiente para que leigos falhem na sua instalação e configuração.

Apesar do sítio do Dataverse apresentar vários relatórios sobre a usabilidade do sistema, não fomos capazes de encontrar nenhum trabalho científico independente que avalie o avalie sob este aspecto. Portanto, avaliar um repositório de dados de pesquisa baseado numa versão recente do Dataverse contribui de forma significativa para a área de pesquisa desta dissertação.

Avaliar um produto que se pretende implantar e disseminar a toda uma comunidade acadêmica é, segundo CORRÊA; MIRANDA (2012), assumir o risco da descoberta, em maior ou menor grau, de possíveis problemas. Como o Dataverse é um produto de prateleira, já em estágio de produção, os resultados de avaliações de usabilidade não se aplicam para fins de alteração do projeto e concepção. Entretanto, esses resultados podem indicar pontos de melhoria da usabilidade para versões futuras, inclusive no que tange a instalação e configuração do produto, item que é uma reclamação constantemente da comunidade de usuários. No caso de *software* com código-fonte aberto, tais como o Dataverse, essas sugestões de melhorias e correção de problemas são realizadas de forma simples e corriqueira, mas não há garantia de prazo e nem mesmo da aceitação.

⁸<http://www.rederio.br>

⁹<https://portal.fiocruz.br>

¹⁰<https://repositoriopesquisas.ibict.br/>

6.2.1 Avaliação heurística

A avaliação de usabilidade através do método de inspeção heurística mostrou-se bastante eficaz. Muitos dos problemas encontrados foram percebidos por mais de um especialista e também por alguns usuários no teste empírico.

Esperava-se que os especialistas da área de Ciência da Informação seriam os que encontrassem mais problemas, já que tinham mais experiência no uso de repositórios. Entretanto, foi verificado exatamente o contrário. Dois deles foram os que relataram pouquíssimos problemas. Muito provavelmente este comportamento seja reflexo da grande familiaridade com esse tipo de sistema, ou seja, a usabilidade do repositório Dataverse era quase tão natural quanto os repositório que já estavam acostumados.

As Tabelas 7, 8 e 9 apresentam o resumo dos problemas encontrados, a heurística afetada e a sugestão de correção, considerando os níveis mais severos. A relação de heurísticas referenciadas foi apresentada na Tabela 2 e os níveis de severidade na Tabela 3.

Tabela 7: Lista de problemas e heurísticas afetadas - Catastrófico

Problema encontrado	Heurística afetada	Solução sugerida
Os guias disponíveis são muito técnicos e não estão disponíveis em português	1, 10	Disponibilizar guias práticos de uso em português
Faltam documentações sobre a arquitetura do Dataverse, de como estão estruturadas as comunidades e as políticas de uso do repositório	10	Disponibilizar a documentação apropriada

Fonte: Autor

O que mais incomodou os avaliadores foi a falta de guias e documentação, sendo julgado como catastrófico pois afeta toda a utilização do sistema. A disponibilização de documentação é imprescindível e precisa ser clara, prática e estar em português. Outro item considerado catastrófico é a falta de documentação descrevendo a estrutura do Dataverse e como ele está organizado em suas (sub)comunidades (dataverses). Isso não seria tão importante mas o próprio sistema enfatiza essa estrutura no resultado das buscas.

Como problemas considerados graves o destaque é a falta de sugestões e de correção de termos. Uma falha bastante citada é não reconhecer sinônimos e em casos extremos só as preposições digitadas foram encontradas, o que casou a recuperação de quase toda a base de dados, mas totalmente irrelevante. A Figura 8 demonstra essa situação. A busca pelos termos “peixes de aquário”, restrita a somente arquivos de dados, retornou 190 resultados, mas todos irrelevantes pois o único termo encontrado foi “de”.

Tabela 8: Lista de problemas e heurísticas afetadas - Grave

Problema encontrado	Heurística afetada	Solução sugerida
A busca avançada contém termos técnicos que talvez só os produtores dos dados conheçam, identifiquem e utilizem	1, 3	Alguns campos poderiam ser retirados, o desenvolvedor poderia gerar uma estatística de quais campos não são usados, ou raramente usados
A pesquisa retorna arquivos, <i>datasets</i> e <i>dataverses</i> , tudo misturado e confuso, pesquisar arquivos de dados e pesquisar comunidades são coisas diferentes para o usuário final	2	Separar as buscas ou tornar mais claras as opções do lado esquerdo da tela que já estão marcadas por padrão
Não há sugestão nem correção de termos nas buscas, assim como sinônimos não são encontrados, nos casos extremos só encontra preposições e outros termos irrelevantes	1, 4, 8	Oferecer sugestão e correção de termos, pesquisar sinônimos e excluir termos irrelevantes
O sistema não informa de maneira destacada a pesquisa que efetivamente foi feita, de forma que se o usuário errar na digitação não verá explicitamente o que digitou errado	5	Mostrar a pesquisa feita destacando termos e operadores
O botão citação está mal posicionado, faltou a informação de que a citação é do conjunto de dados e não de um arquivo lá disponível e link de como citar um arquivo ou parte dele direciona para um sítio externo em inglês	8, 10	Destacar o botão, informar sua condição e regras

Fonte: Autor

Na categoria das severidades simples foi relatado que o posicionamento do campo busca estava inadequado, como lembrado por um avaliador: - Ele é o motivo dos usuários acessarem o sistema e deveria ter um local de destaque!

Um item que foi notado por alguns avaliadores é a falta de tradução em partes do sítio. Atualmente, alguns termos em inglês são tão comuns que passam despercebidos e são aceitos naturalmente pela maior parte dos usuários. Mas no caso da busca avançada e de algumas mensagens de erro isso chamou bastante a atenção pelo uso de termos da área de preservação digital. A Figura 9 mostra um campo da busca avançada que não teve o texto de ajuda traduzido e apresenta jargões técnicos que segundo os especialistas não

Figura 8: Tela com resultado da busca por “peixes de aquário”

1 a 10 de 190 Resultados ↑↓ Ordenar ▾

- RELATÓRIO DE TRANSIÇÃO_2018_final.pdf**
 Mar 20, 2019 - Relatório de Transição
 Adobe PDF - 20.0 MB - MD5: f7058dc939113453b37abd6ca6d9ea0e
 Filename Without Extension: RELATÓRIO DE TRANSIÇÃO_2018_final
- UM OLHAR CONSTRUTIVISTA DO PROCESSO DE BUSCA E USO.pdf**
 Oct 18, 2018 - Um olhar construtivista do processo de busca e uso da informação: a aquisição de competência em informação
 Adobe PDF - 56.8 KB - MD5: a686f732a2a7df77f1829747ba6cac59
 Filename Without Extension: UM OLHAR CONSTRUTIVISTA DO PROCESSO DE BUSCA E USO
- Relatorio de Atividade 2017.docx**
 Sep 28, 2018 - Relatório de Atividades Anual – Exercício 2017.
 MS Word (docx) - 7.3 MB - MD5: 0367abf2d3900f0ad4b67c7ce25c3951
 Filename Without Extension: Relatorio de Atividade 2017

Fonte: Autor

tem interesse por parte da maior parte do público do sistema.

Figura 9: Tela com ajuda de campo da busca avançada

Conjunto de dados: Citation Metadata ^

Title

Author Name

Author Affiliation

Description Text

Subject Agricultural Sciences
 Arts and Humanities
 ...
 ...

Keyword Term

Topic Classification Term

Related Publication Citation

Keyword Term
 Key terms that describe important aspects of the Dataset. Can be used for building keyword indexes and for classification and retrieval purposes. A controlled vocabulary can be employed. The vocab attribute is provided for specification of the controlled vocabulary in use, such as LCSH, MeSH, or others. The vocabURI attribute specifies the location for the full controlled vocabulary.

Fonte: Autor

Um modo de customizar e obter maior precisão nos resultados é utilizar operadores

Tabela 9: Lista de problemas e heurísticas afetadas - Simples

Problema encontrado	Heurística afetada	Solução sugerida
O termo Dataverse não representa nada para o usuário comum, deveria ser trocado para outro, como Comunidade por exemplo, o termo Datasets também	1	Trocar alguns termos não comuns por outros mais naturais
O campo de busca está muito escondido	3, 4	Reposicionar campo de busca dando destaque maior
As dicas e textos de ajuda na busca avançada são insuficientes	4, 10	Melhorar dicas e textos de ajuda
Vários textos como na busca avançada e mensagens de erro sem tradução, além de pouco esclarecedoras	1, 2, 8	Melhorar e traduzir mensagens de erro e outros textos
Faltam informações em alguns momentos, por exemplo o que significa o cadeado em alguns arquivos, simbologia de alguns ícones	1, 2, 5	Tornar mais consistentes, amigáveis e naturais os ícones, explicar o que significam ou representam
É possível utilizar operadores lógicos, mas não é dito isso nem explicado seu uso	6	O uso de operadores lógicos torna a busca muito mais precisa, informar que é possível usar e como

Fonte: Autor

lógicos na pesquisa, combinando o uso deles com parênteses e a especificação dos qualificadores de campo aumenta a chance de obtermos os dados de interesse, sem recuperar itens irrelevantes. Como não é comum o uso de tal ferramenta, ele foi considerado com severidade simples pois não afeta tantos usuários, e os mais experientes poderão aprender e se adaptar ao seu uso, mesmo sem documentação disponível no sistema. Ao realizar uma busca avançada, o padrão para o uso de operadores lógicos é a aplicação do operador “AND” entre as opções de um mesmo quadro e “OR” quando a busca utiliza campos de quadros diferentes, como exemplificado abaixo:

```
(dvName:Q1-a AND dvAffiliation:Q1-b) OR
(title:Q2-a AND authorName:Q2-b) OR
(fileName:Q3-a AND fileDescription:Q3-b)
```

onde:

- Q1-a e Q1-b pertencem ao primeiro quadro: “Dataverses”;

- Q2-a e Q2-b pertencem ao segundo quadro: “Datasets”;
- Q3-a e Q3-b pertencem ao terceiro quadro: “Files”.

Como ressaltado pelos avaliadores a simples documentação da possibilidade de uso, e as regras básicas, possibilitaria uma melhor resposta na recuperação da informação.

As Tabelas 10 e 11 apresentam os problemas encontrados pelos especialistas considerando os níveis menos severos. No final foi solicitado aos especialistas responder sobre seu nível de satisfação ao usar o repositório. As perguntas feitas foram: “Qual seu nível de (dificuldade inicial / dificuldade geral / satisfação) ao usar o repositório Dataverse do IBICT?”. A Tabela 12 resume essas respostas, cujas opções compreendem escalas Likert de 5 pontos (LIKERT, 1932).

Tabela 10: Lista de problemas e heurísticas afetadas - Cosmético

Problema encontrado	Heurística afetada	Solução sugerida
As imagens com logos dos <i>dataverses</i> tem destaque muito grande	3	Poderiam estar só na página principal do sistema
Falta um botão voltar e uma maneira de como voltar à pesquisa feita anteriormente	4, 6, 7	Disponibilizar um botão voltar e um re-fazer pesquisa, que mantenha os dados informados anteriormente
Os ícones que representam o e-mail para contato e compartilhamento não são tão relevantes para estarem no alto da página, também não está explícito se são relacionados à página ou ao (qual) <i>dataverse</i>	2, 3	Posicionar os ícones ao lado do logo do <i>dataverse</i>

Fonte: Autor

As disparidades nas respostas é explicada por causa do nível de conhecimento de cada avaliador em relação a repositórios digitais. Enquanto para os usuários acostumados com esse tipo de ferramenta a experiência foi natural, para os demais as dificuldades foram maiores. Quanto à insatisfação, ela também foi proporcional à experiência anterior de cada especialista, onde a comparação com ambientes conhecidos é inevitável. Além disso, sempre é levado em conta as expectativas alcançadas por cada um.

6.2.2 Avaliação empírica

O avaliação empírica foi realizada com um grupo de jovens pesquisadores, estudantes de mestrado do Programa de Pós-Graduação em Computação da FURG, os quais já tinham certa experiência em pesquisar publicações científicas na *Web*. Na primeira parte

Tabela 11: Lista de problemas e heurísticas afetadas - Sem importância

Problema encontrado	Heurística afetada	Solução sugerida
O quadro “Metrics” está muito destacado já que não contém informações relevantes	3	Torná-lo clicável com informações relevantes e reposicioná-lo com um destaque menor
Telas muito poluídas	3	Tela principal com informações de apresentação e busca, nas seguintes só com informação relevante
Falta de botão imprimir que ofereça versão mais limpa da página	7	Oferecer uma versão mais limpa para impressão

Fonte: Autor

Tabela 12: Notas das questões gerais sobre a satisfação dos especialistas

Questão	Esp1	Esp2	Esp3	Esp4	Esp5	Esp6
Dificuldade inicial	Indiferente	Muito fácil	Fácil	Fácil	Difícil	Fácil
Dificuldade geral	Fácil	Muito fácil	Indiferente	Indiferente	Difícil	Fácil
Satisfação	Insatisfeito	Muito satisfeito	Insatisfeito	Satisfeito	Insatisfeito	Satisfeito

Fonte: Autor

do experimento foi solicitado que os participantes que acessassem o repositório de dados de pesquisa do IBICT e realizassem 5 tarefas básicas e recorrentes de um usuário. Essas tarefas exigiam realizar buscas no repositório, selecionar o conjunto de dados correto, baixar o arquivo solicitado e finalmente localizar e baixar o arquivo de citação necessário para referenciá-lo.

É importante ressaltar que toda pesquisa envolvendo opiniões de usuários está sujeita à subjetividade destas opiniões. Cada pessoa tem suas próprias necessidades e capacidades ao usar um sistema, que podem ser maiores ou menores que seus colegas. Portanto, é natural a alta diversidade de opiniões e classificações de severidade que são encontradas em testes desse tipo. Com isso, diversas características da interface avaliada são constantemente julgadas boas ou ruins, dependendo de diversos fatores ligados ao usuário, principalmente as ligadas aos seus conhecimentos e experiências anteriores.

Os problemas encontrados foram identificados e separados por suas características e novamente reunidos resultando os itens relacionados abaixo:

- Faltou informação de uso de operadores lógicos na busca;
- Faltou destaque na página principal ao campo de busca;
- Palavras erradas não têm sugestões e sinônimos não são encontrados;
- Botão “encontrar” na busca avançada mal posicionado;
- A busca avançada é muito “avançada”: extensa e confusa;
- Excesso de campos e falta de outros na busca avançada;
- A mistura entre *dataverses*, conjuntos de dados e arquivos torna o resultado da busca confuso. Os ícones diferentes e conflitantes entre o lado esquerdo e direito de cada item confunde mais ainda;
- A mistura de resultados da busca, levando em conta as palavras encontradas com os operadores “e” e “ou” torna tudo confuso;
- No resultado da busca alguns arquivos retornados possuem um ícone de “cadeado” e não podem ser baixados, mas nenhuma informação textual avisa o que está acontecendo, estão presentes checkbox e textos clicáveis mas sem ações correspondentes;
- Na tela de resultados não há botão “voltar” ou “nova pesquisa”;
- No resultado da busca faltam informações, principalmente sobre os arquivos presentes no conjunto de dados (quantidade/tipo);
- Só foram oferecidas opções de download dos dados no formato original, sem opção de tipos diferentes (xls/csv);
- A listagem de resultados da busca está imprecisa;
- Os arquivos retornados acabam sendo linkados ao seu respectivo conjunto de dados, misturado aos demais arquivos, sem destaque ao próprio;
- A posição do botão “Cite os dados” não é apropriada;
- O botão “Cite os dados” não é natural, tanto o texto quanto o formato;
- O número de citações não foi encontrado;
- Não está explicado que a citação é do conjunto de dados e não do arquivo de dados;
- Falta de explicação sobre a estrutura organizacional do “Dataverse” e suas comunidades;
- Falta de tutoriais simples, os guias oferecidos são muito extensos e complexos;

- Alguns textos e ajudas não traduzidos;
- O quadro de métricas tem muito destaque mas não tem informações relevantes e não é clicável.

A busca é o item principal do repositório no que tange ao usuário final, é o componente primariamente procurado e mais utilizado por quem necessita recuperar qualquer informação. É claro que a ferramenta é bastante completa em recursos de submissão, armazenamento, preservação e compartilhamento, mas ela peca um pouco na recuperação de informações.

A grande maioria dos problemas relatados, 14 dos 22, foram referentes à busca e aos resultados retornados. Os principais tratam da falta de sugestões e correções ortográficas das palavras buscadas. Neste quesito o Google acostumou os usuários a digitarem de qualquer maneira esperando a sugestão da correção ou o complemento da consulta. O segundo problema mais relatado foi quanto ao resultado da busca, a ordenação não era a esperada e alguns sugeriram a implementação de um algoritmo de busca mais eficiente.

Alguns participantes relataram que o próprio campo de busca não tem o destaque que deveria e merecia ter. Faltaram informações de como a busca funciona e se é possível utilizar operadores lógicos para refinar a pesquisa.

Em alguns momentos os usuários pareceram insatisfeitos com o resultado das busca, principalmente se elas não retornavam informação pertinente. Alguns tentaram a busca avançada mas a acharam muito extensa e confusa. Ainda relativo à busca avançada foi notado que não havia um campo específico para buscar por tipo de arquivo e os botões de “Encontrar” não estavam bem posicionados.

Quando a consulta é retornada aparecem a maior parte dos problemas. Oito deles são diretamente relacionados a como os resultados são apresentados. O principal, já mencionado anteriormente, foi citado por 7 participantes. A maneira como o resultado é apresentado listando junto *dataverses*, conjuntos de dados e arquivos confunde e toma um tempo para ser entendido, principalmente com a confusão gerada pelos ícones apresentados à direita e esquerda de cada item. Os arquivos individuais, quando listados, ao serem clicados acabam sendo direcionados ao conjunto de dados a que pertencem, sem nenhum destaque ao próprio arquivo.

Outro problema relatado é a falta de informações na listagem de resultados, principalmente no caso de conjuntos de dados, sobre a quantidade e os tipos de arquivos disponíveis. A sugestão dada foi colocar uma caixa de combinação (*combobox*) com os arquivos disponibilizados. Sobre os tipos de arquivos também foi sugerida a possibilidade conversão automática de arquivos, por exemplo, baixar um arquivo Excel como csv.

Outro problema que causou certa confusão foi quando o resultado traz arquivos não disponíveis para *download*. Falta informar melhor a situação, se precisa estar logado na plataforma ou solicitar liberação dos dados. Além disso, é aconselhado desabilitar as

caixas de seleção a esquerda de cada item e tornar o item não clicável. A Figura 10 mostra essa situação.

Figura 10: Tela de resultados com arquivo indisponível



Fonte: Autor

Uma última sugestão em relação à apresentação de resultados é a inclusão de um botão de nova pesquisa na página do conjunto de dados.

Quando foi solicitado baixar o arquivo de citações os problemas mais citados foram em relação ao próprio botão de *download*. Metade destas reclamações foram quanto a posição e a outra metade sobre o próprio botão. As sugestões foram desde mudar a posição dele para próximo ao botão de *download* do conjunto de dados até mudar o texto "cite os dados" e também transformá-lo em um subconjunto de botões de acordo com o tipo de citação desejado. Também foi sugerido que seja disponibilizada uma maneira de informar a quantidade e quem citou aqueles dados.

Alguns usuários reportaram a falta de tutoriais simples de uso e uma explicação básica da estrutura do repositório. Talvez a presença deles ajudasse a reduzir o tempo gasto e a insatisfação com os resultados obtidos. A estrutura hierárquica de comunidades (*data-verses*) presentes na maioria das soluções para repositórios, apesar de lógica, precisa ser melhor esclarecida e detalhada ao usuário final.

Um problema percebido por dois usuários foi a falta de tradução para o português de algumas parte do sistema, principalmente a busca avançada e o texto de ajuda no campo de pesquisa da página do conjunto de dados, como mostrado no Figura 11. Além da sugestão de tradução de "*download*" para "baixar", por exemplo.

Figura 11: Tela de busca no conjunto de dados com ajuda não traduzida



Fonte: Autor

O quadro com as “Métricas” também foi criticado, além de ter um destaque injustificado pois não contém informações tão importantes, somente o número de “downloads” feitos, ele é muito chamativo e ocupa espaço útil na página. A sugestão foi de torná-lo mais simples e transformá-lo num botão que forneça mais informações, e mais úteis.

Após o teste empírico ser aplicado, o participante era convidado a preencher um formulário com um conjunto de 10 questões relativas a sua satisfação quanto ao uso do sistema. Na Tabela 13 são mostradas as notas dadas e a média delas para cada questão. As notas atribuídas podiam variar numa escala Likert de 5 pontos, representadas por ícones com faces representativas da satisfação do participante. A correspondência numérica foi uma escala onde 1 representa insatisfação completa e 5 satisfação completa.

Tabela 13: Notas do questionário de satisfação dos usuários

Questão	Participantes											Satisfação	
	1	2	3	4	5	6	7	8	9	10	11	Média	(%)
1.Cores, ícones e símbolos	5	5	5	4	5	5	5	5	5	5	4	4.818	96,36
2.Tipo e o tamanho das fontes	5	5	4	4	5	5	5	4	3	5	4	4.455	89,09
3.Posicionamento dos botões	4	3	3	3	5	5	4	4	2	5	1	3.545	70,91
4.Posicionamento dos menus	5	4	4	4	5	5	5	5	5	5	1	4.364	87,27
5.Funcionalidade dos botões	5	5	4	5	5	5	5	5	5	5	3	4.727	94,55
6.Funcionalidade dos menus	5	5	4	4	5	5	5	3	5	3	5	4.455	89,09
7.Quantidade e qualidade das informações	2	5	4	3	4	5	5	3	3	1	3	3.455	69,09
8.Organização e posicionamento das informações	4	4	3	4	4	5	5	5	3	5	5	4.273	85,45
9.Primeiras impressões	5	4	4	4	5	5	5	4	3	4	3	4.182	83,64
10.Facilidade de aprender	5	5	4	5	5	5	5	4	4	5	3	4.545	90,91

Fonte: Autor

De acordo com as notas médias os itens que mais agradaram aos usuários foram “As cores, ícones e símbolos usados” e “Funcionalidade dos botões”. O índice de aprovação desses quesitos foi de 96,36 e 94,55%, respectivamente, evidenciando a aprovação dos participantes, pelo menos em alguns quesitos da avaliação da interface. Como verificado pelos outros índices, somente dois deles ficaram em torno de 70% e todos outros ficaram acima de 80%.

Contrastando com a aprovação da funcionalidade dos botões está a insatisfação com o posicionamento dos mesmos. Os usuários reclamaram dos botões de citação, pesquisa e *download*. Outro ponto negativo foi a quantidade e qualidade das informações, provavelmente reflexo das reclamações sobre falta de tutoriais e de maiores informações nas listagens de resultados das buscas.

Dois pontos que chamam a atenção são sobre as impressões gerais de uso do sistema, o efeito natural, causado pela primeira impressão de uso, não foi tão baixo quanto seria de se esperar, isso para um sistema em que os participantes não estavam acostumados a usar.

O outro ponto de destaque foi a facilidade de aprender, representando 90,9% de satisfação.

6.2.3 Síntese das sugestões de melhoria

A análise dos resultados dos dois experimentos mostra que a ferramenta possui atributos que satisfazem a maior parte dos usuários, mas que peca em alguns aspectos. As sugestões de correção e melhoria dadas pelos participantes é resumida na listagem abaixo:

- Disponibilizar tutorial de uso, pois os guias oferecidos são muito extensos e complexos;
- Disponibilizar documentação da estrutura organizacional do Dataverse e suas comunidades;
- Completar a tradução da ferramenta;
- Melhorar dicas e textos de ajuda;
- Alterar termos técnicos e não usuais por outros mais comuns e naturais;
- Reduzir o destaque do quadro de métricas e que ele ofereça informações com mais qualidade e quantidade;
- Telas estão muito poluídas, somente manter informações essenciais a cada situação;
- Oferecer versão mais limpa da tela para impressão;
- Reposicionar ícones de e-mail e compartilhamento, assim como reposicionar e diminuir o destaque dos logos dos Dataverses, poderiam estar só na página principal;
- Alterar a posição e dar mais destaque ao campo de busca;
- Sugerir correção de palavras no campo de busca, retornar sinônimos e não considerar termos irrelevantes;
- Mostrar a pesquisa feita anteriormente destacando termos e operadores;

- Rever método de busca para separar melhor *dataverses*, conjuntos de dados e arquivos de dados;
- Rever método de busca para pontuar melhor os resultados, principalmente aqueles contendo os termos com operadores “E” e “OU”;
- Informar da possibilidade de uso de operadores lógicos e sua sintaxe;
- Na busca avançada é preciso rever as opções oferecidas e retirar as com pouca probabilidade de uso;
- Reposicionar o botão “Encontrar” na busca avançada;
- Informar o estado e situação de cada arquivo e conjunto de dados, destacando se pode ser baixado, data de liberação ou qual a maneira de requisitá-lo;
- Tornar mais consistentes, amigáveis e naturais os ícones, explicar o que significam ou representam;
- Quando a busca retornar um conjunto de dados, informar detalhes dos arquivos contidos nele;
- Quando clicar em um arquivo de dados, e abrir a tela com o conjunto de dados a que ele pertence, destacar o arquivo procurado;
- Oferecer opções de formatos diferentes do arquivo, principalmente “CSV”;
- Disponibilizar botões “Voltar” e “Nova/Refazer pesquisa”;
- Alterar a posição do botão “Cite os dados” e a sua forma, traduzir e tornar mais prático o link externo “Padrões de citação de dados”, tentar explicitar o número de citações já feitas;
- Explicar que a citação gerada é para o conjunto de dados e não para o arquivo individual, oferecer alternativas de citação.

7 CONCLUSÃO

Neste capítulo são apresentadas as considerações finais. Os resultados alcançados e principais contribuições desta dissertação são sumarizados e associados à produção bibliográfica resultante. Por fim, são discutidos alguns trabalhos futuros.

Esta dissertação apresentou o processo de implantação do piloto de um serviço nacional para repositório de dados de pesquisa em acesso aberto. Dentre as principais contribuições da pesquisa realizada durante o mestrado, destacam-se:

1. Uma análise de diferentes soluções tecnológicas para repositórios de dados, sendo o primeiro trabalho da literatura a comparar profundamente Dataverse e DSpace;

Como resultado dessa análise foi escolhido e indicado para a RNP o *software* Dataverse. Esta ferramenta, no futuro, deverá ser implementada no ambiente da RNP e disponibilizada como um serviço permanente para toda a comunidade científica brasileira.

2. A implantação de um protótipo de repositório de dados de pesquisa com acesso aberto para a comunidade científica brasileira;

Com a maturidade do protótipo e dos pilotos implantados, será fornecido um pacote para instalação local nas instituições de ensino e/ou pesquisa, nos moldes de como foi feito com o DSpace para publicações científicas (artigos em conferências e revistas, teses e dissertações).

3. A avaliação da usabilidade de um repositório de dados de pesquisa baseado na mesma solução implantada. Ressalta-se que não fomos capazes de encontrar nenhum trabalho com qualquer análise deste tipo de repositório.

Como resultado dessas avaliações foram apresentadas sugestões de correções e melhorias na interface do repositório de dados de pesquisa do IBICT e do *software* Dataverse. Estes resultados estão sendo preparados para submissão de um artigo nos próximos meses.

Após pesquisar sobre as técnicas mais conhecidas e suas variações, os métodos de avaliação da usabilidade escolhidos foram suficientes para alcançar os objetivos propostos. Apesar de outras técnicas aparentarem ser mais completas - e cada pesquisador tem suas favoritas - as escolhidas são plenamente aplicáveis num ambiente acadêmico e até mesmo em pequenas empresas, locais sem muitos recursos que possam ser disponibilizados para contratar consultorias especializadas.

O que fica bem claro com este e todos os trabalhos sobre usabilidade pesquisados é que os desenvolvedores precisam ter uma preocupação adicional com a usabilidade de suas interfaces. Para isso não é necessário dispendir grandes recursos, com pouco esforço é possível fazer um trabalho satisfatório que é indispensável ao sucesso do seu empreendimento.

É importante salientar que a cada versão de *software* lançada que altere a interface de um sistema, um novo teste de usabilidade deverá ser realizado. Isso significa que o trabalho é contínuo e por tempo indeterminado, justificando o investimento em treinamento e outros recursos para sistematizar esse tipo de empreendimento. O treinamento é importante não só para para os envolvidos nas inspeções de usabilidade. Ele deve ser ministrado aos desenvolvedores diretamente ligados ao projeto das interfaces. Um bom projeto elimina diversos problemas de usabilidade e o custo para isso é muito menor que o de consertá-los no futuro.

Entre as possibilidades de trabalhos futuros a partir da pesquisa realizada nesta dissertação destaca-se a avaliação da usabilidade do agregador, que irá integrar a pesquisa e a recuperação em todos os repositórios de dados de pesquisa das instituições brasileiras. Outra sugestão é a montagem de um laboratório para testes de usabilidade, com sala de teste e sala de controle e observação, equipada com câmeras, microfones, sensores, equipamentos e *softwares* de registro. Esse local e pessoas treinadas para seu uso poderiam disponibilizar seus recursos para testes tanto acadêmicos quanto para o setor de desenvolvimento da instituição. Outra possibilidade é um estudo propondo heurísticas específicas para repositórios, tanto institucionais quanto de dados de pesquisa.

Por fim, espera-se que este estudo tenha auxiliado os trabalhos do GT-RDP Brasil e aqueles que contribuem com o desenvolvimento do Dataverse e outros repositórios. Espera-se também que incentive outros trabalhos na área já que a usabilidade de uma interface está diretamente ligada à satisfação do usuário, que por sua vez contribui em grande parte com o sucesso de um sistema.

REFERÊNCIAS

- AMORIM, R. C.; CASTRO, J. A.; SILVA, J. R. d.; RIBEIRO, C. A Comparative study of platforms for research data management: interoperability, metadata capabilities and integration potential. In: ROCHA, A.; CORREIA, A. M.; COSTANZO, S.; REIS, L. P. (Ed.). **New contributions in information systems and technologies**. Cham, Suíça: Springer International Publishing, 2015. p.101–111. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-319-16486-1_10>. Acesso em: 17 nov. 2017.
- AMORIM, R. C.; CASTRO, J. A.; SILVA, J. R. d.; RIBEIRO, C. **A comparison of research data management platforms: architecture, flexible metadata and interoperability**. **Universal Access in the Information Society**, Cham, Suíça, v.16, n.4, p.851–862, jun 2017. Disponível em: <<https://link.springer.com/content/pdf/10.1007%2Fs10209-016-0475-y.pdf>>. Acesso em: 17 nov. 2017.
- Associação Brasileira de Normas Técnicas. **ABNT NBR ISO 9241-11:2011** requisitos ergonômicos para o trabalho com dispositivos de interação visual - parte 11: orientações sobre usabilidade. Rio de Janeiro: ABNT, 2011. Disponível em: <<https://www.abntcatalogo.com.br/norma.aspx?ID=86090>>. Acesso em: 16 mai. 2019.
- BONSIEPE, G. **Design do material ao digital**. Florianópolis: CTAI/LBDI Laboratório Brasileiro de Design, 1997. Disponível em: <https://monoskop.org/images/e/e9/Bonsiepe_Gui_Design_do_material_ao_digital.pdf>. Acesso em: 18 jun. 2019.
- CAFÉ, L. C.; MUÑOZ, I. K. **Avaliação de usabilidade no repositório institucional da universidade de Brasília**. **Informação & Tecnologia**, Salvador, BA, v.3, n.2, p.39–61, 2016. Disponível em: <<http://www.periodicos.ufpb.br/index.php/itec/article/view/38431>>. Acesso em: 02 mai. 2019.
- CCSDS, The Consultative Committee for Space Data Systems. **Reference Model for an Open Archival Information System (OAIS)**. CCSDS, Washington, jun 2012. Disponível em: <<https://public.ccsds.org/pubs/650x0m2.pdf>>. Acesso em: 29 out. 2018.
- CONARQ, Câmara Técnica de Documentos Eletrônicos - Conselho Nacional de Arquivos. **Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis - RDC-Arq**. Conarq, Rio de Janeiro, 2015. Disponível

em: <http://www.conarq.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf>. Acesso em: 29 out. 2018.

CORRÊA, T. P. P.; MIRANDA, A. C. D. **Usabilidade da seção de avaliação do Sistema Eletrônico de Editoração de Revistas através da opinião dos avaliadores do portal de periódicos científicos da FURG.** *Revista ACB*, Florianópolis, SC, v.17, n.1, p.210–226, 2012. Disponível em: <<https://revista.acbsc.org.br/racb/article/view/823>>. Acesso em: 04 jun. 2019.

COSTA, S. M. S. **Filosofia aberta, modelos de negócios e agências de fomento:** elementos essenciais a uma discussão sobre o acesso aberto à informação científica. *Revista Ciência da Informação*, Brasília, v.35, n.2, p.39–50, mai–ago 2006. Disponível em: <<http://www.scielo.br/pdf/ci/v35n2/a05v35n2.pdf>>. Acesso em: 16 nov. 2017.

CYBIS, W.; BETIOL, A. H.; FAUST, R. **Ergonomia e usabilidade:** conhecimentos, métodos e aplicações. São Paulo: Novatec editora, 2007.

DATAcite. **re3data.org - Registry of Research Data Repositories.** 2018. Disponível em: <<https://www.re3data.org/about>>. Acesso em: 01 out. 2018.

DataONE, Data Observation Network for Earth. **Data Life Cycle.** 2012. Disponível em: <<https://www.dataone.org/data-life-cycle>>. Acesso em: 13 nov. 2017.

DENTON. **Denton Declaration:** an open data manifesto. 2012. Disponível em: <<https://openaccess.unt.edu/denton-declaration>>. Acesso em: 12 nov. 2017.

DIAS, C. **Usabilidade na WEB:** criando portais mais acessíveis. Rio de Janeiro: Alta Books, 2007.

DIX, A.; FINLAY, J.; ABOARD, G. D.; BEALE, R. **Human-computer interaction.** 3.ed. Harlow, England: Pearson, 2004.

DUDZIAK, E. **Dados de Pesquisa agora devem ser armazenados e citados.** *SIBiUSP*, São Paulo, 2016. Disponível em: <<http://www.sibi.usp.br/?p=6189>>. Acesso em: 11 nov. 2017.

ECReI, European Commission Research & Innovation. **HORIZON 2020 em breves palavras.** 2014. Disponível em: <http://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_PT_KI0213413PTN.pdf>. Acesso em: 08 nov. 2017.

ECReI, European Commission Research & Innovation. **H2020 Online Manual, Open Access.** 2017. Disponível em: <http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm>. Acesso em: 15 nov. 2017.

ERP, J. A. A. van; LANGEN, C. D.; BOON, A.; BOCHOVE, K. van. **Testing the FAIR metrics on data catalogs.** *PeerJ Preprints*, Londres, v.6, p.e27151v2, set 2018. Disponível em: <<https://doi.org/10.7287/peerj.preprints.27151v2>>. Acesso em: 02 out. 2018.

- FINCH, D. J. **Accessibility, Sustainability, Excellence - How to Expand Access to Research Publications**: a review of progress in implementing the recommendations of the finch report. **Finch Group**, Reino Unido, out 2013. Disponível em: <<https://www.acu.ac.uk/research-information-network/review-of-implementation-finch-report-recommendations>>. Acesso em: 12 set. 2019.
- FORCE11. **The FAIR Data Principles**. La Jolla, CA: Force11, 2016. Disponível em: <<https://www.force11.org/group/fairgroup/fairprinciples>>. Acesso em: 17 ago. 2018.
- FURTADO, M. M. F. D. **Usabilidade e acessibilidade no Repositório de Informação Acessível da UFRN**: avaliação ergonômica de interfaces web. 2016. Dissertação (Mestrado em Gestão de Processos Institucionais) — UFRN, Natal, RN. Disponível em: <<https://repositorio.ufrn.br/jspui/handle/123456789/21773>>. Acesso em: 02 ago. 2019.
- GRAY, W. D.; SALZMAN, M. C. **Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods**. **Human-Computer Interaction**, [S.l.], v.13, n.3, p.203–261, 1998. Disponível em: <https://doi.org/10.1207/s15327051hci1303_2>. Acesso em: 28 jun. 2019.
- HARNAD, S.; BRODY, T.; VALLIÈRES, F.; CARR, L.; HITCHCOCK, S.; GINGRAS, Y.; OPPENHEIM, C.; HAJJEM, C.; HILF, E. R. **The Access/Impact Problem and the Green and Gold Roads to Open Access**: an update. **Serials Review**, [S.l.], v.34, n.1, p.36–40, 2008. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0098791307001530>>. Acesso em: 20 nov. 2017.
- HEDSTROM, M. **Digital preservation**: a time bomb for digital libraries. **Computers and the Humanities**, Países Baixos, v.31, mai 1997. Disponível em: <https://www.researchgate.net/publication/30843346_Digital_Preservation_A_Time_Bomb_for_Digital_Libraries>. Acesso em: 02 out. 2018.
- HEE KIM, H.; HO KIM, Y. **Usability study of digital institutional repositories**. **The electronic library**, [S.l.], v.26, n.6, p.863–881, 2008. Disponível em: <<https://www.emeraldinsight.com/doi/abs/10.1108/02640470810921637>>. Acesso em: 02 mai. 2019.
- IBICT. **Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica**. 2005. Disponível em: <<http://livroaberto.ibict.br/Manifesto.pdf>>. Acesso em: 17 nov. 2017.
- IBICT. **Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã**. 2016. 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: 17 nov. 2017.
- International Organization for Standardization. **ISO 9241-11:1998** ergonomic requirements for office work with visual display terminals (vdts) — part 11: guidance on usability. Genebra: ISO, 1998. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en>>. Acesso em: 16 mai. 2019.

- International Organization for Standardization. **ISO 9241-11:2018 ergonomics of human-system interaction - part 11: usability: definitions and concepts**. Genebra: ISO, 2018. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>>. Acesso em: 16 mai. 2019.
- JORDAN, P. W. **An introduction to usability**. Londres: Taylor & Francis Ltd, 1998. 136p.
- KAFURE, I.; CUNHA, M. **Usabilidade em ferramentas tecnológicas para o acesso à informação Software tools usability for information retrieval p. 273-282. Revista ACB**, Florianópolis, SC, v.11, n.2, p.273–282, 2006. Disponível em: <<https://revista.acbsc.org.br/racb/article/view/483>>. Acesso em: 04 jun. 2019.
- Kaufman-Wills Group. **The facts about open access: a study on financial and non-financial effects of alternative business models for scholarly journals**. [S.l.]: Association of Learned and Professional Society Publishers, 2005. Disponível em: <<https://www.alpsp.org/write/MediaUploads/FAOAaddendum.pdf>>. Acesso em: 17 nov. 2017.
- KOSHIYAMA, D. C. A. D. G. **Análise da usabilidade e da arquitetura da informação do Repositório Institucional da UFRN**. 2014. Dissertação (Mestrado em Design) — UFRN, Natal, RN. Disponível em: <<https://repositorio.ufrn.br/jspui/handle/123456789/19940>>. Acesso em: 02 ago. 2019.
- KRUG, S. **Não me faça pensar: uma abordagem de bom senso à usabilidade na web e mobile**. Rio de Janeiro: Alta Books, 2014.
- KÖRBER, N.; SULEMAN, H. Usability of Digital Repository Software: a study of dspace installation and configuration. In: BUCHANAN, G.; MASOODIAN, M.; CUNNINGHAM, S. J. (Ed.). **Digital Libraries: universal and ubiquitous access to information**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p.31–40. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-540-89533-6_4>. Acesso em: 26 jun. 2019.
- LAVERY, D.; COCKTON, G.; ATKINSON, M. P. **Comparison of evaluation methods using structured usability problem reports. Behaviour & Information Technology**, [S.l.], v.16, n.4-5, p.246–266, 1997.
- LAWRENCE, S. **Free online availability substantially increases a paper's impact. Nature Publishing Group**, [S.l.], v.411, 2001. Disponível em: <<https://www.nature.com/articles/35079151>>. Acesso em: 17 nov. 2017.
- LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto**. Brasília: IBICT, 2009. Disponível em: <[http://livroaberto.ibict.br/bitstream/1/775/4/Como gerenciar e ampliar a visibilidade da informação científica brasileira.pdf](http://livroaberto.ibict.br/bitstream/1/775/4/Como%20gerenciar%20e%20ampliar%20a%20visibilidade%20da%20informa%C3%A7%C3%A3o%20cient%C3%ADfica%20brasileira.pdf)>. Acesso em: 18 nov. 2017.
- LIKERT, R. **A technique for the measurement of attitudes. Archives of Psychology**, Nova Iorque, v.22, 1932. Disponível em: <https://legacy.voteview.com/pdf/Likert_1932.pdf>. Acesso em: 22 mai. 2019.

- MACHADO, L.; FERREIRA, E.; VERGARA, L. **Métodos de Avaliação de Usabilidade: características e aplicações.** 3º CONEPRO-SUL Congresso de Engenharia de Produção da Região Sul, Joinville, abr 2014. Disponível em: <https://www.researchgate.net/publication/306375261_METODOS_DE_AVALIACAO_DE_USABILIDADE_CHARACTERISTICAS_E_APLICACOES>. Acesso em: 22 mai. 2019.
- MAGUIRE, M. **Context of use within usability activities.** *International Journal of Human-Computer Studies*, Loughborough, Inglaterra, v.55, n.4, p.453–483, 2001. Disponível em: <<https://hdl.handle.net/2134/2206>>. Acesso em: 20 mai. 2019.
- MARTINS NETTO, O. A. **Heurísticas e guidelines para apresentação de hiperdocumentos multimídia na Web.** 2002. 152p. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) — USP, São Carlos, SP. Disponível em: <<http://doi.org/10.11606/D.55.2002.tde-17062015-112030>>. Acesso em: 17 jun. 2019.
- MELERO, R.; HERNÁNDEZ-SAN-MIGUEL, F. J. **Acceso abierto a los datos de investigación, una vía hacia la colaboración científica.** *Revista Española de Documentación Científica*, Madri, v.37(4), 2014. Disponível em: <<http://redc.revistas.csic.es/index.php/redc/article/view/869/1175>>. Acesso em: 16 nov. 2017.
- NASCIMENTO, J. A. M. do; AMARAL, S. A. do. **Avaliação de usabilidade na Internet.** Brasília: Thesaurus, 2010. 142p.
- NIELSEN, J. **Usability engineering.** Mountain View, California: Morgan Kaufmann, 1993. 362p.
- NIELSEN, J. Usability Inspection Methods. In: CONFERENCE COMPANION ON HUMAN FACTORS IN COMPUTING SYSTEMS, 1994, New York, NY, USA. **Anais...** ACM, 1994. p.413–414. (CHI '94). Disponível em: <<http://doi.acm.org/10.1145/259963.260531>>. Acesso em: 13 jun. 2019.
- NIELSEN, J. **Durability of Usability Guidelines.** 2005. Disponível em: <<https://www.nngroup.com/articles/durability-of-usability-guidelines>>. Acesso em: 13 jun. 2019.
- NIELSEN, J.; MOLICH, R. Heuristic Evaluation of User Interfaces. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 1990, New York, NY, USA. **Anais...** ACM, 1990. p.249–256. (CHI '90). Disponível em: <<http://doi.acm.org/10.1145/97243.97281>>. Acesso em: 13 jun. 2019.
- NSB, National Science Board. **Long-Lived Digital Data Collections: enabling research and education in the 21st century.** Alexandria, Virginia, EUA: National Science Foundation-NSF, 2005. 92p. Disponível em: <<https://www.nsf.gov/geo/geo-data-policies/nsb-0540-1.pdf>>. Acesso em: 12 jun. 2018.
- OECD, Organisation for Economic Co-operation & Development. **OECD Principles and Guidelines for Access to Research Data from Public Funding.** 2007. Disponível em: <<http://www.oecd.org/sti/sci-tech/38500813.pdf>>. Acesso em: 14 nov. 2017.

- PACKER, A. L.; COP, N.; LUCCISANO, A.; RAMALHO, A.; SPINAK, E. **SciELO - 15 Anos de Acesso Aberto**: um estudo analítico sobre acesso aberto e comunicação científica. Paris: UNESCO, 2014. 19p. Disponível em: <<http://www.scielo15.org/livro-scielo-15/>>. Acesso em: 17 nov. 2017.
- PALMA, J. L.; SANTOS, A. C. de Lima dos; JORDÃO, S. S. de; SOUSA, R. J. do Prado Leite de. **AVALIAÇÃO DE REPOSITÓRIOS INSTITUCIONAIS**: proposta de padronização de critérios. **V Encontro Regional dos Estudantes de Biblioteconomia, Documentação, Ciência e Gestão da Informação das Regiões Sudeste, Centro-Oeste e Sul**, Belo Horizonte, nov 2018. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/moci/article/view/3811>>. Acesso em: 28 jun. 2019.
- PAVÃO, C. G.; VANZ, S. A. d. S.; PASSOS, P. C. S. J.; CAREGNATO, S. E.; AZAMBUJA, L. A. B.; BORGES, E. N.; GABRIEL JUNIOR, R. F.; ROCHA, R. P. d. **Acesso aberto a dados de pesquisa no Brasil**: repositórios brasileiros de dados de pesquisa. Porto Alegre, RS: UFRGS, 2018. Relatório técnico, Disponível em: <<http://hdl.handle.net/10183/185138>>. Acesso em: 20 dez. 2018.
- PEREIRA, F. **Avaliação de usabilidade em bibliotecas digitais**: um estudo de caso. 2011. Dissertação (Mestrado em Ciência da Informação) — UFMG, Belo Horizonte, MG. Disponível em: <<http://hdl.handle.net/1843/ECID-8LAKHD>>. Acesso em: 20 mai. 2019.
- POLSON, P. G.; LEWIS, C. H.; RIEMAN, J.; WHARTON, C. **Cognitive Walkthroughs**: a method for theory-based evaluation of user interfaces. **International Journal of Man-Machine Studies**, [S.l.], v.36, p.741–773, 1992. Disponível em: <http://sonify.psych.gatech.edu/~ben/references/polson_cognitive_walkthroughs_a_method_for_theory-based_evaluation_of_user_interfaces.pdf>. Acesso em: 09 jul. 2019.
- ROCHA, H. V.; BARANAUSKAS, M. C. C. **Design e avaliação de interfaces humano-computador**. [S.l.]: Creative Commons, Brasil, 2003. Disponível em: <<http://200.18.252.196:8080/pergamumweb/vinculos/000000/000000ce.pdf>>. Acesso em: 28 mai. 2019.
- ROCHA, R. P. d.; AZAMBUJA, L. A. B.; BORGES, E. N.; GABRIEL JUNIOR, R. F.; CAREGNATO, S. E.; PAVÃO, C. G.; VANZ, S. A. d. S.; PASSOS, P. C. S. J. **Acesso aberto a dados de pesquisa no Brasil**: soluções tecnológicas. Porto Alegre, RS: UFRGS, 2018. Relatório técnico, Disponível em: <<http://hdl.handle.net/10183/185126>>. Acesso em: 20 dez. 2018.
- ROCHA, R. P. da; CIUFFO, L. N.; REIS, C.; LAZARTE, L.; FERREIRA, V.; GARCIA, M. **Acesso Aberto a Dados de Pesquisa**. Belém, PA: 18º WRNP, 2017. Cartaz, Disponível em: <http://wrnp.rnp.br/sites/wrnp2017/files/11_wrnp2017_cartaz_aadp_design.pdf>. Acesso em: 10 nov. 2017.
- RODRIGUES, R. S.; OLIVEIRA, A. B. de. **Periódicos científicos na América Latina**: títulos em acesso aberto indexados no isi e scopus. **Perspectivas em Ciência da Informação**, Belo Horizonte, MG, v.17, n.4, p.77–

- 99, dez 2012. Disponível em: <http://www.scielo.br/scielo.php?pid=S1413-99362012000400006&script=sci_arttext&tlng=pt>. Acesso em: 17 fev. 2019.
- SANTOS, D. B. dos. **Publicações ampliadas** : aspectos da integração de dados de pesquisa. 2017. 187p. Dissertação (Mestrado em Comunicação e Informação) — UFRGS, Porto Alegre, RS. Disponível em: <<http://hdl.handle.net/10183/172948>>. Acesso em: 07 out. 2018.
- SANTOS, D. B. dos; MOURA, A. M. M. de; PAVÃO, C. M. G. **Proposta metodológica para avaliação de usabilidade de interfaces de repositórios. Cadernos de Biblioteconomia, Arquivística e Documentação**, Coimbra, Portugal, n.2, p.215–218, 2014. Disponível em: <<https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1210>>. Acesso em: 17 mai. 2019.
- SANTOS, D. B. dos; PAVÃO, C. M. G.; MOURA, A. M. M. de. **Usabilidade do Lume** : repositório digital da ufrgs : uma avaliação por meio das heurísticas e de testes com usuários. **InCID : Revista de Ciência da Informação e Documentação - FF-CLRP - USP**, Ribeirão Preto, SP, v.7, n.1, p.150–166, mar–ago 2016. Disponível em: <<https://lume.ufrgs.br/handle/10183/142583>>. Acesso em: 17 mai. 2019.
- SAYÃO, L. F. **Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos. Revista Ponto de Acesso**, Salvador, BA, v.4, n.3, p.68–94, dez 2010. Disponível em: <<https://portalseer.ufba.br/index.php/revistaici/article/view/4709/3565>>. Acesso em: 15 nov. 2017.
- SAYÃO, L. F.; SALES, L. F. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: <http://www.cnen.gov.br/images/CIN/PDFs/GUIA_DE_DADOS_DE_PESQUISA.pdf>. Acesso em: 22 nov. 2017.
- SAYÃO, L. F.; SALES, L. F. **Algumas considerações sobre os repositórios digitais de dados de pesquisa. Informação & Informação**, Londrina, PR, v.21, n.2, p.90–115, 2016. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/27939>>. Acesso em: 15 jun. 2018.
- SEABRA, N. S. **Usabilidade do Repositório Institucional da Universidade Federal da Bahia**: método de inspeção tecnológica baseado nos princípios heurísticos de Jakob Nielsen. 2018. Dissertação (Mestrado em Ciência da Informação) — UFBA, Salvador, BA. Disponível em: <<http://repositorio.ufba.br/ri/handle/ri/27723>>. Acesso em: 02 mai. 2019.
- SHANKAR MAHATO, S.; GAJBE, S. **A Comparative study of Open source data repository software: dataverse and ckan. Library Herald**, Nova Deli, Índia, v.56, n.1, p.36–54, jan 2018. Disponível em: <<http://dx.doi.org/10.5958/0976-2469.2018.00005.2>>. Acesso em: 17 jul. 2019.
- SILVA, R. T. S. e. **Estudo de usabilidade em repositório digital de acesso aberto**: o caso do repositório. 2014. Dissertação (Mestrado em Sistemas de

Informação) — UMinho-Universidade do Minho, Braga, Portugal. Disponível em: <<http://hdl.handle.net/1822/35211>>. Acesso em: 02 mai. 2019.

SILVA, V. M. **Inspeção de usabilidade do site furg.br através do método percurso cognitivo**. 2017. Dissertação (Mestrado em Modelagem Computacional) — FURG, Rio Grande, RS. Disponível em: <<https://sistemas.furg.br/sistemas/sab/arquivos/bdtd/0000011883.pdf>>. Acesso em: 02 ago. 2019.

SMITH, S. L.; MOSIER, J. N. **Guidelines for designing user interface software**. Bedford, Massachusetts, EUA: MITRE Corporation, 1986. Relatório técnico, Disponível em: <<https://apps.dtic.mil/docs/citations/ADA177198>>. Acesso em: 28 jun. 2019.

SOUZA, A. H. L. R. de; OLIVEIRA, A. F. de; D'AVILA, R. T.; SILVA SALES CHAVES, E. da. **O modelo de referência OAIS e a preservação digital distribuída**. *Revista Ciência da Informação*, Brasília, DF, v.41, n.1, p.65–73, jan–abr 2012. Disponível em: <http://www.brapci.inf.br/_repositorio/2014/11/pdf_20ae7f8b97_0031823.pdf>. Acesso em: 17 jul. 2019.

STANTON, N. A.; HEDGE, A.; BROOKHUIS, K.; SALAS, E.; HENDRICK, H. W. **Handbook of human factors and ergonomics methods**. Boca Raton, Florida: CRC press, 2005.

THE ROYAL SOCIETY. **Science as an open enterprise**. Londres: The Royal Society Science Policy Centre, 2012. 105p. Disponível em: <<https://royalsociety.org/~media/policy/projects/sape/2012-06-20-saoe.pdf>>. Acesso em: 12 jul. 2018.

VEIGA, V. S. d. O.; MACHADO, R. R.; ALVES, A. d. S.; PIMENTA, D. N.; SILVA, C. H. d.; CARVALHO, M. d. C. R. d. **Avaliação da usabilidade em repositórios institucionais: revisão de literatura**. *RECIIS – Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, Rio de Janeiro, RJ, v.8, n.4, p.540–553, 2014. Disponível em: <<https://www.arca.fiocruz.br/handle/icict/12869>>. Acesso em: 19 jun. 2019.

WINCKLER, M.; PIMENTA, M. S. **Avaliação de usabilidade de sites web**. *ESCOLA REGIONAL de Informática - SBC Sul (ERI 2002)*, Porto Alegre, RS, p.1–54, 2002. Disponível em: <<https://www.irit.fr/~Marco.Winckler/2002-winckler-pimenta-ERI-2002-cap3.pdf>>. Acesso em: 20 mai. 2019.

APÊNDICE A DOCUMENTOS PARA AVALIAÇÃO HEURÍSTICA

- Descrição da metodologia;
- Termo de consentimento e formulário.

A.1 Descrição da metodologia

AValiação Heurística do DataVerse do IBICT

O objetivo desta inspeção de usabilidade é avaliar se a interface do sistema está de acordo com os princípios de usabilidade propostos, assim denominados heurísticas.

Os avaliadores deverão, individualmente, acessar e utilizar o repositório de dados de pesquisa do IBICT, <https://repositoriopesquisas.ibict.br>, em suas diversas funcionalidades e apontar os problemas encontrados, conflitantes com cada uma das regras descritas abaixo.

Para cada um dos problemas encontrados, um grau de severidade deve ser atribuído. A escala é apresentada ao final deste documento.

Heurísticas

1. Correspondência entre o sistema e o mundo real

O sistema deve falar a linguagem dos usuários, com palavras, frases e conceitos familiares à ele, em vez de termos orientados pelo sistema, portanto a linguagem utilizada deve ser a mais natural possível, evitando códigos do sistema. Siga as convenções do mundo real, fazendo as informações aparecerem em uma ordem natural e lógica.

É necessário muito cuidado na utilização de jargões técnicos, acrônimos, linguagem complexa e termos obscuros, os usuários do sistema não devem ter que compreender esse linguajar. Cuidado também com elementos não verbais, como imagens e ícones, para torná-los intuitivos e aproximá-los de objetos e atividades do mundo real.

<https://www.nngroup.com/articles/match-system-real-world>

2. Consistência e padrões

Os usuários não devem se perguntar se palavras, situações ou ações diferentes significam a mesma coisa. Siga as convenções, a criatividade do design nunca deve prejudicar o entendimento da plataforma conseguido ao seguir os padrões estabelecidos.

O sistema deve garantir integridade de suas funções, que deverão ter o mesmo efeito quando executadas, além de apresentar informações de acordo com um padrão estabelecido, facilitando assim o reconhecimento.

<https://www.nngroup.com/articles/do-interface-standards-stifle-design-creativity>

3. Design estético e minimalista

Os diálogos não devem conter informações irrelevantes ou raramente necessárias. Cada unidade extra de informação em um diálogo compete com as unidades relevantes de informação e diminui sua visibilidade relativa. Interfaces simples contendo somente a informação necessária e oferecendo somente as opções importantes naquele contexto facilitam a interação do usuário com o sistema tornando-a natural e facilmente compreendida e utilizada.

4. Reconhecimento ao invés de lembrar

Minimize a carga de memória do usuário, tornando objetos, ações e opções visíveis. O usuário não deve ter que lembrar informações de uma parte do diálogo para outra. As instruções de uso do sistema devem ser visíveis ou facilmente recuperáveis sempre que apropriado. Histórico de pesquisas anteriores, botões e opções de menu facilmente reconhecidos, uso de dicas e textos de ajuda no preenchimento de um formulário ou durante a execução de uma tarefa, são exemplos de reconhecimento ao invés de lembrar.

<https://www.nngroup.com/articles/recognition-and-recall>

5. Visibilidade do status do sistema

O sistema deve sempre manter os usuários informados sobre o que está acontecendo, através de feedback apropriado dentro de um prazo razoável.

Trata-se de comunicação e transparência, o sistema deve manter um diálogo contínuo com o usuário, sobre o que o ele está fazendo e inserindo no mesmo, principalmente antes de ações definitivas como subscrição de arquivos ou exclusão de dados. A informação deve ser constante e possibilitar tanto a tomada de decisões quanto informar se a operação foi bem sucedida ou não.

<https://www.nngroup.com/articles/visibility-system-status>

6. Flexibilidade e eficiência de uso

Os aceleradores - invisíveis para o usuário novato - podem acelerar a interação do usuário especialista, de modo que o sistema possa atender a usuários inexperientes e experientes, permitindo assim que os usuários personalizem ações frequentes.

Devem ser disponibilizados atalhos para ações frequentes: teclas de função ou combinação de teclas, diálogos com teclas para acessar as funções mais importantes.

7. Controle do usuário e liberdade

É normal os usuários escolherem as funções do sistema por engano e precisarão de uma "saída de emergência" claramente marcada para deixar o estado indesejado sem ter que passar por um diálogo extenso. O sistema deve suportar desfazer e refazer.

Se forem oferecidas as opções de restauração e contorno de uma determinada ação para o usuário, a segurança em explorar o sistema aumentará, facilitando o aprendizado de forma exploratória.

8. Ajude os usuários a reconhecer, diagnosticar e recuperar erros

Mensagens de erro devem ser expressas em linguagem clara e simples (sem códigos), indicar precisamente o problema, de forma que o usuário entenda facilmente o ocorrido e sugerir de forma construtiva uma solução.

Sons (beeps) e mudança de cores, quando usados isoladamente, não são bons exemplos de mensagens de erro, lembre-se de usuários com necessidades especiais. Mensagens de erro em áreas não destacadas também não são boas escolhas.

<https://www.nngroup.com/articles/error-message-guidelines>

9. Prevenção de erros

Ainda melhor do que boas mensagens de erro é um projeto cuidadoso que impede que um problema ocorra em primeiro lugar. Elimine as condições propensas a erros ou verifique-as e apresente aos usuários uma opção de confirmação antes de se comprometerem com a ação.

Um caso clássico é a obrigatoriedade e consistência de dados em um formulário antes de submetê-lo.

<https://www.nngroup.com/articles/slips>

10. Ajuda e documentação

Mesmo que seja melhor o sistema ser usado sem documentação, pode ser necessário fornecer ajuda e documentação. Qualquer informação desse tipo deve ser fácil de pesquisar, focada na tarefa do usuário, listar as etapas concretas a serem executadas e não ser muito grande. Quanto mais simples de usar melhor será, mas todo ambiente deve conter informações de auxílio para as tarefas disponíveis.

Classificação de gravidade

Para cada problema de usabilidade encontrado e categorizado de acordo com as regras descritas acima, cada avaliador deverá atribuir um nível de severidade único por problema, de acordo com a tabela abaixo, levando em conta:

- A frequência com que o problema ocorre, comum ou raro?
- O impacto do problema ao ocorrer, fácil ou difícil do usuário superar?
- A persistência do problema - ao superar o problema, ele é único ou ficará se repetindo e incomodando o usuário?

A combinação dos fatores acima deve gerar um grau de severidade do problema que representará a gravidade que o mesmo possui e a prioridade que deve ser dada para a correção do mesmo.

Nível de severidade	Tipo	Descrição
0	Sem importância	Não afeta a operação, não é um problema de usabilidade impactante.
1	Cosmético	Apenas um problema cosmético sem grande impacto, não precisa ser corrigido a menos que haja tempo extra disponível no projeto.
2	Simple	Pequeno problema de usabilidade, pode ser corrigido aplicando baixa prioridade.
3	Grave	Grande problema de usabilidade, deve ser dada alta prioridade para a correção do problema.
4	Catastrófico	Problema de usabilidade catastrófico, necessária correção urgente.



A.2 Termo de consentimento e formulário

Avaliação Heurística de Usabilidade do Dataverse

*Obrigatório

1. Endereço de e-mail *

Termo de Consentimento Livre e Esclarecido

Você está sendo convidado(a) para participar do projeto de pesquisa acima identificado. O documento abaixo contém todas as informações necessárias sobre a pesquisa que estamos realizando.

IDENTIFICAÇÃO DO PROJETO DE PESQUISA

Título do Projeto: Avaliação de Usabilidade de um repositório de dados de pesquisa brasileiro
 Área do Conhecimento: Computação
 Curso: Mestrado em Engenharia da Computação
 Patrocinador da pesquisa: O próprio pesquisador
 Instituição onde será realizado: Universidade Federal do Rio Grande - FURG
 Nome dos pesquisadores e colaboradores: Luís Alberto Barbosa Azambuja e Eduardo Nunes Borges

IDENTIFICAÇÃO DO PESQUISADOR RESPONSÁVEL

Nome: Luís Alberto Barbosa Azambuja
 Profissão: Técnico em Tecnologia da Informação
 E-mail: cpdlaa@furg.br
 Telefone: (53) 3233-6580

DOS TERMOS DA PARTICIPAÇÃO

- ▶ Da justificativa e dos objetivos para realização desta pesquisa.
 Uma ferramenta de repositório de dados de pesquisa é importantíssima para a comunidade científica. A aceitação desta ferramenta por parte dos usuários é determinante para seu sucesso e a usabilidade do sistema é um dos principais fatores a serem considerados. Esse estudo visa avaliar e identificar problemas na interface do Dataverse, ferramenta escolhida pelo Grupo de Trabalho Rede de Dados de Pesquisa Brasileira (GT-RDP Brasil), criado a partir de uma iniciativa da Rede Nacional de Ensino e Pesquisa (RNP) e do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), com a participação da Universidade Federal do Rio Grande (FURG) e da Universidade Federal do Rio Grande do Sul (UFRGS), para escolher e implantar um repositório de dados de pesquisa em âmbito nacional.
- ▶ Do objetivo de minha participação.
 Avaliar a usabilidade da interface do Dataverse através de uma técnica analítica (avaliação heurística) e de outra empírica.
- ▶ Do procedimento para coleta de dados.
 Os dados serão coletados através da plataforma eletrônica Google Forms, com autorização prévia.
- ▶ Da utilização, armazenamento e descarte das amostras.
 Os autores do projeto de pesquisa se comprometem a manter o sigilo dos dados coletados. Concordam, igualmente, que estas informações serão utilizadas única e exclusivamente com finalidade científica, preservando-se integralmente o anonimato dos participantes.

- ▶ Da isenção e ressarcimento de despesas.
Minha participação é isenta de despesas, portanto não receberei qualquer ressarcimento.
- ▶ Da liberdade de recusar, desistir ou retirar meu consentimento.
Tenho a liberdade de recusar, desistir ou de interromper a colaboração nesta pesquisa no momento em que desejar, sem necessidade de qualquer explicação. A minha desistência não me causará nenhum prejuízo.
- ▶ Da garantia de sigilo e de privacidade.
Os resultados obtidos durante este estudo serão mantidos em sigilo, mas concordo que sejam divulgados em publicações científicas, desde que eu não seja identificado(a).
- ▶ Da garantia de esclarecimento e informações a qualquer tempo.
Tenho a garantia de tomar conhecimento e obter informações, a qualquer tempo, dos procedimentos e métodos utilizados neste estudo, bem como dos resultados finais, desta pesquisa. Para tanto, poderei consultar o pesquisador responsável acima identificado.

Consentimento

2. Eu, após receber informações e esclarecimento sobre o projeto de pesquisa, acima identificado, concordo de livre e espontânea vontade em participar como voluntário(a): *

Marcar apenas uma oval.

Concordo voluntariamente com a minha participação, sabendo que poderei retirar o meu consentimento a qualquer momento, antes ou durante a pesquisa, sem penalidades ou prejuízos.

Não concordo em participar do estudo. *Pare de preencher este formulário.*

Descrição da metodologia

No link abaixo estão disponíveis as instruções de como proceder a avaliação de usabilidade. São apresentadas as heurísticas que devem ser consideradas e explicado como classificar os problemas encontrados de acordo com o seu nível de severidade:

<https://docs.google.com/document/d/1n63pwO8pfpno0vHtFDbMg437C7gMLxIEbXqT62kwbiA/edit?usp=sharing>

No formulário a seguir estarão disponíveis 5 possibilidades de preenchimento de problemas de cada uma das 10 heurísticas, se forem encontrados mais de 5 problemas para uma mesma heurística priorize para as que você considerar as mais graves.

Aconselhamos que mantenha a nova janela aberta para consulta durante a avaliação!

É importante diferenciar os problemas de conteúdo, como falta e inconsistência de dados submetidos, dos problemas da interface que é o que se almeja avaliar.

A página do repositório de dados de pesquisa que será avaliada é a do IBICT:

<https://repositoriopesquisas.ibict.br/>

3. Eu, após ler as instruções acima sobre a avaliação de usabilidade, concordo de livre e espontânea vontade em participar como voluntário(a): *

Marcar apenas uma oval.

Concordo voluntariamente com a minha participação, sabendo que poderei retirar o meu consentimento a qualquer momento, antes ou durante a pesquisa, sem penalidades ou prejuízos.

Não concordo em participar do estudo. *Pare de preencher este formulário.*

1 - Correspondência entre o sistema e o mundo real

4. O sistema utiliza uma linguagem familiar aos usuários ou usa jargões técnicos, acrônimos, linguagem complexa e termos obscuros? Os ícones e imagens correspondem ao mundo real?

5. **Gravidade:**

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

6. **Encontrou outro problema referente a esta heurística? ***

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 6.*

2 - Consistência e padrões

7. O sistema é consistente e segue os mesmos padrões em toda interface? As ações executadas têm os mesmos resultados independente do local ou situação?

8. **Gravidade:**

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

9. **Encontrou outro problema referente a esta heurística? ***

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 9.*

3 - Design estético e minimalista

10. A interface exibe somente a informação necessária ao usuário? Há excesso de conteúdo ou ele está confuso?

11. Gravidade:

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

12. Encontrou outro problema referente a esta heurística? *

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 12.*

4 - Reconhecimento ao invés de lembrar

13. É necessário lembrar informações entre partes do sistema? As funcionalidades são facilmente reconhecidas e as instruções estão disponíveis?

14. Gravidade:

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

15. Encontrou outro problema referente a esta heurística? *

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 15.*

5 - Visibilidade do status do sistema

16. O sistema mantém os usuários informados sobre o que está acontecendo, dentro de um prazo razoável? Toda interação tem feedback apropriado?

17. Gravidade:

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

18. Encontrou outro problema referente a esta heurística? *

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 18.*

6 - Flexibilidade e eficiência de uso

19. São oferecidos atalhos e outros meios de otimizar o uso do sistema?

20. Gravidade:

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

21. Encontrou outro problema referente a esta heurística? *

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 21.*

7 - Controle do usuário e liberdade

22. O sistema oferece, de maneira simples e clara, opções de cancelar e desfazer ações indesejadas?

23. **Gravidade:**

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

24. **Encontrou outro problema referente a esta heurística? ***

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 24.*

8 - Reconhecimento, diagnóstico e recuperação de erros

25. O sistema oferece mensagens de erro simples e esclarecedoras? Os erros são facilmente identificáveis?

26. **Gravidade:**

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

27. **Encontrou outro problema referente a esta heurística? ***

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 27.*

9 - Prevenção de erros

28. O sistema possui erros ou falhas graves? O quão facilmente poderiam ser corrigidos?

29. **Gravidade:**

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

30. **Encontrou outro problema referente a esta heurística? ***

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 30.*

10 - Ajuda e documentação

31. O sistema disponibiliza documentação? É oferecida ajuda de forma simples e clara para as funcionalidades disponíveis?

32. **Gravidade:**

Marcar apenas uma oval.

- 0 - Sem importância
- 1 - Cosmético
- 2 - Simples
- 3 - Grave
- 4 - Catastrófico

33. **Encontrou outro problema referente a esta heurística? ***

Marcar apenas uma oval.

- Sim
- Não *Ir para a pergunta 33.*

Questões gerais:

34. **Qual seu nível de dificuldade inicial ao começar a usar o repositório Dataverse do IBICT? ***

Marcar apenas uma oval.

- Muito fácil
- Fácil
- Indiferente
- Difícil
- Muito difícil

35. **Qual seu nível de dificuldade geral ao usar o repositório Dataverse do IBICT? ***

Marcar apenas uma oval.

- Muito fácil
- Fácil
- Indiferente
- Difícil
- Muito difícil

36. **Qual seu nível de satisfação ao usar o repositório Dataverse do IBICT? ***

Marcar apenas uma oval.

- Muito satisfeito
- Satisfeito
- Indiferente
- Insatisfeito
- Muito insatisfeito

37. **Declaro que estou ciente que meus dados neste questionário serão utilizados para pesquisa descrita conforme o termo de consentimento livre e esclarecido. ***

Marque todas que se aplicam.

Concordo

Envie para mim uma cópia das minhas respostas.

APÊNDICE B DOCUMENTOS PARA AVALIAÇÃO EMPÍRICA

- Descrição da metodologia;
- Roteiro do experimento;
- Ficha de acompanhamento;
- Termo de consentimento;
- Lista de tarefas;
- Formulário.

B.1 Descrição da metodologia

AVALIAÇÃO EMPÍRICA DO DATAVERSE DO IBICT

O objetivo desta inspeção de usabilidade é avaliar a interface do repositório de dados de pesquisa do IBICT.

Os usuários deverão, individualmente, acessar e utilizar o repositório Dataverse do IBICT, disponível em <https://repositoriopesquisas.ibict.br>, e realizar 5 tarefas simples, básicas a um visitante que deseje localizar, baixar e citar dados de pesquisa.

Durante a execução dessas tarefas ele será acompanhado por um avaliador que irá registrar o horário de início e fim de cada tarefa e se ela foi completamente ou parcialmente completada ou ainda se não foi. Também irá fazer anotações e questionar o participante sobre sua interação: pensamentos, dúvidas e expectativas. Entre as perguntas feitas estão incluídas:

- Poderia me dizer o que está pensando?
- O que está tentando fazer?
- Por que fez isso?

Ao mesmo tempo ele poderá fazer anotações que posteriormente serão agregadas as do avaliador.

No final do experimento, após o usuário terminar todas as tarefas, será solicitado que ele preencha um questionário com 10 questões relativas a sua satisfação com diversos aspectos da ferramenta.

Após todos usuários terem encerrado sua participação, as respostas serão analisadas para elencar os problemas encontrados e as tentativas de solução. Também serão geradas as notas médias e analisados os resultados provenientes desses dados.



B.2 Roteiro do experimento

Roteiro a ser seguido pelo avaliador durante a avaliação empírica

Antes de iniciar o experimento:

- Recepcionar o participante, cumprimentar, se apresentar e agradecer.
- Explicar o propósito da avaliação e o objetivo do estudo.
 - Essa avaliação é para subsidiar a dissertação de mestrado e o intuito é avaliar a interface do Dataverse, um repositório para dados de pesquisa. Dados de pesquisa são os dados usados em um projeto de pesquisa e que geraram trabalhos científicos.
- Explicar como se dará a avaliação e como será a participação dele.
 - Explicar que serão dadas tarefas mas que não há tempo limite para a realização do experimento.
 - Esclarecer que não é necessário completar todas tarefas, o importante é avaliar a facilidade ou dificuldade na execução delas e os aspectos que poderiam melhorar essa interação.
- Esclarecer que o objetivo é avaliar a interface e não o usuário.
- Garantir ao participante a confidencialidade da sua participação, somente serão divulgados os dados do experimento, mas nenhum dados pessoal.
- Mostrar o documento de consentimento de participação, solicitar que o leia e assine. Informar novamente sobre a possibilidade de desistir a qualquer momento, sem necessidade de explicar motivos.
- Solicitar que o participante desligue celular e outros aparelhos que possam interromper ou atrapalhar o teste.
- Mostrar o equipamento que será utilizado e posicioná-lo para começar o teste.
- Mostrar o documento com as tarefas, esperar ler e esclarecer dúvidas iniciais.

Durante o experimento:

- Registrar data e horário de início.
- Registrar horário de término ou desistência de cada tarefa.
- Solicitar que o participante verbalize seus pensamentos, dúvidas e expectativas.
- Esclarecer dúvidas conceituais do sistema sem indicar caminhos ou soluções. O participante não conseguir completar uma tarefa também é um resultado válido para o teste de usabilidade.
- Registrar horário de término da avaliação.

Ao final do experimento:

- Solicitar que o participante preencha o questionário.
- Agradecer novamente a participação.

B.3 Ficha de acompanhamento

**REPOSITÓRIO DE DADOS DE PESQUISA
FICHA DE ACOMPANHAMENTO DA AVALIAÇÃO EMPÍRICA**

Nº: _____ Participante: _____ Data: ____/____/____ Área: <input type="radio"/> Exatas <input type="radio"/> Biológicas <input type="radio"/> Engenharias <input type="radio"/> Saúde <input type="radio"/> Agrárias <input type="radio"/> Sociais <input type="radio"/> Humanas <input type="radio"/> Letras
Avaliador: _____

Tarefa 1 Completou a tarefa: <input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Parcialmente Hora de Início: _____
Dificuldades encontradas: _____ _____ _____
Sugestões: _____ _____ _____
Se “completou parcialmente”, o que conseguiu e o que faltou: _____ _____ _____
Hora de Fim: _____

Tarefa 2 Completou a tarefa: <input type="radio"/> Sim <input type="radio"/> Não <input type="radio"/> Parcialmente Hora de Início: _____
Dificuldades encontradas: _____ _____ _____
Sugestões: _____ _____ _____
Se “completou parcialmente”, o que conseguiu e o que faltou: _____ _____ _____
Hora de Fim: _____

Tarefa 3 Completou a tarefa: Sim Não Parcialmente Hora de Início: _____

Dificuldades encontradas: _____

Sugestões: _____

Se “completou parcialmente”, o que conseguiu e o que faltou: _____

Hora de Fim: _____

Tarefa 4 Completou a tarefa: Sim Não Parcialmente Hora de Início: _____

Dificuldades encontradas: _____

Sugestões: _____

Se “completou parcialmente”, o que conseguiu e o que faltou: _____

Hora de Fim: _____

Tarefa 5 Completou a tarefa: Sim Não Parcialmente Hora de Início: _____

Dificuldades encontradas: _____

Sugestões: _____

Se “completou parcialmente”, o que conseguiu e o que faltou: _____

Hora de Fim: _____

Considerações Finais: _____

Assinatura do participante

B.4 Termo de consentimento

Avaliação Empírica de Usabilidade - Termo de Consentimento

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado(a) para participar do projeto de pesquisa abaixo identificado. O documento abaixo contém todas as informações necessárias sobre a pesquisa que estamos realizando.

IDENTIFICAÇÃO DO PROJETO DE PESQUISA

Título do Projeto: Avaliação de Usabilidade de um repositório de dados de pesquisa brasileiro

Área do Conhecimento: Computação

Curso: Mestrado em Engenharia da Computação

Patrocinador da pesquisa: O próprio pesquisador

Instituição onde será realizado: Universidade Federal do Rio Grande - FURG

Nome dos pesquisadores e colaboradores: Luís Alberto Barbosa Azambuja e Eduardo Nunes Borges

IDENTIFICAÇÃO DO PESQUISADOR RESPONSÁVEL

Nome: Luís Alberto Barbosa Azambuja

Profissão: Técnico em Tecnologia da Informação

E-mail: cpdlaa@furg.br

Telefone: (53) 3233-6580

DOS TERMOS DA PARTICIPAÇÃO

- ▶ Da justificativa e dos objetivos para realização desta pesquisa.

Uma ferramenta de repositório de dados de pesquisa é importantíssima para a comunidade científica. A aceitação desta ferramenta por parte dos usuários é determinante para seu sucesso e a usabilidade do sistema é um dos principais fatores a serem considerados. Esse estudo visa avaliar e identificar problemas na interface do Dataverse, ferramenta escolhida pelo Grupo de Trabalho Rede de Dados de Pesquisa Brasileira (GT-RDP Brasil), criado a partir de uma iniciativa da Rede Nacional de Ensino e Pesquisa (RNP) e do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), com a participação da Universidade Federal do Rio Grande (FURG) e da Universidade Federal do Rio Grande do Sul (UFRGS), para escolher e implantar um repositório de dados de pesquisa em âmbito nacional.

- ▶ Do objetivo de minha participação.

Avaliar a usabilidade da interface do Dataverse através de uma técnica analítica (avaliação heurística) e de outra empírica.

- ▶ Do procedimento para coleta de dados.

Os dados serão coletados por colaboradores através da observação direta da interação dos usuários com o sistema, com autorização prévia, além da anotação de comentários deste e suas próprias anotações. No final será solicitado o preenchimento de uma ficha de satisfação em relação ao uso do Dataverse.

- ▶ Da utilização, armazenamento e descarte das amostras.

Os autores do projeto de pesquisa se comprometem a manter o sigilo dos dados coletados. Concordam, igualmente, que estas informações serão utilizadas única e exclusivamente com finalidade científica, preservando-se integralmente o anonimato dos participantes.

► Da isenção e ressarcimento de despesas.

Minha participação é isenta de despesas, portanto não receberei qualquer ressarcimento.

► Da liberdade de recusar, desistir ou retirar meu consentimento.

Tenho a liberdade de recusar, desistir ou de interromper a colaboração nesta pesquisa no momento em que desejar, sem necessidade de qualquer explicação. A minha desistência não me causará nenhum prejuízo.

► Da garantia de sigilo e de privacidade.

Os resultados obtidos durante este estudo serão mantidos em sigilo, mas concordo que sejam divulgados em publicações científicas, desde que eu não seja identificado(a).

► Da garantia de esclarecimento e informações a qualquer tempo.

Tenho a garantia de tomar conhecimento e obter informações, a qualquer tempo, dos procedimentos e métodos utilizados neste estudo, bem como dos resultados finais, desta pesquisa. Para tanto, poderei consultar o pesquisador responsável acima identificado.

DESCRIÇÃO DA METODOLOGIA

O experimento consiste em acessar a página do repositório Dataverse do IBICT (<https://repositoriopesquisas.ibict.br/>) e realizar as tarefas propostas em anexo.

Não importa se conseguir completar todas tarefas, o importante é exatamente verificar e apontar os problemas que dificultaram ou impediram a execução delas.

Durante a execução das tarefas é importante descrever para o colaborador o que está tentando fazer e as dificuldades encontradas.

Ao final será solicitado o preenchimento de uma ficha de satisfação .

CONSENTIMENTO

Eu, após receber informações e esclarecimento sobre o projeto de pesquisa, acima identificado, concordo de livre e espontânea vontade em participar como voluntário(a):

Concordo voluntariamente com a minha participação, sabendo que poderei retirar o meu consentimento a qualquer momento, antes ou durante a pesquisa, sem penalidades ou prejuízos.

Não concordo em participar do estudo.

Nº: _____ Participante: _____ Data: ____/____/____

Área: Exatas Biológicas Engenharias Saúde Agrárias Sociais Humanas Letras

Assinatura do participante

B.5 Lista de tarefas

Tarefas a serem realizadas pelos participantes

Em cada tarefa a seguir indique se conseguiu realizar a tarefa e descreva suas impressões sobre as dificuldades encontradas, problemas ou falta de informações; organização dos vários elementos e qualquer aspecto desagradável ou incorreto nos textos, cores, botões, menus, imagens e ícones.

1. Você quer escrever um artigo sobre “botânica marinha” e precisa de dados sobre o assunto.
Tarefa: A partir da página principal pesquise se existe uma base de dados sobre o assunto, verifique se encontrou algo relevante sobre sua pesquisa.
2. Você quer escrever um artigo sobre “plantas aquáticas” e precisa de dados sobre o assunto.
Tarefa: A partir da página principal pesquise se existe uma base de dados sobre o assunto, verifique se encontrou algo relevante sobre sua pesquisa.



















































3. Você quer escrever um artigo sobre “normas de associação de palavras da língua portuguesa” e precisa de dados sobre o assunto.
Tarefa: A partir da página principal pesquise se existe uma base de dados sobre o assunto, verifique se encontrou algo relevante sobre sua pesquisa.

Baseado nas 3 pesquisas já feitas descreva as dificuldades encontradas e como melhorar a pesquisa.

4. Sabendo que está disponível um arquivo de dados em formato “csv” sobre “Second Life” no repositório, agora é necessário encontrá-lo, baixá-lo e conferir se está no formato correto.
Tarefa: A partir da página principal encontre o arquivo indicado, inicie o download para verificar se o arquivo está no formato desejado. (basta verificar se o arquivo a ser baixado tem a extensão “csv”)
5. Você conseguiu encontrar os dados necessários à sua pesquisa, já baixou e vai usá-los, agora precisa citá-los no seu trabalho
Tarefa: Baixe o arquivo de dados de citação, no formato BibTeX, do dataset sobre “Second Life” da tarefa anterior.

B.6 Formulário

REPOSITÓRIO DE DADOS DE PESQUISA QUESTIONÁRIO DE SATISFAÇÃO DO PARTICIPANTE

Nº: _____ Participante: _____ Data: ____/____/____					
Área: <input type="radio"/> Exatas <input type="radio"/> Biológicas <input type="radio"/> Engenharias <input type="radio"/> Saúde <input type="radio"/> Agrárias <input type="radio"/> Sociais <input type="radio"/> Humanas <input type="radio"/> Letras					
Marque de acordo com seu nível de satisfação, o quanto você ficou satisfeito para cada quesito, ou "NA" (Não Aplicável) se não tem opinião formada sobre o assunto.					
Quanto ao leiaute em geral do sistema, o que você achou sobre?					
1. As cores, ícones e símbolos usados:					
Inadequados				Adequados	NA
					<input type="radio"/>
2. O tipo e o tamanho das fontes:					
Inadequados				Adequados	NA
					<input type="radio"/>
3. Os botões, quanto a seu posicionamento:					
Inadequados				Adequados	NA
					<input type="radio"/>
4. Os menus, quanto a seu posicionamento:					
Inadequados				Adequados	NA
					<input type="radio"/>
Quanto à funcionalidade dos componentes, são adequados, suficientes e efetivos?					
5. Os botões, quanto a suas funcionalidades:					
Inadequados				Adequados	NA
					<input type="radio"/>
6. Os menus, quanto a suas funcionalidades:					
Inadequados				Adequados	NA
					<input type="radio"/>
Quanto às informações presentes na interface (textuais, imagens, logos, etc.), o que você achou sobre?					
7. A quantidade e qualidade das informações:					
Inadequados				Adequados	NA
					<input type="radio"/>
8. A organização e posicionamento das informações:					
Inadequados				Adequados	NA
					<input type="radio"/>
Quanto às suas impressões de uso do sistema, o que você achou sobre?					
9. A primeira impressão ao ser apresentado ao sistema:					
Difícil				Fácil	NA
					<input type="radio"/>
10. A facilidade de aprender:					
Difícil				Fácil	NA
					<input type="radio"/>
<hr style="width: 20%; margin: 0 auto;"/> Assinatura do participante					