UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO CURSO DE BIBLIOTECONOMIA

Aline Macedo Brito

PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS: estratégias e iniciativas implementadas nas universidades da região Sul do Brasil

Aline Macedo Brito

PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS: estratégias e

iniciativas implementadas nas universidades da região Sul do Brasil

Trabalho de Conclusão de Curso, apresentado como requisito parcial para obtenção do título de Bacharel em Biblioteconomia, pela Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

Orientadora: Prof^a. Dr^a Ana Maria Mielniczuk de Moura

Co-orientadora: Daiane Barrili dos Santos

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Rui Vicente Oppermann

Vice-Reitor: Profa. Dra. Jane Fraga Tutikian

Faculdade de Biblioteconomia e Comunicação

Diretor: Profa. Dra. Karla Maria Müller

Vice-Diretor: Ilza Maria Tourinho Girardi

Departamento de Ciência da Informação

Chefe: Profa. Dra. Jeniffer Alves Cuty

Chefe-substituto: Profa. Dra. Eliane Lourdes da Silva Moro

CIP - Catalogação na Publicação

Brito, Aline Macedo

Preservação digital em repositórios institucionais: estratégias e iniciativas implementadas nas universidades da região Sul do Brasil / Aline Macedo Brito. -- 2017.

59 f.

Orientadora: Ana Maria Mielniczuk de Moura.

Coorientadora: Daiane Barrili dos Santos.

Trabalho de conclusão de curso (Graduação) -Universidade Federal do Rio Grande do Sul, Faculdade
de Biblioteconomia e Comunicação, Curso de
Biblioteconomia, Porto Alegre, BR-RS, 2017.

1. Preservação digital. 2. Repositório institucional. 3. Estratégias de preservação digital. I. Moura, Ana Maria Mielniczuk de, orient. II. Santos, Daiane Barrili dos, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Faculdade de Biblioteconomia e Comunicação

Rua Ramiro Barcelos, 2705 – Bairro Santana

CEP 90035-007 - Porto Alegre - RS -

Telefone: (51) 3308-5067

Fax: (51) 3308-5435

E-mail: fabico@ufrgs.br

Aline Macedo Brito

PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS: estratégias e

iniciativas implementadas nas universidades da região Sul do Brasil

Trabalho de Conclusão de Curso, apresentado como requisito parcial para obtenção do título de Bacharel em Biblioteconomia, pela Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

Orientadora: Prof^a. Dr^a Ana Maria Mielniczuk de

Moura

Co-orientadora: Daiane Barrili dos Santos

Aprovado em: de de
BANCA EXAMINADORA:
Dr ^a Caterina Marta Groposo Pavão
Universidade Federal do Rio Grande do Sul
Drof Dr. Dana Faustina Cabriel Junior

Prof. Dr. Rene Faustino Gabriel Junior Universidade Federal do Rio Grande do Sul

AGRADECIMENTOS

Agradeço a Deus pela vida.

Aos meus pais pelos ensinamentos.

Ao INSS pela licença capacitação, tão importante para finalizar essa etapa.

Aos professores da faculdade pelos ensinamentos ao longo dessa jornada.

À minha orientadora Ana Moura e co-orientadora Daiane dos Santos pelo direcionamento na escolha do tema, atenção e incentivo.

À Mileide por todo apoio no desenvolvimento do trabalho.

RESUMO

O presente trabalho objetiva identificar as estratégias e iniciativas de preservação digital implementadas nos repositórios institucionais das universidades da região Sul do Brasil a partir de um estudo descritivo e quantitativo. Destaca os benefícios advindos das novas tecnologias de comunicação e informação dentre eles o movimento do acesso livre e os repositórios institucionais. Compara brevemente a preservação tradicional com preservação digital, enfatizando a complexidade de se preservar documentos digitais, pois requerem o uso do computador que engloba a fragilidade dos suportes e obsolescência tecnológica. Questiona as soluções possíveis para se preservar em longo prazo, o patrimônio cultural registrado em meio digital. Detalha as principais estratégias de preservação digital encontradas na literatura dentre as quais: preservação tecnológica, migração, encapsulamento, emulação, ressaltando também a importância de políticas de preservação digital formalizadas. Comenta as iniciativas de preservação digital LOCKSS e Rede Cariniana. Trata-se de um estudo descritivo, cujo instrumento de coleta de dados foi um questionário enviado via Google docs para os responsáveis pelos repositórios institucionais da região Sul, totalizando 14 instituições participantes. Analisa os dados em forma de gráficos, quantitativamente. Conclui que parte dos repositórios está atento quanto à utilização das estratégias e iniciativas de preservação digital, porém a ausência de uma política de preservação digital formalizada é um fator que merece maior atenção, pois auxilia na padronização de critérios e procedimentos trazendo mais confiabilidade e segurança aos repositórios.

Palavras-chave: Preservação digital. Repositório institucional. Estratégias de preservação digital.

ABSTRACT

The present work aims to identify the strategies and initiatives of digital preservation implemented in the institutional repositories of the universities of the Southern region of Brazil based on a descriptive and quantitative study. It highlights the benefits of new communication and information technologies, including the free access movement and institutional repositories. It briefly compares traditional preservation with digital preservation, emphasizing the complexity of preserving digital documents, since they require the use of the computer that encompasses the fragility of the media and technological obsolescence. It questions possible solutions to preserve in the long term the cultural heritage registered in digital media. It details the main digital preservation strategies found in the literature, among them: technological preservation, migration, encapsulation, emulation, and the importance of formalized digital preservation policies. Discusses the digital preservation initiatives LOCKSS and Cariniana Network. It is a descriptive study, whose instrument of data collection was a questionnaire sent via Google docs to those responsible for the institutional repositories of the South region, totaling 14 participating institutions. Analyze the data in graph form, quantitatively. It concludes that part of the repositories is aware of the use of digital preservation strategies and initiatives, but the absence of a formalized digital preservation policy is a factor that deserves greater attention, since it helps to standardize criteria and procedures, bringing greater reliability and security to repositories.

Keywords: Digital preservation. Institutional repository. Digital preservation strategies.

LISTA DE FIGURAS

Figura 1– Estrutura do documento digital	17
Figura 2 – Diferença entre migração e emulação	20
Figura 3 – Modelo de atores entidades funcionais OAIS	23
Figura 4 – Funcionamento da rede centralizada e rede P2P	27
Figura 5 – Consulta avançada de instituições de ensino superior	34

LISTA DE GRÁFICOS

Gráfico 1– Localização geográfica	37
Gráfico 2 – Tamanho da coleção	37
Gráfico 3 – Existência de política formalizada	39
Gráfico 4 – Política formalizada	40
Gráfico 5 – Rede de preservação digital	41
Gráfico 6 – Estratégias de preservação digital	42
Gráfico 7 – Identificador persistente	43
Gráfico 8 – Extensões de arquivos de texto autorizados para inclusão no RI	43
Gráfico 9 – Adequação dos formatos de texto	44
Gráfico 10 – Extensões de arquivos de imagem autorizados para inclusão no RI	45
Gráfico 11 – Adequação dos formatos de imagem	46
Gráfico 12 - Extensões de arquivos de vídeo autorizados para inclusão no RI	46
Gráfico 13 – Resultados - Adequação dos formatos de vídeo	47

LISTA DE QUADROS

Quadro 1- Recomendações para utilização de metadados de preservação	22
Quadro 2– Formatos de texto - adequação para preservação	24
Quadro 3– Formatos de imagem - adequação para preservação	25
Quadro 4 – Formatos de vídeo - adequação para preservação	25
Quadro 5 – Universidades da região Sul	35

LISTA DE ABREVIATURAS

ASCII American Standard Code for Information Interchange

AUSJAL Associação de universidades confiadas a Companhia de Jesus na América

Latina

AVI Audio Vídeo Interleave

BMP Bitmap

CSS Cascading Style Sheets

DOI Digital Object identifier

DTD Document Type Definition

DVD Digital Video Disk

EPUB Formato de Publicação Eletrônica

FINEP Financiadora de Estudos e Projetos

GIF Graphics Interchange Format

HTML HyperText Markup Language

IBICT Instituto Brasileiro de Informação em Ciência e Tecnologia

ISO Internacional Organization for Standardization

JFIF JPEG File Interchange Format

JPG/JPEG Joint Photographic Experts Group

MPEG Moving Picture Expert Group

NISO National Information Standards Organization

ODT Documento de Texto Aberto

OpenDOAR Directory of Open Access Repositories

P2P Peer-to-peer

PD Preservação Digital

PDF Portable Document Format

PURL Persistent Uniform Resource Locators

RI Repositório Institucional

ROAR Registry of Open Access Repositories

RM Real Media File

RTF Rich Text Format

RV Real Vídeo Movie

SGML Standard Generalized Markup Language

TIFF Tagged Image File Format

TXT Extensão para arquivo de texto

UNESCO United Nations Educational, Scientific and Cultural Organization

URL Uniform Resource Locator

XHTML Linguagem Extensível para Marcação de Hipertexto

XML Extensible Markup Language

XMP Extensible Metadata Platform

SUMARIO

1	INTRODUÇÃO	13
1.1	PROBLEMA	14
1.2	OBJETIVOS	14
1.2.1	1 Objetivo Geral	14
1.2.2	2 Objetivos Específicos	14
1.3	JUSTIFICATIVA	14
2	PRESERVAÇÃO DIGITAL	16
2.1	ESTRATEGIAS DE PRESERVAÇÃO DIGITAL	18
2.2	INICIATIVAS DE PRESERVAÇÃO DIGITAL	26
3	REPOSITÓRIOS DIGITAIS	29
3.1	REPOSITORIOS TEMÁTICOS	31
3.2	REPOSITORIOS INSTITUCIONAIS	31
4	METODOLOGIA	33
4.1	QUANTO À NATUREZA	33
4.2	QUANTO À ABORDAGEM	33
4.3	QUANTO AO OBJETIVO	33
4.4	QUANTO AO PROCEDIMENTO	33
4.5	DELINEAMENTO DA PESQUISA	34
4.6	INSTRUMENTO DE COLETA DE DADOS	35
4.7	LIMITAÇÕES DA PESQUISA	36
5	ANÁLISE DOS DADOS	37
6	CONSIDERAÇÕES FINAIS	48
	REFERÊNCIAS	50
	APENDICE A - Questionário	57

1 INTRODUÇÃO

As bibliotecas estão presentes em nossa sociedade desde os tempos primórdios, evoluindo assim, juntamente com a humanidade. Com essa evolução, ela foi se aprimorando e buscando abarcar todos os diferentes papéis que a sociedade vem apresentando. As modificações ocorridas nos últimos anos nos meios de comunicação e o uso intensivo das tecnologias de informação fez surgir a era da informação digital, onde o fluxo e transmissão de dados se tornam mais dinâmicos e por consequência "[...] o peso do sistema econômico produtivo é cada vez mais centrado no fator informação". (LAZARTE, 2000, p. 45).

Das tábulas de argila, papiro e pergaminho, até o advento da Internet e das novas tecnologias de comunicação e informação, os suportes informacionais tradicionais estão aos poucos perdendo sua importância perante o crescimento significativo da produção da informação em meios digitais.

Diante desse cenário surgiu através da rede mundial de computadores uma nova forma de acesso à informação, tanto para uso do dia a dia como de caráter técnico científica. Do movimento de acesso livre nascem os Repositórios Institucionais (RIs) que organizam e disseminam a produção científica das universidades ou instituições de pesquisa.

E, se por um lado, a comunicação científica digital facilitou o livre e imediato acesso à informação, característica da era contemporânea, surge também o obstáculo da preservação das informações armazenadas em meio digital, pouco discutido pelo movimento do acesso livre, mas fundamentais para a garantia da memória institucional e que apresentam desafios de natureza complexa e fundamentalmente diferente se comparados com a preservação dos suportes tradicionais. Não basta os documentos estarem acessíveis sem restrições, há que se garantir que com o passar dos anos as informações científicas não se percam no tempo.

Dentre as iniciativas no Brasil a respeito da preservação digital está a publicação da Carta para a Preservação do Patrimônio Arquivístico Digital do Conselho Nacional de Arquivos (2004) baseada na proposta da Organização das Nações Unidas para a Educação, Ciência e Cultura (UNESCO), cujo conteúdo apresenta conceitos padronizados de preservação digital e o Seminário Internacional de Preservação Digital organizado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) que realizou sua terceira edição em 2017. Apesar de sua importância, pouco se tem discutido sobre o assunto em âmbito nacional para um programa de grande escala. O que ocorre são pesquisas para métodos em pequena escala, sem garantias efetivas de que tais conteúdos poderão ser acessados a longo prazo.

Nesse contexto, são abordadas quais estratégias e iniciativas de preservação digital (PD) identificadas na literatura da área são utilizadas para preservar o acervo dos repositórios institucionais da região Sul do Brasil.

1.1 PROBLEMA

Quais estratégias e iniciativas de PD são utilizadas nos repositórios institucionais das universidades da região Sul do Brasil?

1.2 OBJETIVOS

Nesta seção são apresentados o objetivo geral e os objetivos específicos desse trabalho.

1.2.1 Objetivo Geral

Identificar estratégias e iniciativas de preservação digital utilizadas nos repositórios institucionais das universidades da região Sul do Brasil.

1.2.2 Objetivos Específicos

- a) identificar na literatura da área as principais estratégias e iniciativas de preservação digital;
- b) identificar estratégias e iniciativas de preservação digital utilizadas nos repositórios institucionais das universidades da região Sul do Brasil;
- c) verificar a existência de uma política de preservação digital formalizada;
- d) analisar se os repositórios utilizam estratégias e iniciativas de preservação recomendados pela literatura da área.

1.3 JUSTIFICATIVA

Repositórios institucionais em comunidades acadêmicas contribuem para democratização da informação e avanço da ciência, potencializam as funções da comunicação científica maximizando a visibilidade das pesquisas e consequentemente o status da

instituição no meio acadêmico já que o conteúdo é disponibilizado online, podendo ser acessado em qualquer parte do mundo.

Com a "[...] web se tornando algo como uma única e imensa biblioteca digital [...]" (MARCONDES, 2005, p. 109) que de um lado possibilita fácil acesso a diversos tipos de informações, por outro lado traz a necessidade de se preservar e garantir acesso aos documentos para que estes não se percam na mesma velocidade em que são criados. Por esse motivo é fundamental estabelecer um planejamento ou política de preservação digital com estratégias e procedimentos que traga confiabilidade e segurança ao sistema. Nem sempre a preocupação com a preservação digital se inicia com a criação do sistema. Por ser uma área com carência de profissionais capacitados, a tendência é que o tema seja considerado relevante somente após perdas irreparáveis nos acervos.

A proposta de analisar quais estratégias e iniciativas de preservação digital são utilizadas nos repositórios de universidades da região Sul do Brasil (Paraná-PR, Santa Catarina-SC e Rio Grande do Sul-RS) tem sua relevância na contemporaneidade do tema, visto que as instituições de ensino superior cada vez mais têm demonstrado interesse em criar seus próprios repositórios. Através dos resultados de busca, constatou-se que é um tema pouco abordado se comparado à crescente quantidade de documentos disponíveis no mundo digital e que poderá juntamente com outros estudos já realizados, auxiliar os repositórios no planejamento de uma política de preservação digital ou até mesmo a revisão de procedimentos já utilizados. Do ponto de vista pessoal, a escolha do autor em realizar esse tema tem relação com sua formação em Arquivologia, onde o tema preservação digital também é bastante discutido nos últimos anos.

2 PRESERVAÇÃO DIGITAL

Desde o surgimento da escrita, o homem sentiu a necessidade de registrar fatos e acontecimentos e essa intenção sempre esteve associada à necessidade de lembrar (GONTIJO, 2004 apud CORRÊA, 2010). A partir daí, surgiu também a necessidade de preservar e conservar seus registros (nos mais variados suportes) das ações do tempo/natureza (desgaste natural dos suportes, sol, umidade, etc.) ou destruições provocadas pelo homem garantindo assim que novas gerações pudessem conhecer seus antepassados.

Segundo Conway (2001, p. 11), "[...] há muito tempo, as bibliotecas e os arquivos têm a responsabilidade de reunir, organizar e proteger a documentação da atividade humana". Ainda que a preservação do patrimônio digital tenha iniciado seus estudos na Arquivologia (FERREIRA, 2011) é possível também que o tema seja abordado pela Biblioteconomia considerando as particularidades de cada área, o que não exclui também a participação de outros especialistas em um programa de preservação digital.

Dos primeiros arquivos e bibliotecas até os tempos atuais, muitos foram os desafios encontrados com o que se refere à preservação dos acervos e conservação da memória da humanidade. O surgimento da internet e do microcomputador trouxeram novas formas de acesso, uso e compartilhamento de informações. Em consequência disso, um grande volume de informações passou a ser criada ou até mesmo convertida em formato digital, mesmo não havendo estratégias consolidadas que garantam o acesso a longo prazo (SAYÃO, 2005, p. 116).

Diferentemente de um livro em que a durabilidade da informação está diretamente ligada à integridade física do suporte, a preservação dos objetos digitais toma uma dimensão mais ampla "[...] onde a preocupação está na geração, gestão e acesso do arquivo eletrônico e no cumprimento das normas que garantem sua integridade intelectual". (TAVARES; MÁRDERO ARELLANO, 2015, p. 33). Dependentes de *hardwares* e *softwares*, devido a sua natureza, requerem o uso do computador tanto para criá-los como acessá-los. Em alguns casos, o nível de dependência é tanto que os arquivos só podem ser interpretados pelo *software* que os criou, caso contrário se tornam documentos criptografados através do seu próprio código (SAYÃO, 2005).

Conforme a Figura 1, o documento digital se caracteriza por 3 elementos básicos: *hardware*, *software* e a informação armazenada.

Documento digital

Hardware (físico) Software (lógico) Informação (suporte + bits)

Figura 1– Estrutura do documento digital

Fonte: Innarelli (2007, p. 42)

Na era do conhecimento, onde a informação é um elemento de valor econômico e sinônimo de poder, o que fazer para preservar os documentos que num lapso de tempo podem tornar-se obsoletos devido à volatilidade do meio digital e fragilidade dos suportes? Como salvaguardar nossa herança cultural, histórica, científica e econômica, expressa através da informação digital? (SAYÃO, 2005, p 116). Para Borba e Lima (2009, p. 17), "[...] a preservação digital tem a ver com a conservação e preservação do patrimônio cultural da humanidade, cuja atenção antes estava voltada apenas para os documentos em suportes físicos orgânicos, e que agora se aplica aos formatos de expressão digital".

Muito tem se discutido em como manter não somente a integridade dos documentos digitais, mas também a garantia de que esses documentos possam ser acessados ao longo de sua existência, mesmo diante das constantes mudanças de plataformas tecnológicas. Como refere Ferreira (2006, p. 20):

A preservação digital consiste na capacidade de garantir que a informação digital permaneça acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação. (FERREIRA, 2006, p. 20).

Parece haver um consenso na literatura de que não há uma solução definitiva para a preservação digital (SAYÃO, 2005), mas existem estratégias, procedimentos e tecnologias específicas que podem ser utilizadas a fim de minimizar os riscos da perda de arquivos digitais. Portanto, o sucesso da preservação digital depende de uma soma de esforços que envolvem políticas e ações institucionais que visam resguardar os documentos ao longo do tempo, principalmente da obsolescência tecnológica, mantendo-os íntegros e autênticos para que os pesquisadores confiem neles como fonte de informação e objeto de estudo, pois muitas vezes são vistos com desconfiança por estarem suscetíveis a alterações e manipulações.

A existência de uma política de preservação digital formalizada contribui para o gerenciamento das coleções digitais, pois consolida processo, procedimentos, estratégias e iniciativas de forma institucionalizada e também:

[...] documentam o compromisso de uma instituição com a conservação de conteúdo digital para uso futuro; especificam os formatos de arquivos a serem mantidos e o nível de preservação a ser providenciado; e documentam o compromisso de organizar o processo de preservação de acordo com padrões e melhores práticas que garantam uma custódia responsável. (CORRÊA, 2010, p. 23-24).

Márdero Arellano (2004, p. 25) destaca a existência de estudos que apontam a importância da implementação de políticas de preservação digital como:

[...] a forma mais efetiva de garantir o armazenamento e uso dos recursos de informação por longos períodos de tempo. A falta dessas políticas nos projetos de repositórios digitais sugere a carência de conhecimentos técnicos sobre a importância das estratégias de preservação digital existentes. (MÁRDERO ARELLANO, 2004, p. 25).

As atividades de preservação digital devem constituir políticas organizacionais que definam as estratégias de preservação de modo que se possa minimizar os efeitos da obsolescência tecnológica, bem como aumentar os níveis de segurança dos documentos digitais.

2.1 ESTRATEGIAS DE PRESERVAÇÃO DIGITAL

As estratégias contribuem para a preservação digital, porém não abarcam a garantia de acesso ao objeto digital como um todo. O que apontam são soluções específicas para problemas específicos "[...] capazes de compreender e recriar a forma original ou a função do objeto de forma que seja assegurada sua autenticidade e acessibilidade". (SAYÃO, 2005, p. 119).

A escolha das estratégias a serem utilizadas depende de cada caso, considerando os objetos a serem preservados e da disponibilidade orçamentária da instituição. Recomenda-se que as estratégias sejam utilizadas de forma combinada. De acordo com Sayão (2005, p. 134):

[...] há um consenso entre os especialistas da área de que a preservação digital pode ser realmente facilitada com a adoção de alguns procedimentos que incluem: 1) a definição de um conjunto limitado e gerenciável de padrões, preferencialmente abertos e/ou de ampla aceitação e de uso corrente: 2) a aplicação desses padrões na criação de novos objetos digitais, ou na conversão de documentos analógicos para formatos digitais; 3) o acompanhamento da obsolescência dos padrões desse conjunto e o monitoramento do surgimento de novos padrões; 4) a migração para os novos padrões tão logo eles estejam consolidados. (SAYÃO, 2005, p. 134).

Para Márdero Arellano (2008), Ferreira (2006), e Sayão (2010), as estratégias de preservação digital podem ser divididas em:

- a) **Preservação da tecnologia** consiste na preservação dos equipamentos tecnológicos tais como *software*, *hardware* e sistema operacional em que o objeto digital foi criado, permitindo assim o acesso em sua totalidade. Possui como desvantagens o alto custo operacional e de manutenção dos equipamentos, espaço físico, e o acesso restrito dos documentos para consulta local;
- b) Emulação são utilizados programas denominados "emuladores" que permitem imitar sistemas obsoletos em tecnologias mais avançadas, permitindo o acesso ao conteúdo em um ambiente que reproduz o sistema original (software e hardware).
 Uma vantagem dessa técnica é que não há perdas de conteúdo, porém possui como desvantagem a dificuldade de construir sistemas emulados diante da variedade de sistemas;
- c) Encapsulamento reunir juntamente ao objeto digital todos os elementos (metadados, software visualizador, arquivos específicos) em estruturas físicas ou lógicas (software e hardware) necessários para seu acesso futuro. Essa estratégia é recomendada para documentos que não necessitam acesso por longo período de tempo;
- d) **Migração** é realização de transferência periódica de arquivos de uma tecnologia para um patamar tecnológico mais avançado, o que pode envolver suporte, *software* e *hardware*. É muito utilizada para acervos de grande porte. Pode ser executada de forma esporádica ou rotineira, automaticamente ou por controladores do sistema. Algumas instituições evitam essa estratégia pelo fato de que algumas propriedades dos objetos digitais podem ser modificadas durante o processo. Pode ser dividida em migração a pedido, migração distribuída, atualização de versão, conversão para formatos concorrentes e normalização. Pode ser combinada com o refrescamento que é a transferência do conteúdo de um suporte que está em vias de

se tornar obsoleto para um atual. Ex.: DVD para Pen-Drive (refrescamento) + formato padronizado MSWord para XML, ASCII ou PDF/A (migração).

Na Figura 2 podemos perceber a diferença entre a migração e a emulação com relação às propriedades de um objeto digital. Observa-se que na migração algumas características originais se corrompem, porém ainda é possível interpretar o documento, enquanto que na emulação o documento original permanece intacto.

Figura 2- Diferença entre migração e emulação

Fonte: Adaptado de Keep ([2011?] apud DIAS, 2011)

As cópias de segurança (backups) são consideradas essenciais na preservação digital, como meio de assegurar a restauração de dados em caso de pane nos equipamentos ou sistemas. Trata-se de uma estratégia bastante difundida por ser um procedimento simples e de baixo custo. Innarelli (2012, p. 46) comenta que a política de backup é de suma importância para a preservação digital, pois "[...] visa garantir de forma íntegra e confiável a restauração de qualquer tipo de dados registrados no sistema de informação".

O conceito mais simplista de metadados é que são dados sobre dados, porém podem variar de comunidade para comunidade, mas tem como objetivo principal a recuperação e preservação na área de informação digital. A *National Information Standards Organization* (NISO) define metadados como a informação estruturada que descreve, explica, localiza ou possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar (NISO, 2004).

De acordo com Saramago (2004, p. 1), metadados para preservação são definidos "[...] como informação de apoio aos processos associados com a preservação digital de longo

prazo". Devem acompanhar todo o ciclo de vida do objeto digital preservando sua cadeia de bits, contendo:

[...] informação técnica e administrativa sobre decisões e acções de preservação, registar os efeitos das estratégias de conversão de dados, assegurar a autenticidade dos recursos digitais ao longo do tempo, registrar informação acerca de gestão de coleções e de direitos e ainda fornecer informação acerca dos próprios metadados. (SARAMAGO, 2004, p. 1).

Borba e Lima (2009, p. 13) complementam que: "Os metadados de preservação são informações estruturadas sobre um objeto digital que indica a necessidade de manter e proteger os dados do documento, registrar a história do objeto, sua identidade, integridade, com o intuito de garantir sua autenticidade e preservação", mesmo diante das mudanças tecnológicas.

Sayão (2010) dividiu metadados em 3 tipos: a) descritivos, b) administrativos, c) estruturais. Os metadados descritivos estão relacionados à recuperação da informação, tais como título, autor. Os metadados administrativos são considerados os mais importantes para a preservação uma vez que registram atos de gestão ao longo do tempo: criação, especificidades e dependência técnica. Os metadados estruturais complementam a informação administrativa acrescentando dados tecnológicos indispensáveis à recuperação dos objetos.

Com relação aos metadados de descrição, provavelmente o *Dublin Core Metadata Initiative* (DCMI) é a iniciativa mundial mais conhecida/utilizada principalmente para documentos textuais (esquema de metadados XMP) em repositórios digitais (PAVÃO et al. 2015, p. 106). A aplicação de alguns desses metadados visa auxiliar na preservação digital, conforme podemos visualizar no Quadro 1.

Quadro 1– Recomendações para utilização de metadados de preservação

Nome do termo	Definição
Proveniência	Informação sobre modificações na posse ou custódia do trabalho desde sua criação. Estas informações são significantes para sua autenticidade, integridade e interpretação.
Modificado	Data cujo recurso foi alterado ou modificado.
Formato Original	Quando o recurso descrito em seu formato original é o mesmo conteúdo intelectual do recurso referenciado, mas apresentado em outro formato.
Disponibilidade	Data que o trabalho começou a ser disponibilizado ou começará. Data de alimentação no sistema.
Descrição	Descrição do trabalho, que deve incluir informações adicionais ao documento, como estado físico.
Formato	Formato do arquivo, da mídia física. (PDF, CD, DVD, 300KB, 20MB).
Extensão	Tamanho do arquivo como o nº. de páginas, duração, em caso de áudio, vídeo

Fonte: Dublin Core ([200-?] apud BORBA, 2009, p. 16)

O Modelo de Referência Open Archival Information System (OAIS) é uma norma ISO (14721:2003) que pode ser aplicada também como estratégia no intuito de padronizar atividades de preservação digital, já que atua como um esquema conceitual com terminologia e conceitos relevantes que descrevem o ambiente, os componentes funcionais e os objetos informacionais associados com um sistema responsável pela preservação de longo prazo de arquivos digitais.

Castro et al. (2010, p. 287) consideram o modelo OAIS como referência para a preservação digital pois tem o objetivo de:

[...] suportar e garantir que as características mais relevantes do arquivamento digital como a disponibilidade, durabilidade e confiabilidade dos dados sejam mantidas e consideradas pelos sistemas de preservação digital, por meio de padrões que permitem a manutenção, o compartilhamento e a distribuição do material preservado. (CASTRO et al., 2010, p. 287).

O modelo é formado por seis entidades funcionais que interagem com três atores conforme ilustra a Figura 3. Cada ator possui uma atribuição específica: produtor (producer) é responsável pela submissão da informação no sistema, a administração (management) estabelece políticas de preservação e monitora o sistema e o consumidor (consumer) é o usuário que acessa o sistema em busca de informações. A entidade produção é responsável pela recepção dos pacotes de submissão da informação *Submission Information Package* (SIP) e por gerar pacotes de arquivamento de informação *Archival Information Package* (AIP)

e enviar para armazenamento. Informações descritivas são extraídas dos documentos para sua futura recuperação e enviadas para o gerenciamento de dados que fornece as funcionalidades necessárias para criar e manter base de dados de metadados descritivos utilizados na busca da informação armazenada. Por fim cabe a entidade acesso fornecer suporte aos consumidores para que localizem a informação desejada que será entregue em forma de *Dissemination Information Package* (DIP). A entidade plano de preservação supervisiona e monitora o ambiente OAIS para que as informações permaneçam acessíveis por um longo prazo e a administração gerencia o sistema no âmbito operacional.

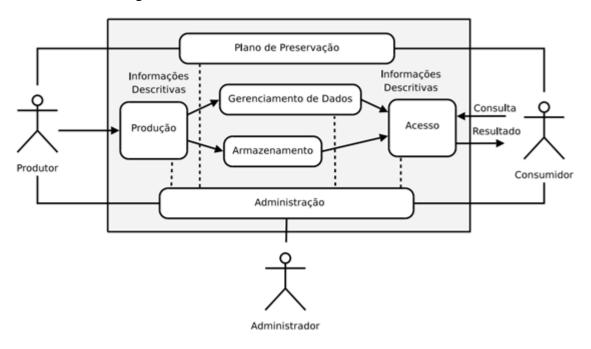


Figura 3– Modelo de atores entidades funcionais OAIS

Fonte: Seara (2008)

A normalização contribui com a aplicação de estratégias de preservação digital, pois permite o controle dos formatos dos arquivos a serem armazenados em um repositório. Quanto menor o número de formatos a serem utilizados, mais prático será o gerenciamento de grandes coleções. Sempre que possível deverão ser escolhidos os formatos abertos que são interoperáveis entre sistemas distintos (FERREIRA, 2006) e apresentam maior probabilidade de preservação conforme demonstram os quadros de 2 a 4. O repositório poderá estabelecer em sua política quais formatos de arquivo serão aceitos ou até mesmo realizar a conversão, pois:

O objeto digital que faz parte de um RI precisa ter a garantia de sua preservação em longo prazo, sem perda de conteúdo ou modificação na sua apresentação. Para tanto, é necessário que se estabeleçam critérios durante o processo de povoamento desses objetos, com a escolha de formatos que possibilitem um bom grau de adequação às práticas de preservação. (WEITZEL; MESQUITA, 2015, p. 192).

Munhoz (2011) aponta os principais formatos considerando a natureza do objeto digital, assim estruturado:

Quadro 2- Formatos de texto - adequação para preservação

Adequação	Extensão	Formato
	HTML, XHTML ou XML	Contendo referência para o DTD ou esquema de validação de formato
Alta	TXT	Textos planos usando codificação de caracteres UTF-8 USASCII ou UTF-16 com Byte Order Mark
7 Httu	PDF	PDF/A -1 segundo a norma ISSO 19005-1
	ODT	Open Document Text
	HTML, XHTML ou XML	Sem referência para o DTD ou esquema de validação de formato
	CSS	Cascading Style Sheets
	TXT	Textos plano com codificação de caracteres em ISO-8859-1
Média	PDF	Com fontes embebidas
1,10010	RTF	Rich Text Format 1.x
	HTML	HTML 4.x incluindo declaração de DOCTYPE
	SXW	Open Office Text Document
	DOCX	Office Open XML
	DTD	Document Type Definition
	SGML	Standard Generalized Markup Language
	DOC	Microsoft Word
Baixa	PS	Postscript
Dunu	PDF	PDF cifrado
	DVI	Device Independent File Format

Fonte: Adaptado de Munhoz (2011, p. 29-30)

Quadro 3- Formatos de imagem - adequação para preservação

Adequação	Extensão	Formato
	PNG	Portable Network Graphics
	JP2	JPEG2000 usando a compressão sem perda de informação
Alta	TIFF	sem compressão
	SVG	Scalable Vector Graphics
	JP2	*JPEG2000 usando a compressão com perda de informações
	GIF	Graphics Interchange Format
	JPG	JPEG/JFIF
	TIFF	com compressão
Média	ВМР	Bitmap
	DNG	Digital Negative
	CGM	Computer Graphic Metafile e WebCGM
	SWF	Macromedia Flash
	PSD	PhotoShop
	JPF, JPX	JPEG 2000 Part2
Baixa	SID	MrSID
Duma	TIFF	em formato planar
	FPX	FlashPix
	RAW	RAW (a extensão pode variar de acordo com o dispositivo de captura)
	EPS	Encapsulated Postscript

Fonte: Adaptado de Munhoz (2011, p. 30-31)

Quadro 4 – Formatos de vídeo – adequação para preservação

Adequação	Extensão	Formato
	MOV	QuickTime Movie sem compressão
	AVI	AVI sem compressão
	MJ2	Motion JPEG 2000 (ISO/IEC 15444-4)
Alta	AVI, MOV	Motion JPEG
	MPG, MPEG	MPEG-1,MPEG-2
Média	MP4	MPEG-4
1110010	OGG, OGM, OGV	Ogg Theora
	WMV	Windows Media Video
Baixa	AVI	AVI com compressão
Buixu	MOV	JP QuickTime Movie com compressão
	RV, RM	RealNetworks 'Real Video'

Fonte: Adaptado de Munhoz (2011, p. 32)

O Portable Document Format (PDF) é um padrão aberto, conhecido como ISO 32000, mantido pela Internacional Organization for Standardization. O PDF/A é uma extensão de arquivo preservável a longo prazo, recomendado como arquivo padrão nos repositórios pois não depende de software/hardware específico para funcionar e é capaz de "[...] armazenar no próprio documento tudo que é necessário para visualizar e imprimir. Utiliza metadados Extensible Metadata Platform (XMP)" (WENSING, 2010, p. 58). A linguagem XML, juntamente com o PDF, contribui para preservação a longo prazo, favorece a interoperabilidade entre sistemas, podendo ser utilizada também como formato de extensão de arquivo (MÁRDERO ARELLANO, 2008).

Identificadores persistentes ou permanentes também são considerados um "[...] componente importante da infraestrutura dos repositórios digitais" (MÁRDERO ARELLANO, 2008, p. 114), pois permitem que o objeto digital permaneça disponível, independentemente de sua localização, "quebra" ou mudança de URL, através da criação de um endereço fixo (link) que tem por objetivo preservar acesso à informação por um longo período, mesmo que esse arquivo seja movido, sua propriedade transferida ou até mesmo que a organização não exista mais quando a informação for acessada (SAYÃO, 2007, p. 67). Em uma linguagem simples, pode-se comparar a URL a um endereço residencial e o link persistente ao CPF que é único e inalterável.

A estabilidade dos links preserva também a integridade dos recursos digitais, uma vez que esses recursos são constituídos também pelos seus relacionamentos com outros objetos digitais através de links acionáveis, como por exemplo, as referências bibliográficas que citam outras publicações com seus respectivos links.

Dentre as organizações que oferecem identificadores persistentes podemos citar: URN da *Internet Engineering Task Force* (IETF), PURL da *Online Computer Library Center* (OCLC), HANDLE da *Corporate for National Research Initiatives* (CNRI), DOI da *International DOI Fundation* (IDF).

2.2 INICIATIVAS DE PRESERVAÇÃO DIGITAL

São conjuntos de ações voltadas à preservação digital que complementam as estratégias a serem utilizadas. Um modelo de preservação considerado bastante promissor é a preservação digital distribuída (ver Figura 4), em que cópias dos objetos digitais são distribuídas em locais geograficamente distintos em ambiente tecnológico padronizado sem a necessidade de um supercomputador como servidor central, considerada de baixo custo para

os participantes e com a vantagem de que estes podem sair da rede a qualquer momento, caso desejarem.

Skinner e Schultz (2010, tradução nossa) sugerem princípios essenciais a serem seguidos em um modelo de preservação digital distribuída:

- a) cópias dos conteúdos devem ser mantidas distantes uma das outras, num raio de 120
 a 200 km;
- b) não selecionar locais de preservação que estejam propensos a catástrofes naturais;
- c) os locais de preservação devem ser distribuídos em diferentes redes de energia;
- d) os locais de preservação devem ser controlados por diferentes administradores de sistemas;
- e) as cópias dos conteúdos devem ser verificadas regularmente (auditoria) a fim de possibilitar a reparação do conteúdo, caso estejam corrompidos;
- f) o conteúdo deve ser replicado pelo menos três vezes com o intuito de garantir a segurança das cópias, de acordo com o nível de importância do conteúdo.

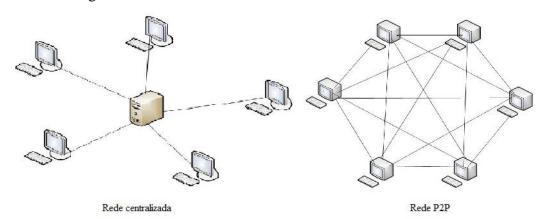


Figura 4 - Funcionamento da rede centralizada e rede P2P

Fonte: Valadão (2008)

Uma rede mundialmente conhecida é o sistema LOCKSS (sigla em inglês para "muitas cópias mantem as coisas seguras"), foi criado em 1998 pela Universidade de Stanford como um projeto com objetivo de auxiliar na preservação digital de periódicos acadêmicos. Com o passar do tempo o projeto foi se aprimorando e em 2004 se consolidou como uma iniciativa de preservação digital colaborativa de âmbito internacional. Utiliza *software* de código aberto que permite a criação de redes privadas locais com intuito de preservar documentos de dados digitais atuando como uma rede de preservação digital distribuída, garantindo a integridade dos acervos de seus associados. Para a formação da rede privada

LOCKSS são necessárias no mínimo 06 "caixas", isto é, 06 instituições/editores interligados em diferentes localidades, de modo que as instituições que formam essa rede colaborativa não podem descartar os registros (LOCKSS, 2014, tradução nossa).

No âmbito nacional apresenta-se a Rede Cariniana que tem por objetivo "[...] propor a construção e customização de uma rede nacional de serviços de preservação digital." (MÁRDERO ARELLANO, 2012, p. 87). Coordenada pelo IBICT, iniciou suas atividades em 2012 com a responsabilidade de inserir e armazenar os periódicos eletrônicos das universidades parceiras no sistema LOCKSS, com a colaboração da *Stanford University*, *University of Edinburgh* e *Harvard University*. Oferece também "[...] apoio técnico, capacitação e assessoramento na negociação das permissões de arquivamento das publicações e construção de futuras sub-redes locais de preservação digital." (MÁRDERO ARELLANO, 2012, p. 88), sendo considerada uma experiência única no país, na temática de preservação digital (TAVARES; MÁRDERO ARELLANO, 2015).

3 REPOSITÓRIOS DIGITAIS

Os avanços tecnológicos viabilizaram novas formas de comunicação e acesso à informação em formato eletrônico. O processo de comunicação técnico científica se modificou com a chegada da internet e da web possibilitando que as publicações fossem disponibilizadas em larga escala nas bases de dados online. Começa a se desenvolver uma sociedade regida pela informação, conceituada por Castells (1999) de sociedade em rede que pode ser definida como:

[...] uma estrutura social baseada em redes operadas por tecnologias de comunicação e informação fundamentadas na microelectrônica e em redes digitais de computadores que geram, processam e distribuem informação a partir de conhecimento acumulado nos nós dessas redes. (CASTELLS, 1999, p. 20).

Com a crise dos periódicos na década de 80, em que as bibliotecas não conseguiram manter suas assinaturas devido ao alto custo, a academia começou a discutir novas possibilidades de acesso e disseminação do conhecimento científico sem a participação das editoras e avaliadores, ou seja, uma forma de acesso mais "democrático". Guedes (2010) considera que a crise dos periódicos e a chegada da internet foram determinantes para o surgimento do acesso livre ao conhecimento e a informação.

Desde o encontro pioneiro conhecido como *Budapest Open Acess Initiative* em 2001 que preconiza os princípios e estratégias para implantação do acesso aberto à informação científica, o movimento vem sendo aderido por diversos países, o que demonstra seu fortalecimento. Outros dois documentos importantes surgiram para o movimento de acesso aberto: a Declaração de Bethesda (abril 2003) e a Declaração de Berlim (outubro de 2003).

No Brasil, o IBICT publica em 2005 o manifesto brasileiro de apoio ao acesso livre à informação científica tornando-se o primeiro desdobramento político de um movimento internacional amplo de apoio ao livre acesso à informação científica. As instituições de ensino são incentivadas a implementar repositórios com iniciativas como: a Declaração de Salvador (2005), a Carta de São Paulo (2005), a Declaração de Florianópolis (2006). Em 2009, é lançado o edital FINEP/PCAL/XBDB 003/2009, uma parceria entre Financiadora de Estudos e Projetos (FINEP) e o IBICT, com o de objetivo apoiar projetos de implementação de repositórios, fornecendo kits tecnológicos e infraestrutura necessária às instituições de ensino e pesquisa.

A partir de Budapeste, houve a criação do primeiro Repositório do Instituto de Tecnologia de Massachusetts (MIT), juntamente com a disponibilização do Dspace. Em 2004, foi lançada a versão brasileira do *software* pelo IBICT.

Os repositórios digitais surgem da "[...] aplicação da internet e de tecnologias emergentes de informação e comunicação no contexto da comunicação científica somada ao [...]" (LEITE, 2006, p. 84) movimento mundial de acesso livre, *Open Access* (OA) e à produção científica em ambiente digital. O movimento é uma via alternativa ao sistema de publicação tradicional por assinaturas, muitas vezes de alto custo para uma instituição. O acesso livre preconiza além do acesso gratuito na internet de literatura de caráter acadêmico ou científico, a permissão a qualquer usuário ler, descarregar, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral dos documentos (PAVÃO, 2010).

Para King (2010, p. 2), "[...] o livre acesso a toda a literatura científica não tem apenas valor substancial para os investigadores financiados por órgãos federais enquanto leitores, mas para outros cientistas, além de ser um ganho para a ciência e suas contribuições para economia e para a sociedade."

Para que o acesso aberto fosse difundido internacionalmente foram recomendadas duas estratégias de povoamento dos repositórios:

- a) via verde os próprios autores podem arquivar sua produção científica nos repositórios também chamado de autoarquivamento com a autorização (sinal verde) dos editores;
- b) via dourada disponibilização do acesso aberto diretamente pelos periódicos científicos/editor, sem restrições.

A implementação de uma política que resguarde o direito do autor sobre a sua obra, a divulgação da relevância do movimento aberto, bem como tutoriais de como proceder ao autoarquivamento nos repositórios, contribui para seu povoamento, pois em muitos casos, tais procedimentos não são de conhecimento dos autores.

São resultado da Iniciativa dos Arquivos Abertos onde os resultados das pesquisas financiadas por agências públicas de fomento são disponibilizados de forma online e gratuita, utilizando-se de *softwares* desenvolvidos pela Iniciativa de Código Aberto (ICA) – do inglês *Open Source Initiative* – que é uma iniciativa de desenvolvimento de "[...] *software* que pode ser utilizado livremente, alterado e compartilhado (de forma modificada ou não) por qualquer pessoa" (OPEN SOURCE INITIATIVE, 2014, documento on-line). Atualmente, o *software* livre mais utilizado em RIs é o Dspace, seguido do E-prints e Bepress (ROAR, 2017). Podemos citar também Fedora, Arno, CDSware, iTor, MyCoRe, DiVA.

Autores como Weitzel (2006, p. 53), Ribeiro e Vidotti (2009, p. 109), dividem os repositórios em temáticos e institucionais. Leite (2009, p. 19) divide os repositórios digitais em três tipos e os define:

- a) Repositórios institucionais: voltados à produção intelectual de uma instituição, especialmente universidades e institutos de pesquisa;
- b) Repositórios temáticos ou disciplinares: voltados a comunidades científicas específicas. Tratam, portanto, da produção intelectual de áreas do conhecimento em particular;
- c) Repositórios de teses e dissertações (Electronic Theses and Dissertation ETDs): repositórios que lidam exclusivamente com teses e dissertações. Muitas vezes a coleta dos documentos é centralizada por um agregador. (LEITE, 2009, p. 19).

3.1 REPOSITORIOS TEMÁTICOS

Os repositórios temáticos reúnem publicações de uma determinada área do conhecimento com acesso online. Dentre eles podemos citar o Sistema Brasileiro de Informação do Café (SBICafé), da Universidade Federal de Viçosa (UFV), onde podemos encontrar documentos relacionados à temática cafeeira, o EPrints in *Library and Information Science* (E-LIS) na área de Biblioteconomia e Ciência da Informação e ArXiv nas áreas de Física, Matemática, Computação, Estatística e Biologia.

Café et al. (2003) considera que o repositório institucional, de caráter multidisciplinar é a união de todos repositórios temáticos dos departamentos de uma universidade, onde cada departamento desenvolve pesquisas relativas ao seu campo específico de estudo.

3.2 REPOSITORIOS INSTITUCIONAIS

Os repositórios institucionais foram criados com o objetivo de reunir a produção científica de uma instituição de pesquisa ou de ensino superior em um só espaço, sendo uma fonte confiável de pesquisa, haja vista que seu conteúdo é gerido e controlado pela instituição. Apresentam como características as "[...] facilidades de importar, identificar, armazenar, preservar, recuperar e exportar um conjunto de documentos digitais, por intermédio da web. Os documentos são identificados com metadados que facilitam sua recuperação [...]" (PAVÃO, 2010, p. 52), podendo ser localizados em sites de buscas tais como Google, Google Scholar, Yahoo, etc.

Os repositórios institucionais são, segundo Santos Júnior (2010, p. 36), "[...] coleções digitais que armazenam, preservam, divulgam e dão acesso à produção intelectual de uma

determinada comunidade, onde os documentos são disponibilizados irrestritamente na web." Promovem maior visibilidade institucional, servindo como indicador de qualidade e relevância científica, permitindo avaliar o impacto das pesquisas através dos dados estatísticos: a quantidade de acessos e downloads, identificando inclusive os países em que determinado documento foi visualizado ou feito download. Outro fator importante é a contribuição dos repositórios para a construção da memória institucional, atuando como uma combinação de arquivos e bibliotecas do ponto de vista conceitual (DODEBEI, 2009).

Nas universidades geralmente são divididos em comunidades e subcomunidades, baseado na estrutura hierárquica da instituição e seu conteúdo é dividido em coleções. Mueller (2006) afirma que a implantação de repositórios institucionais (RIs) em universidades é visto como um indicador importante da aceitação e expansão do movimento aberto que defende a disponibilidade de acesso irrestrito ao conteúdo científico resultante das pesquisas sobretudo financiadas com recursos públicos.

De acordo com Costa e Leite (2010), todo repositório institucional é um tipo de biblioteca digital, mas nem toda biblioteca digital pode ser considerada um repositório institucional, pois parte do princípio que duas condições devem ser estabelecidas para que um repositório seja institucional:

A primeira é que sejam oficialmente reconhecidos pela instituição, por meio da implementação de políticas de depósito compulsório e outras que garantam sua existência. A segunda é que seus conteúdos cubram a maior parte das áreas de ensino e pesquisa da instituição, como demonstração de seu reconhecimento e aceitação. (COSTA; LEITE 2010, p. 164).

De acordo com Santos Junior (2010), instituições não acadêmicas/científicas não possuem repositórios institucionais, pois não lidam com informação científica, nesse caso, seus repositórios são considerados bibliotecas digitais¹. Para Kuramoto (2015) o conceito de repositório institucional também está ligado à produção científica de instituições de ensino e pesquisa.

-

¹Informações obtidas a partir de entrevista com o especialista em repositórios institucionais Fernando César Lima Leite.

4 METODOLOGIA

Para atingir o objetivo desse trabalho, foram escolhidos os procedimentos metodológicos especificados abaixo:

4.1 QUANTO À NATUREZA

O método utilizado no presente trabalho é de natureza básica, pois objetiva gerar novos conhecimentos, porém sem aplicação prática prevista (MORESI, 2003, p. 8).

4.2 QUANTO À ABORDAGEM

A abordagem da pesquisa é quantitativa uma vez que permite descobrir quantos repositórios "[...] compartilham de uma característica ou um grupo de características [...]" (MORESI, 2003, p. 64), trabalhando com dados precisos e confiáveis onde se busca quantificar resultados da amostra da população alvo da pesquisa tais como quantos repositórios utilizam determinada estratégia de preservação digital e até mesmo quais são as mais utilizadas. Para Fonseca (2002) a pesquisa quantitativa está centrada na objetividade, empregando a linguagem matemática para descrever as causas de determinado(s) fenômeno(s). A amostra é não probabilística, pois os critérios e julgamentos foram definidos pelo autor (HONORATO, 2004, p. 105).

4.3 QUANTO AO OBJETIVO

Pesquisas descritivas, segundo Gil (2002, p. 47), "[...] tem como objetivo primordial a descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis", que no presente estudo tem como objeto os repositórios institucionais da região Sul do Brasil.

4.4 QUANTO AO PROCEDIMENTO

A técnica empregada no estudo foi o levantamento, que tem por característica a "[...] interrogação direta das pessoas cujo comportamento se deseja conhecer" (GIL, 2002, p. 50), mediante a ferramenta do questionário para a coleta de dados.

4.5 DELINEAMENTO DA PESQUISA

A seleção da amostra foi realizada através de uma consulta avançada de instituições de ensino superior no site emec.mec.gov.br (conforme a Figura 5), utilizando-se os filtros por Unidade da Federação: Paraná, Santa Catarina, Rio Grande do Sul, por categoria administrativa: pública federal, pública estadual, privada com fins lucrativos, privada sem fins lucrativos e por organização acadêmica universidades com credenciamento presencial, situação ativa.

Buscar por:

Instituição de Ensino Superior

Curso de Graduação

Curso de Especialização

Nome ou Sigla da Instituição:

UF:
Rio Grande do Sul

Município:

Selecione...

Pública Federal

Pública Estadual

Privada sem fins lucrativos

Privada com fins lucrativos

Privada com fins lucrativos

Categoria Administrativa:

Presencial - Superior

EAD - Superior

EAD - Lato-sensu

Escola de Governo

Situação:

Ativa

Código de verificação:

Digite o código

Trocar imagem

Pesquisar

Figura 5 – Consulta avançada de instituições de ensino superior

Fonte: e-MEC (2017, documento on-line)

Do resultado da pesquisa foram encontradas 57 instituições assim dispostas: 24 localizadas no Paraná, 12 em Santa Catarina e 21 no Rio Grande do Sul, sendo uma instituição com sede em SC, mas com campus no RS e PR. A partir desse resultado, o próximo passo foi averiguar, através de buscas no Google¹, Open Doar², IBICT³ e sites das instituições, quais possuem repositórios em funcionamento.

Finalizada a etapa, totalizaram 24 instituições que possuem RIs, consideradas, portanto, os sujeitos do presente trabalho: 6 instituições no PR, 5 instituições em SC e 13

² http://www.opendoar.org/find.php

¹ www.google.com.br

³ http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais/repositorios-brasileiros

instituições no RS. O contato foi feito via correio eletrônico, mediante carta de apresentação (APÊNDICE A), ao e-mail de contato informado no site dos repositórios institucionais, ou e-mail prévio ao "fale conosco" da universidade ou biblioteca, solicitando contato do gestor do repositório para os que não havia tal informação. Após 3 tentativas de envio, dos 24 repositórios institucionais, 14 responderam ao questionário, validando este trabalho. No quadro 5 estão listadas as 24 universidades da região sul que possuem repositórios. As instituições participantes da pesquisa estão identificadas por asterisco.

Quadro 5 – Universidades da região Sul

Instituições	Categoria
Paraná	
Universidade Estadual de Ponta Grossa – UEPG	Pública
Universidade Estadual do Centro Oeste – UNICENTRO	Pública
Universidade Federal da Integração Latino-Americana – UNILA*	Pública
Universidade Federal do Paraná – UFPR*	Pública
Universidade Pitágoras – UNOPAR	Privada
Universidade Tecnológica Federal do Paraná UTFPR	Pública
Santa Catarina	
Universidade Alto Vale do Rio do Peixe – UNIARP	Privada
Universidade do Extremo Sul Catarinense – UNESC*	Privada
Universidade do Sul de Santa Catarina – UNISUL	Privada
Universidade Federal da Fronteira Sul – UFFS*	Pública
Universidade Federal de Santa Catarina – UFSC*	Pública
Rio Grande do Sul	
Fundação Universidade Federal do Pampa – UNIPAMPA*	Pública
Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS*	Privada
Universidade de Caxias do Sul – UCS*	Privada
Universidade de Passo Fundo – UPF	Privada
Universidade de Santa Cruz do Sul – UNISC	Privada
Universidade Federal de Ciências da Saúde de Porto Alegre - UFCSPA	Pública
Universidade do Vale do Rio dos Sinos – UNISINOS*	Privada
Universidade do Vale do Taquari – UNIVATES	Privada
Universidade Federal de Pelotas – UFPel*	Pública
Universidade Federal de Santa Maria – UFSM*	Pública
Universidade Federal do Rio Grande – FURG*	Pública
Universidade Federal do Rio Grande do Sul UFRGS*	Pública
Universidade Regional do Noroeste do Estado do Rio Grande do Sul – UNIJUÍ*	Privada

Fonte: Elaborado pelo autor

4.6 INSTRUMENTO DE COLETA DE DADOS

A coleta de dados, segundo Marconi e Lakatos (2007, p. 167), é a "[...] etapa da pesquisa em que se inicia a aplicação dos instrumentos elaborados e das técnicas

selecionadas." O instrumento de coleta de dados a ser utilizado deve estar de acordo com o objeto a ser estudado, sendo escolhido para essa pesquisa o questionário que consiste em "[...] uma série ordenada de perguntas, que devem ser respondidas por escrito e sem a presença do entrevistador." (LOPES, 2006, p. 241). Cabe destacar que um questionário deve ser elaborado com questões organizadas de forma lógica, clara, evitando perguntas irrelevantes, complexas ou extensas facilitando o entendimento do público-alvo.

Laville e Dionne (1999, p. 183-184) destacam que:

Dentre as vantagens desse tipo de questionário padronizado – diz-se também uniformizado- pode-se, lembrar que se mostra econômico no uso e permite alcançar rápida e simultaneamente um grande número de pessoas, uma vez que elas respondem sem que seja necessário enviar-lhes um entrevistador. (LAVILLE; DIONNE, 1999, p. 183-184).

O questionário foi desenvolvido em formato eletrônico no *Google.docs*, que permite o envio automático das repostas ao pesquisador, contendo 12 perguntas fechadas que se vinculam à abordagem quantitativa baseada na concepção de uma realidade objetiva e mensurável que tem suas origens no positivismo. A aplicação ocorreu entre os dias 30 de outubro a 17 novembro de 2017. Os participantes tiveram acesso à pesquisa através de um link enviado ao e-mail dos repositórios.

Após a coleta, os dados foram tabulados em forma de gráficos com o intuito de visualizar melhor as informações e analisados quantitativamente.

4.7 LIMITAÇÕES DA PESQUISA

Como limitações deste estudo podemos mencionar:

- a) dificuldade em encontrar o endereço do repositório institucional no site da instituição;
- b) dificuldade em contatar os gestores dos RIs, visto que o e-mail para contato não era informado na página ou quando informado era inexistente;
- c) a não obtenção do retorno de todos os repositórios institucionais das universidades da região Sul.

5 ANÁLISE DOS DADOS

Os dados apresentados a seguir foram obtidos através das repostas das 14 instituições participantes. As questões 1 e 2 referem-se à localização geográfica dos repositórios, distribuídos da seguinte forma:

Localização Geográfica

2;14%

9;64%

PR

SC

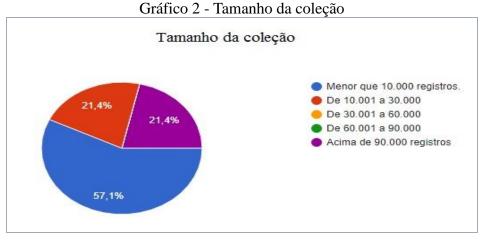
RS

Gráfico 1- Localização geográfica

Fonte: Elaborado pelo autor

Conforme revela o Gráfico 1, o RS concentra o maior número de RIs com 64% da amostra, seguido de SC com 22% e PR, com 14%. Cabe ressaltar que o edital FINEP/PCAL/XBDB 003/2009 destinado às instituições públicas, contemplou 2 instituições do PR, 1 de SC e 3 do RS.

O Gráfico 2 representa a questão 3 do questionário, que se refere ao tamanho da coleção.



Fonte: Elaborado pelo autor

Observa-se que 57,1% dos RIs afirmam possuir menos de 10.000 registros em sua coleção, o que demonstra uma paridade entre a maioria dos RIs nesse quesito. Do restante da amostra, destaca-se que apenas 21,4% possui acima de 90.000 registros. O sucesso de um repositório depende das estratégias de povoamento e políticas de depósito bem definidas que englobam direitos autorais e licenças de uso, autoarquivamento e sua divulgação no meio acadêmico a fim de que os pesquisadores sejam incentivados a depositar suas obras (LEITE, 2009). Kuramoto (2009) assinala que políticas institucionais de informação são imprescindíveis para o povoamento dos RIs por parte dos pesquisadores. Boso (2011, p. 41) associa o êxito de um RI "[...] à quantidade e à qualidade do conteúdo depositado. Para atrair conteúdo aos RIs, é relevante que principalmente os administradores tomem iniciativas, sejam proativos e criem políticas, o que despende grande esforço para obter resultados expressivos."

A questão 4 do questionário solicita aos RIs qual é o *software* utilizado. De acordo com as instituições, o Dspace é utilizado em 100% dos RIs. Segundo o site Opendoar⁴ é a ferramenta mais utilizada em RIs atualmente, se consolidando como uma plataforma flexível, com a capacidade de gerenciar, armazenar, preservar e redistribuir a informação de um repositório nos mais diversos formatos digitais. Permite a customização da interface e seu modelo de informação organizacional baseado em comunidades ou coleções. Com relação à preservação digital, Blattmann e Weber (2008, p. 473) mencionam que ela ocorre nos seguintes níveis:

1) Preservação dos Bits: para garantir que o arquivo continue exatamente o mesmo com o passar do tempo –sem nenhuma modificação –enquanto a mídia física evolui ao seu redor. 2) Preservação funcional: o arquivo muda com o passar do tempo mas, apesar disto, o material continua a ser usável da mesma forma em que era originalmente, enquanto os formatos digitais (e a mídia física) evoluem com o passar do tempo. 3) URL persistentes (Handles). (BLATTMANN; WEBER, 2008, p. 473).

O Dspace é um *software* constituído sobre os princípios da Iniciativa de Código Aberto e utiliza o protocolo OAI-PMH.

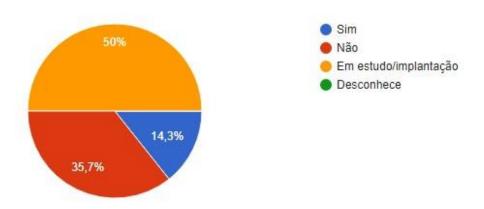
O Gráfico 4 representa a existência ou não de uma política de preservação formalizada de acordo com a questão 5.

-

⁴ http://www.opendoar.org/find.php

Gráfico 3 – Existência de política formalizada

Política de preservação digital formalizada



Fonte: Elaborado pelo autor

Com base nas respostas ilustradas pelo Gráfico 3, observa-se que apenas 14,3% das instituições dispõem de uma política de preservação digital formalizada. Outros 50% estão em fase de estudo/implantação e 35,7% não possuem políticas, apesar de possuírem repositório em funcionamento. A ausência de uma política formalizada em 85,7% da amostra total demonstra que grande parte dos RIs trabalham sem critérios, procedimentos e métodos padronizados no que se refere a preservação digital. Márdero Arellano (2004) aponta que políticas de preservação digital contribuem de forma efetiva para o armazenamento e uso dos recursos de informação a longo prazo. Nesse sentido, Santos Júnior (2010, p. 47) enfatiza a importância de se estabelecer quais os "[...] padrões, formatos, tipologias, entre outras características inerentes a documentos digitais que serão aceitos [...]", dando preferência à "[...] escolha de formatos de alta e média adequação, visando minimizar a necessidade de migração." (WEITZEL; MESQUITA, 2015, p. 193).

Weitzel e Mesquita (2015, p. 182) afirmam que:

[...] dados do OpenDOAR revelam que entre 2.509 repositórios ali cadastrados somente, 7,4% possuem uma política de preservação digital formalizada, ao passo que no Brasil há mais de 80% de RI sem uma política de preservação definida e cerca de 10,8% sem política formalizada. (WEITZEL; MESQUITA, 2015, p. 182).

Mesquita (2014) identifica situação semelhante em pesquisa realizada com RIs da região Sudeste, onde constata que 89% não possuem política formalizada. Baggio (2016) reafirma a ausência de políticas de PD, quando ratifica que 70% dos RIs de universidades

federais selecionadas pelo edital FINEP/IBICT não contemplam em suas políticas informações específicas sobre preservação digital.

Simultaneamente à pergunta 5, questionou-se aos participantes que afirmaram já possuírem políticas de preservação digital formalizada, se esta é própria do RI ou é ligada à instituição de ensino.

Políticas próprias ou da instituição

é uma política própria do RI?
é uma política da Instituição?

Gráfico 4 – Política formalizada

Fonte: Elaborado pelo autor

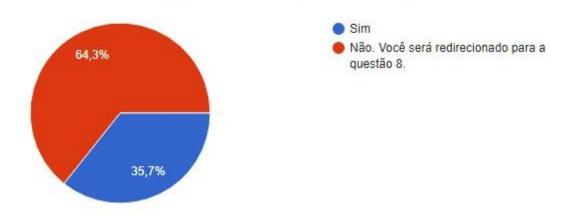
Ao comparar os gráficos 3 e 4, percebe-se que dos 14,3% que afirmaram já possuir políticas de preservação digital formalizadas, 25% disseram que estas políticas são próprias do repositório e 75% as elaboraram em conjunto com a instituição. Grácio, Fadel e Valentim (2013, p. 116) afirmam que:

A implantação de uma política de preservação digital depende, inicialmente, da vontade da instituição, ou seja, deve estar inserida nos objetivos da IES e ir ao encontro da missão e visão institucional, pois são esses dois elementos que definem os objetivos institucionais. (GRÁCIO; FADEL; VALENTIN, 2013, p. 116).

Weitzel e Mesquita (2015) complementam que a política de preservação digital formalizada é uma forma de divulgar à comunidade acadêmica os procedimentos de preservação a longo prazo que serão adotados e "[...] a falta desse documento formalizado pode evidenciar problemas em relação à padronização de critérios, métodos ou periodicidade definida para o desenvolvimento do processo de preservação digital." (BAGGIO, 2016, p. 250).

Gráfico 5 – Rede de preservação digital





O Gráfico 5, que corresponde à questão 6, ilustra que 9 RIs (64,3%), não participam das iniciativas de preservação em rede. Cinco RIs (35,7%) afirmaram participar de redes de preservação digital e também responderam à questão 7 que solicitava quais são as redes a que fazem parte. Dessa questão obteve-se que 3 RIs integram a Rede Cariniana, administrada pelo IBICT, que oferece orientações, pacotes de *softwares*, aplicações e ambientes multimídia para a implementação e desenvolvimento de centros de memória digital e capacitação de recursos humanos. Márdero Arellano (2012, p. 90) assinala que "O IBICT propôs a rede porque se trata de um instrumento para a busca das melhores práticas de preservação digital e de metodologias criteriosas para ajudar a preservar documentos digitais importantes."

O projeto LOCKSS foi citado por 1 RI e rede AUSJAL também por 1 instituição, porém não se trata de uma rede específica para preservação digital.

Estratégias de preservação digital

Backup

Metadados

Não realiza

Migração

Encapsulamento

Preservação de tecnologia

Preservação Digital Distribuída

Gráfico 6 – Estratégias de preservação digital

A partir da questão 8, a apresentação de resultados passa a ter novo arranjo, devido às questões permitirem mais de uma resposta. Conforme ilustrado no Gráfico 6, o backup (cópias de segurança) é utilizado por 9 instituições (64,28%) como estratégia de preservação digital, sendo citado por 4 como a única estratégia a ser utilizada. Em seguida, está a inclusão de metadados em 7 RIs. Nesse quesito, chama atenção o fato de 2 RIs não utilizarem nenhuma estratégia, o que significa que os acervos estão vulneráveis a danos irreparáveis a qualquer momento, uma vez que não há "[...] garantia de acesso, confiabilidade e integridade dos documentos de longo prazo." (MÁRDERO ARELLANO, 2004, p. 15). Destaca-se também que 1 instituição afirma utilizar todas as estratégias citadas, além de desenvolver o uso do software Archivematica (software de acesso aberto utilizado para Repositório Arquivísticos Digitais Confiáveis- RDC-Arq), capaz de interoperar com Dspace, oferecendo derivadas de acesso (cópias). Outro dado inusitado, é que apenas 1 RI afirma utilizar a preservação digital distribuída como estratégia, mas na sexta questão, 5 instituições afirmaram utilizar Rede Cariniana e a LOCKSS. Tal fato infere que os RIs não reconhecem a Cariniana e LOCKSS como redes de preservação digital distribuída ou pelo fato da Cariniana preservar os periódicos do SEER (Sistema Eletrônico de Editoração de Revistas).

Identificadores persistentes 8 7 ■ Handle System Não utiliza ■ PURL DOI 1

Gráfico 7 – Identificador persistente

Com relação ao uso de identificadores persistentes, 7 RIs (50%) citaram a utilização do Handle System, considerado por Sayão (2007, p. 71) "[...] como um sistema de informação de grande amplitude, projetado para alcançar interoperabilidade global através de uma rede hierarquicamente distribuída de servidores."

Embora a utilização de links persistentes seja uma garantia de acesso a longo prazo, 6 instituições afirmaram não utilizar. Pode-se inferir que tal fato esteja relacionado à necessidade de pagamento de taxas junto às agências de registro que podem criar nome para os links persistentes.

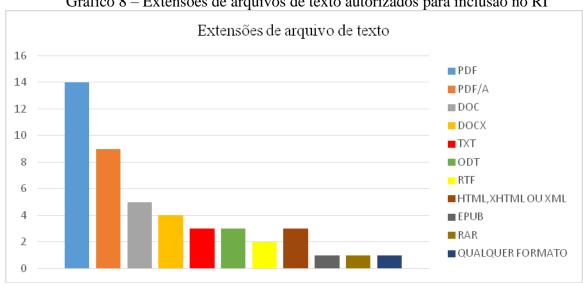


Gráfico 8 – Extensões de arquivos de texto autorizados para inclusão no RI

Fonte: Elaborado pelo autor

Os resultados obtidos da questão 10 estão representados no gráfico 8, onde apontam que 100% dos RIs utilizam o formato PDF, seguido do PDF/A, citado por 9 instituições. A inclusão de qualquer tipo de formato foi mencionada por 1 respondente. Baggio (2016) sugere que os RIs convertam os arquivos PDF em PDF/A, que é uma extensão preservável a longo prazo.

Considerando a classificação de Munhoz (2011), apresentada nos quadros 2 a 4, os resultados do Gráfico 8 foram divididos em três grupos de acordo com a cor da primeira coluna da tabela, a saber: verde (alta adequação para preservação), amarelo (média adequação para preservação) e vermelho (baixa adequação para preservação). Extensões que se repetiam em mais de uma opção de classificação, levando em consideração aspectos técnicos do arquivo, foram classificadas na melhor classificação possível, e não repetida. Por exemplo, o formato HTML, XHTML ou XML, se salvo com referência para o DTD com esquema de validação de formato, estava classificado como alta adequação de preservação.

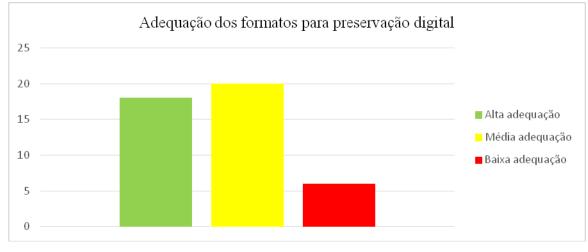


Gráfico 9 – Adequação dos formatos de texto

Fonte: Elaborado pelo autor

A partir do gráfico 9 é possível constatar homogeneidade no uso dos formatos de alta e média preservação, indicando que os RIs tem seguido as recomendações no que se refere à utilização de formatos que contribuem para preservação de longo prazo e que não dependam de *softwares* específicos para serem acessados, porém como destaca Santos e Flores (2015, p. 51) "Uma política de preservação que defina o uso de formatos padronizados para a preservação desde a produção dos documentos reduz o risco de obsolescência [...]" e reduz também a variedade de formatos quando da implementação das estratégias de PD.

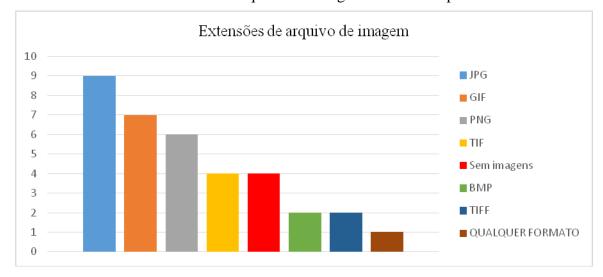


Gráfico 10 - Extensões de arquivos de imagem autorizados para inclusão no RI

A questão 10 questionou os participantes acerca dos formatos utilizados para imagem, conforme ilustrado no Gráfico 11. A extensão JPG foi a mais citada, sendo utilizada por 9 instituições, seguido do GIF citado por 7. Ambos os formatos utilizam compressão de imagens e não necessitam de *software* específico para serem acessados conforme recomenda Bodê (2008, p. 35) ao ressaltar que:

Um ponto crucial sobre Formatos de Arquivo e que está diretamente ligado aos problemas com sua preservação se refere ao fato de se tratar de um formato proprietário ou não. Os formatos abertos de arquivo (aqueles em que o público tem acesso aos detalhes técnicos) são mais adequados para a preservação futura, pois as possibilidades de compreender o significado de sua estrutura de bits é maior. (BODÊ, 2008, p. 35).

Das extensões citadas pelos repositórios, nenhuma necessita de aplicação específica para ser visualizada.

Adequação dos formatos para preservação digital

20
18
16
14
12
10
8
6
4
2
0

Gráfico 11 – Adequação dos formatos de imagem

O cenário apurou disparidade entre o uso das extensões de média e alta adequação. Além disso, das 4 extensões mencionadas por Munhoz (2011) como de alta adequação, foram citados o uso de apenas 2.

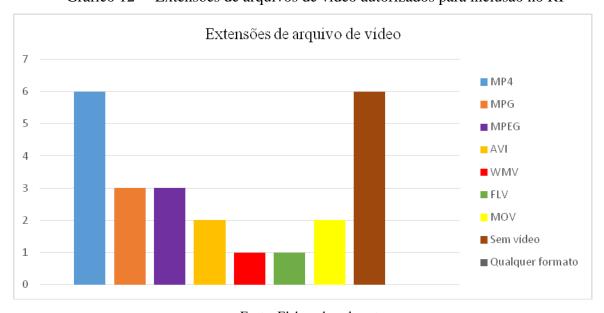


Gráfico 12 – Extensões de arquivos de vídeo autorizados para inclusão no RI

Fonte: Elaborado pelo autor

O Gráfico 12 indica a extensão MP4 como a mais utilizada, a qual segue o mesmo padrão de compressão utilizado nas imagens. Formatos considerados de alta adequação, AVI e

MOV estão presentes em 2 RIs somente. A extensão FLV, citada 1 vez, não foi classificada por Munhoz (2011) e não foi considerada nos resultados do Gráfico 14.

Adequação dos formatos para preservação digital

14
12
10
8
6
4
2
0

Gráfico 13 – Resultados - Adequação dos formatos de vídeo

Fonte: Elaborado pelo autor

Observa-se uma discrepância entre as adequações, prevalecendo o uso da média adequação para arquivos de vídeo que foram citados por 12 vezes.

Um fator importante a ser considerado pelos repositórios na escolha dos formatos a serem utilizados pode estar relacionado à condição de que os formatos altamente recomendados para a preservação digital podem divergir dos formatos recomendados para o acesso, conforme pode ser observar na Política de Preservação Digital do Arquivo Nacional que recomenda a extensão MOV para preservação digital e a extensão MPEG-2 para o acesso, considerada de média adequação por Munhoz, 2011.

6 CONSIDERAÇÕES FINAIS

A proposta desse estudo foi identificar as estratégias e iniciativas implementadas, bem como a existência de uma política de preservação digital formalizada nos repositórios das universidades região Sul do Brasil.

Dos resultados encontrados foi possível perceber que apesar dos RIs promoverem a visibilidade das pesquisas e o prestígio de uma instituição, 57,89% das universidades não possuem repositório em funcionamento, demonstrando que esse tema encontra-se ainda em desenvolvimento. Alguns repositórios não possuem link no portal principal da instituição e por esse motivo optou-se por verificar a existência dos repositórios primeiramente no site de busca Google. A divulgação do repositório na página principal da instituição é um fator importante a ser considerado.

Grande parte dos repositórios pesquisados possuem até 10 mil itens em suas coleções, o que denota uma homogeneidade nesse quesito. Todos afirmaram utilizar o Dspace, que é um *software* voltado à criação de repositórios e orientado para o acesso aberto.

No âmbito da preservação digital, a pesquisa demonstrou que os repositórios ainda carecem de políticas formalizadas, que são elementos importantes no sentido de padronizar critérios e procedimentos de preservação a longo prazo. Ainda que algumas estratégias de preservação sejam adotadas, a ausência do documento não permite que tais ações sejam de conhecimento dos usuários e pesquisadores que ali depositam suas obras. A criação de uma política de preservação digital documentada, vinculada à política institucional demonstra uma preocupação por parte da instituição em estabelecer metodologias necessárias a fim de preservar o acesso a longo prazo da produção científica, bem como salvaguardar a memória institucional.

O uso das redes de preservação digital não está consolidado no âmbito dos RIs. Podese inferir que há até mesmo desconhecimento por parte dos repositórios, visto que 5 instituições afirmaram participar de redes de preservação, mas não a reconhecem como um recurso de preservação digital distribuída que é o caso da LOCKSS e Rede Cariniana.

No que se refere às estratégias de preservação, o backup e o uso dos metadados foram os mais citados. Observa-se uma falta de diversificação no uso das demais estratégias recomendadas pela literatura: preservação de tecnologia, migração, encapsulamento e emulação. Uma instituição afirma utilizar todas as estratégias ao passo que duas mencionam não fazer uso de nenhum recurso de preservação digital a longo prazo. A construção de uma política de preservação digital formalizada é uma vertente necessária no sentido de normatizar

padrões, frequências e circunstâncias em que a preservação deve ocorrer inclusive no que se refere ao uso de formatos de arquivo. Apesar dos repositórios utilizarem em sua maioria, formatos de média e alta preservação, a padronização diminui a variedade de formatos, contribuindo com a aplicação das estratégias de preservação.

Em termos gerais, a preservação digital em repositórios institucionais é um tema que merece maior atenção por parte das instituições principalmente no que se refere a formalização das políticas de preservação digital, que pode-se considerar o primeiro passo a sedimentar a adoção das estratégias expostas nesse trabalho, de acordo com a necessidade de cada comunidade, sempre em concordância com o planejamento acadêmico institucional.

REFERÊNCIAS

BAGGIO, Claudia Carmem. **Análise das políticas de informação dos repositórios institucionais das Universidades Federais do Brasil**. 2016. 352 f. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Programa de Pós-Graduação em Ciência da Informação, Florianópolis, 2016.

BLATTMANN, Ursula; WEBER, Claudiane. Dspace como repositório digital na organização. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 13, n. 2, p. 467-485, jul./dez. 2008. Disponível em: http://revista.acbsc.org.br/racb/article/view/593. Acesso em: 05 mar. 2014.

BODÊ, Ernesto Carlos. **Preservação de documentos digitais**: o papel dos formatos de arquivo. 2008. 153 f. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2008. Disponível em: http://repositorio.unb.br/handle/10482/2034>. Acesso em: 11 dez. 2017.

BORBA, Vildeane da Rocha; LIMA, Marcos Galindo. Preservação digital: modelo orientador para o BDTD/UFPE. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. **Anais**... João Pessoa: UFPB, 2009. Disponível em: http://200.20.0.78/repositorios/handle/123456789/652. Acesso em: 24 set. 2017.

BOSO, Karla. **Repositórios de instituições federais de ensino superior e suas políticas**: análise sob o aspecto das fontes informacionais. 2011. 150 f. Dissertação (Mestrado em Ciência da Informação) — Universidade Federal de Santa Catarina, Florianópolis, 2011. Disponível em: https://repositorio.ufsc.br/xmlui/handle/123456789/95776>. Acesso em: 01 dez. 2017.

BRASIL. ARQUIVO NACIONAL. **Política de preservação digital**. 2ª versão. Rio de Janeiro: Arquivo Nacional, 2016. Disponível em: www.siga.arquivonacional.gov.br/images/an_digital/and_politica_preservacao_digital_v2.pdf. Acesso em: 22 jan. 2018.

CASTELLS, M. **A Sociedade em Rede**: A Era da Informação, Sociedade e Cultura. 3. ed. São Paulo: Paz e Terra, 1999.

CASTRO, Cristiane et al. Repositórios institucionais confiáveis: repositório institucional como ferramenta para a preservação digital. In: SAYÃO, Luis et al. **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2010, p. 283-304. Disponível em:

https://repositorio.ufba.br/ri/bitstream/ufba/473/3/implantacao_repositorio_web.pdf. Acesso em: 08 nov. 2017.

CAFÉ, L. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 26., 2003, Belo Horizonte. **Anais**... Belo Horizonte: INTERCOM – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2003. Disponível em:

http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003_ENDOCOM_TRABALHO_cafe.pdf. Acesso em: 06 nov. 2017.

CONSELHO NACIONAL DE ARQUIVOS. Carta para Preservação do Patrimônio Arquivístico Digital. Rio de Janeiro, 2004. Disponível em:

http://www.conarq.arquivonacional.gov.br/images/publicacoes_textos/Carta_preservacao.pdf >. Acesso em: 18 dez. 2017.

CONWAY, Paul. **Preservação no universo digital**. 2. ed. Rio de Janeiro: Projeto Conversação Preventiva em Bibliotecas e Arquivos: Arquivo Nacional, 2001. 33 p. Disponível em: http://www.prefeitura.sp.gov.br/cidade/upload/cpba_52_1253284406.pdf>. Acesso em: 25 jun. 2017.

CORRÊA, Amarílis Montagnolli Gomes. **Preservação digital: autenticidade e integridade de documentos em bibliotecas digitais de teses e dissertações**. 2010. 96 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2010. Disponível em: http://www.teses.usp.br/teses/disponiveis/27/27151/tde-05112010-105831/publico/2916162.pdf>. Acesso em: 26 jun. 2017.

COSTA, Sely Maria de Souza; LEITE, Fernando Cesar Lima. Insumos conceituais e práticos para iniciativas de repositórios institucionais de acesso aberto à informação científica em bibliotecas de pesquisa. In: SAYÃO, Luis et al. **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2010. Disponível em: http://eprints.rclis.org/14906/1/Costa_%26_Leite_2010.pdf>. Acesso em: 12 set. 2017.

DIAS, Rafael Cobbe. **Preservação Digital**: análise dos periódicos da base BRAPCI. 2011. 68 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) — Universidade Federal de Santa Catarina, Florianópolis, 2011. Disponível em: http://eprints.rclis.org/16429/1/Rafael%20Cobbe%20Dias_2011_Preserva%C3%A7%C3%A30%20Digital.pdf. Acesso em: 14 nov. 2017.

DODEBEI, Vera. Repositórios institucionais: por uma memória criativa no ciberespaço. In: SAYÃO, L. et al. **Implantação e gestão de repositórios institucionais**: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. p. 83-106. Disponível em: https://repositorio.ufba.br/ri/bitstream/ufba/473/3/implantacao_repositorio_web.pdf. Acesso em: 04 nov. 2017.

e-MEC. **Instituições de Educação Superior e cursos cadastrados**. 2017. Disponível em: <emec.mec.gov.br>. Acesso em: 15 nov. 2017.

FERREIRA, Carla Alexandra Silva. **Preservação da informação digital**: uma perspectiva orientada para as bibliotecas. Dissertação (Mestrado em Informação, Comunicação e Novos Media) – Faculdade de Letras da Universidade de Coimbra, Coimbra, 2011. Disponível em: https://estudogeral.sib.uc.pt/bitstream/10316/15001/1/Preserva%C3%A7%C3%A3o%20da%20Informa%C3%A7%C3%A3o%20Digital.pdf. Acesso em: 9 nov. 2017.

FERREIRA, Miguel. **Introdução à preservação digital**: conceitos, estratégias e actuais consensos. Guimarães: Escola de Engenharia da Universidade do Minho, 2006. Disponível em: http://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>. Acesso em: 26 jun. 2017.

FONSECA, João José Saraiva da. **Metodologia da pesquisa científica**. Fortaleza: UEC, 2002. Apostila. Disponível em:

http://leg.ufpi.br/subsiteFiles/lapnex/arquivos/files/Apostila_-

_METODOLOGIA_DA_PESQUISA(1).pdf.>. Acesso em: 9 out. 2017.

GIL, A. C. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2002.

GRÁCIO, José; FADEL, Bárbara; VALENTIM, Marta. Preservação digital nas instituições de ensino superior: aspectos organizacionais, legais e técnicos. **Perspectivas em Ciência da Informação**, v. 18, n. 3, 2013. Disponível em:

http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/1612. Acesso em: 30 nov. 2017.

GUEDES, Rodrigo Duarte. O surgimento dos repositórios institucionais e uma breve análise dos instrumentos legais. In: CONGRESSO DE DIREITO DE AUTOR E INTERESSE PÚBLICO, 4., 2010, Florianópolis. **Anais**... Disponível em:

http://www.gedai.com.br/sites/default/files/publicacoes/final_codaip2.pdf. Acesso em: 13 nov. 2017.

HONORATO, Gilson. Conhecendo o marketing. Barueri: Manole, 2004.

INNARELLI, Humberto. Os dez mandamentos da preservação digital. In: SANTOS, V. B.; INNARELLI, H. C.; SOUSA, T. R. B. **Arquivística**: temas contemporâneos. Brasília: SENAC, 2007. p. 39-122.

_____. **Instrumenta 2**: preservação de documentos digitais. São Paulo: ARQ-SP, 2012.

KING, Donald W. An approach to open access author payment. **D-LIB Magazine**, v. 16, n. 3/4, mar./abr. 2010. Disponível em:

http://www.dlib.org/dlib/march10/king/03king.print.html. Acesso em: 18 nov. 2017.

KURAMOTO, Hélio. Repositórios institucionais: políticas e mandatos. In: SAYÃO, Luis Fernando et al. (Org.). **Implantação e gestão de repositórios institucionais**: política, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. p. 203-217. Disponível em:https://repositorio.ufba.br/ri/bitstream/ufba/473/3/implantacao_repositorio_web.pdf. Acesso em: 29 nov. 2017.

_____. A informação científica e o seu acesso livre: que direção o brasil está adotando?. **Bibliotecas Universitárias: pesquisas, experiências e perspectivas**, v. 2, n. 1, p. 6-28, 2015. Disponível em: http://www.brapci.inf.br/v/a/23473. Acesso em: 23 Jan. 2018.

LAVILLE, Christian; DIONNE, Jean. A construção do saber: manual de metodologia da pesquisa em ciências humanas. Porto Alegre: Artmed, 1999. 340 p.

LAZARTE, Leonardo. Ecologia cognitiva na sociedade da informação. **Ciência da Informação**, Brasília, v. 29, n. 2, p.43-51, maio/ago. 2000. Disponível em: http://revista.ibict.br/ciinf/article/view/886/921>. Acesso em: 18 nov. 2017.

LEITE, Fernando César. **Gestão do conhecimento científico no contexto acadêmico**: proposta de um modelo conceitual. 2006. 240 f. Dissertação (Mestrado em Ciência da

Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília, 2006. Disponível em: http://repositorio.unb.br/handle/10482/3975>. Acesso em: 27 set. 2017.

LEITE, Fernando. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2009. Disponível em:

http://repositorio.unb.br/bitstream/10482/4841/1/LIVRO_ComoAmpliareGerenciar.pdf>. Acesso em: 15 set. 2017.

LOPES, Jorge. **O fazer do trabalho em ciências sociais aplicadas**. Recife: Editora Universitária UFPE, 2006. 303 p. Disponível em:

< https://books.google.com.br/books?id=A321LE03ab8C&pg=PA241&dq=INSTRUMENTO+DE+COLETA+DE+DADOS+QUESTIONARIO&hl=pt-DADOS

BR&sa=X&ei=GcfxUcngGona9QTrl4CIDQ#v=onepage&q=questionario&f=false>. Acesso em: 17 nov. 2017.

LOCKSS - LOTS OF COPIES KEEP STUFF SAFE. **Home**. Stanford, 2014. Disponível em: <www.lockss.org>. Acesso em: 12 dez. 2017.

MARCONDES, Carlos Henrique. Metadados: descrição e recuperação de informações na Web. In: MARCONDES, Carlos Henrique et al. (Org.). **Bibliotecas digitais**: saberes e práticas. Salvador: UFBA; Brasília: IBICT, 2005. p. 97-113.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos da metodologia científica**. 6. ed. São Paulo: Atlas, 2007.

MÁRDERO ARELLANO, Miguel Ángel. **Critérios para a preservação digital da informação científica**. 2008. 354 f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2008. Disponível em: http://hdl.handle.net/10760/12649>. Acesso em: 04 jul. 2017.

. Preservação de documentos digitais. Revista Ciência da Informação , Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: http://revista.ibict.br/ciinf/article/view/1043/1113 . Acesso em: 29 jun. 2017.
Cariniana: uma rede nacional de preservação digital. Revista Ciência da [nformação, Brasília, v. 41, n. 1, p. 83-91, jan./abr. 2012, Disponível

MESQUITA, Marco Aurélio Alencar de. **Preservação digital em repositórios institucionais**: estratégias, iniciativas e práticas na região sudeste do Brasil. 2014. 69 f. Trabalho de Conclusão de Curso (Escola de Biblioteconomia) — Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em: http://www.unirio.br/unirio/cchs/eb/arquivos/tccs-2014.2/TCC1%20marco%20aurelio%20alencar.pdf>. Acesso em: 01 jul. 2017.

em:http://revista.ibict.br/ciinf/article/viewFile/1354/1533. Acesso em: 23 jul. 2014.

MORESI, Eduardo. **Metodologia da Pesquisa**. Brasília: Universidade Católica de Brasília, 2003. Disponível em:

http://xa.yimg.com/kq/groups/22703089/512340126/name/MetodologiaPesquisa-Moresi2003.pdf. Acesso em: 30 jun. 2017.

MUELLER, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, ago. 2006. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000200004&lng=en&nrm=iso. Acesso em: 09 jul. 2017.

MUNHOZ, Augusto Maciel. **Preservação digital de acervos**: uma análise das políticas e estratégias utilizadas pelas universidades do Rio Grande do Sul. 2011. 63 f. Trabalho de Conclusão de Curso (Faculdade de Biblioteconomia e Comunicação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011. Disponível em: http://www.lume.ufrgs.br/handle/10183/37473>. Acesso em: 29 set. 2017.

NISO - NATIONAL INFORMATION STANDARDS ORGANIZATION. **Understanding Metadata**. Bethesda: NISO Press, 2004. 20 p. Disponível em: http://www.niso.org/publications/press/UnderstandingMetadata.pdf>. Acesso em: 09 jul. 2017.

OPEN SOURCE INITIATIVE. **Open Source Initiative**. 2014. Disponível em: http://opensource.org/>. Acesso em: 19 dez. 2017.

PAVÃO, Caterina Groposo. **Contribuição dos repositórios institucionais à comunicação científica**: um estudo na Universidade Federal do Rio Grande do Sul. 2010. 152 f. Dissertação (Mestrado em Comunicação e Informação) — Programa de Pós-Graduação em Comunicação e Informação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010. Disponível em: http://hdl.handle.net/10183/20932>. Acesso em: 08 jul. 2017.

PAVÃO, Caterina Groposo et al. Metadados e repositório institucionais: uma relação indissociável para a qualidade da recuperação e visibilidade da informação. **Ponto de Acesso**, Salvador, v. 9, n. 2, p.103-116, dez. 2015. Disponível em: https://portalseer.ufba.br/index.php/revistaici/article/view/15163>. Acesso em: 16 dez. 2017.

RIBEIRO, Odília; VIDOTTI, Silvana. G. Otimização do acesso à informação científica: discussão sobre a aplicação de elementos da Arquitetura da Informação em repositórios digitais. **Biblos**, v. 23, p. 105-116, 2009. Disponível em:https://www.seer.furg.br/biblos/article/view/1309/593>. Acesso em: 29 nov. 2017.

ROAR - REGISTRY OF OPEN ACCESS REPOSITORIES. **Registry of Open Access Repositories**. 2017. Disponível em: http://roar.eprints.org/>. Acesso em: 17 dez. 2017.

SANTOS, Henrique; FLORES, Daniel. As vulnerabilidades dos documentos digitais: obsolescência tecnológica e ausência de políticas de preservação digital. **Biblios: Revista de Bibliotecología y Ciencias de la Información**, n. 59, 2015. Disponível em:http://biblios.pitt.edu/ojs/index.php/biblios/article/view/215/229. Acesso em: 01 dez. 2017.

SANTOS JUNIOR, Ernani Rufino dos. **Repositórios institucionais de acesso livre no Brasil**: estudos delfos. 2010. 177 f. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2010. Disponível em: http://repositorio.unb.br/bitstream/10482/5343/6/2010_ErnaniRufinodosSantosJunior.pdf>. Acesso em: 05 jul. 2014.

SARAMAGO, M. L. Metadados para preservação digital e aplicação do modelo OAIS. **Bad**, 2004. Disponível em:

https://www.bad.pt/publicacoes/index.php/congressosbad/article/view/640/637. Acesso em: 07 jul. 2017.

SAYÃO, Luis Fernando. Preservação digital no contexto das bibliotecas digitais: uma breve introdução. In: MARCONDES, Carlos Henrique et al. (Org.). **Bibliotecas digitais**: saberes e práticas. Salvador: UFBA; Brasília: IBICT, 2005. p. 115-146. Disponível em: http://livroaberto.ibict.br/bitstream/1/1013/1/Bibliotecas Digitais.pdf>. Acesso em: 27 jun. 2017.

_____. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes – URN, PURL, DOI, Handle System, CrossRef e OpenURL. **Transinformação**, v. 19, n. 1, p. 65-82, 2007. Disponível em: http://www.brapci.inf.br/v/a/459>. Acesso em: 18 dez. 2017.

_____. Uma outra face dos metadados: informações para a gestão da preservação digital. **Revista Encontros Bibli**, Florianópolis, v. 15, n. 30, p. 1-31, 2010. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2010v15n30p1/19527. Acesso em: 09 jul. 2017.

SEARA, Everton. **Uma arquitetura OAI para preservação digital utilizando redes Peerto-Peer estruturadas**. 2008. 80 f. Dissertação (Mestrado em Informática) — Programa de Pósgraduação em Informática, Universidade Federal do Paraná, Curitiba, 2008. Disponível em: http://acervodigital.ufpr.br/handle/1884/23744>. Disponível em: 14 nov. 2017.

SKINNER, Katherine; SCHULTZ, Matt. **A Guide to distributed digital preservation**. Georgia: Educopia Institute, 2010. Disponível em: https://educopia.org/sites/educopia.org/files/publications/A_Guide_to_Distributed_Digital_Preservation_0.pdf>. Acesso em: 03 out. 2017.

TAVARES, Maria de Fátima Duarte; MÁRDERO ARELLANO, Miguel Ángel. Preservação do patrimônio científico das humanidades: a emergência da Rede Cariniana. **Cadernos de História**, Belo Horizonte, v. 16, n. 25, p. 30-45, dez. 2015. Disponível em: http://periodicos.pucminas.br/index.php/cadernoshistoria/article/view/P.2237-8871.2015v16n25p30. Acesso em: 02 dez. 2017.

VALADÃO, Renan Bernardo. **Poluição em redes P2P**. 2008. Disponível em: https://www.gta.ufrj.br/ensino/eel879/trabalhos_vf_2008_2/renan_bernardo/info.html. Acesso em: 11 nov. 2017.

WEITZEL, Simone. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 51-71, jan/jun. 2006. Disponível em: http://seer.ufrgs.br/index.php/EmQuestao/article/view/19. Acesso em: 29 nov. 2017.

WEITZEL, Simone; MESQUITA, Marco Aurélio Alencar de. Preservação digital em repositórios institucionais: práticas na região sudeste do brasil. **Liinc em Revista**, v. 11, n. 1, 2015. Disponível em: https://doi.org/10.18617/liinc.v11i1.778>. Acesso em: 30 nov. 2017.

WENSING, Jairo. **Preservação e recuperação de informação em fontes de informações digitais**: estudo de caso do Greenstone. 2010. 219 f. Dissertação (Mestrado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Centro de Ciências da Educação, Universidade Federal de Santa Catarina, Florianópolis, 2010. Disponível em: http://repositorio.ufsc.br/xmlui/handle/123456789/93702. Acesso em: 05 out. 2017.

APENDICE A - Questionário

Pesquisa - Preservação Digital em Repositório Institucional

Prezado,

* Obrigatório

A presente pesquisa contempla o Trabalho de Conclusão de Curso da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul (FABICO/UFRGS) e se propõe Identificar estratégias e iniciativas de preservação digital utilizadas nos repositórios institucionais (RI) da região Sul do Brasil.

O instrumento de coleta de dados é composto de 12 (doze) questões, não ocupando mais do que 5 minutos do seu tempo. Salienta-se que será mantido sigilo, de forma que os nomes dos participantes e dos respectivos RI não sejam divulgados.

A participação nesta pesquisa não oferece risco ou prejuízo à pessoa participante. Se no decorrer desta pesquisa o(a) participante resolver não mais continuar terá toda a liberdade de o fazer, sem que isto lhe acarrete nenhum prejuízo.

Os pesquisadores responsáveis por esta pesquisa são a aluna Aline Macedo Brito (Graduanda em Biblioteconomia pela FABICO/UFRGS) e a professora Ana Mielniczuk de Moura (FABICO/UFRGS) que se comprometem a esclarecer devida e adequadamente qualquer dúvida ou necessidade de esclarecimento que eventualmente o participante venha a ter no momento da pesquisa ou posteriormente através do email: alinecaperucita@hotmail.com.

Questionário (Adaptado de Mesquita, 2014)

1 - Instituição*:	
2 - Estado/Sigla*:	
3 - Tamanho da coleção*	
Total de registros no Repositório. Pode ser um valor aproximado.	
() menor que 10.000 registros	
() de 10.001 a 30.000	
() de 30.001 a 60.000	
() de 60.001 a 90.000	
() acima de 90.000 registros	

4 - Em qual <i>software</i> o Repositório foi criado?
() Dspace
() E-prints
() Bepress
() Fedora
() Outro. Qual?
() 0 4401 2 4411
5 – O RI possui uma política de preservação digital formalizada?
() Sim () Não () Em estudo/implantação () Desconhece
() () - ···· ()
Caso a resposta acima seja afirmativa:
() É uma política própria do RI? () É uma política da Instituição?
() 2 ama ponaea propria do ra: () 2 ama ponaea da mentangae.
Sobre a Política de Preservação Digital em Rede
, C
6 - A instituição ou RI faz parte de alguma rede que tenha finalidade da preservação digital? *
() Sim
() Não
Caso negativo, você será redirecionado à pergunta nº 8
7 - Indique qual é a rede que tenha finalidade da preservação digital que o Repositório
Institucional faz parte?
() Redes Cariniana
() LOCKSS
() Outro:
() out of
Sobre as Estratégias de Preservação Digital
8 - Qual a estratégia de preservação digital utilizada?
Se necessário, marque mais de uma opção
() Não realizamos
() Inclusão de Metadados
() Backup
() Migração
() Encapsulamento
() Preservação da tecnologia
() Preservação Digital Distribuída
() Outro:
() Outio
9 - Utiliza algum método de identificação persistente para os objetos digitais?
Se necessário, marque mais de uma opção
() Não utilizamos
() URN - Uniform Resource Name
() RUL - Uniform Resource Locator
() PURL - Persistent URL
() DOI - Digital Object Identifier
() Handle System
CONTO

	10 - Quais são os tipos documentos de TEXTO?	de	ar	quivos	(exte	ensão)	autoriza	idos	para	inclusão	no	RI	para
) PDF) PDF/A) DOC) DOCX) TXT) ODT) RTF		() HTM) Outro) Qualo)		ou XML						
	11 - Quais são os tipos documentos de IMAGEM?		ar	quivos	(exte	ensão)	autoriza	dos	para	inclusão	no	RI	para
() Sem imagens		() BMP	•								
() JPG		() DNG									
() JPX		() PSD									
() JP2		() RAW	7								
() JPF		() SVG									
() GIF		() SWF									
() TIFF		() Outro									
() TIF		() Qual									
() PNG		`	, ,	•								
	12 - Quais são os tipos documentos de VÍDEO?	de	ar	quivos	(exte	ensão)	autoriza	idos	para	inclusão	no	RI	para
() Sem vídeos		() MJ2									
() MOV		() OGG	j								
() MPG		() OGM									
() MPEG		() OGV									
() AVI		() Outro									
() WMV		() Qual									
() MP4		•	. ~	-								