

UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO  
ESCOLA DE BIBLIOTECONOMIA

MARCO AURÉLIO ALENCAR DE MESQUITA

**PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS:**  
estratégias, iniciativas e práticas na região sudeste do Brasil.

Rio de Janeiro

2014

MARCO AURÉLIO ALENCAR DE MESQUITA

**PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS:**

estratégias, iniciativas e práticas na região sudeste do Brasil.

Trabalho de Conclusão de Curso apresentado como requisito para a obtenção do título de Bacharel em Biblioteconomia, pela Escola de Biblioteconomia da Universidade Federal do Estado do Rio de Janeiro.

Orientadora: Simone da Rocha Weitzel

Rio de Janeiro

2014

M582p

MESQUITA, Marco Aurélio Alencar de.

Preservação digital em repositórios institucionais: estratégias, iniciativas e práticas na região sudeste do Brasil. / Marco Aurélio Alencar de Mesquita. – Rio de Janeiro, 2014.

69 f. : il. color; 30cm.

Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade Federal do Estado do Rio de Janeiro.

Orientador: Prof. Simone da Rocha Weitzel.

1. Preservação digital. 2. Repositório Institucional. 3. DSpace. 4. Estratégias de Preservação Digital. I. Universidade Federal do Estado do Rio de Janeiro. II.

Título

CDU – 021:004

MARCO AURÉLIO ALENCAR DE MESQUITA

**PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS:**

estratégias, iniciativas e práticas na região sudeste do Brasil.

Trabalho de Conclusão de Curso apresentado como requisito para a obtenção do título de Bacharel em Biblioteconomia, pela Escola de Biblioteconomia da Universidade Federal do Estado do Rio de Janeiro.

Aprovado em: \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

**BANCA EXAMINADORA**

---

Prof. Simone da Rocha Weitzel

(Orientadora)

Universidade Federal do Estado do Rio de Janeiro

---

Prof. Carlos Alberto Ferreira

Universidade Federal do Estado do Rio de Janeiro

---

Prof. Eduardo da Silva Alentejo

Universidade Federal do Estado do Rio de Janeiro

## **AGRADECIMENTOS**

Agradeço a Deus pelo dom da vida e pela força da fé durante esta caminhada.

Aos meus familiares, pela confiança, reconhecimento e auxílio incondicional nas minhas realizações.

A minha esposa, Vanusa, pelo suporte que apenas o amor verdadeiro proporciona.

A Universidade Federal do Estado do Rio de Janeiro - UNIRIO, pela oportunidade de desenvolvimento profissional e pessoal. As tantas amizades construídas durante o período do curso de bacharelado; aos amigos do Núcleo de Estudos Inter-transdisciplinar de Educação, Sexualidade, Saúde e Cultura – NIESC/UNIRIO – ao qual tive orgulho de pertencer; ao corpo técnico desta instituição: auxiliares, seguranças e tantos outros que sempre permanecerem unidos no objetivo de conduzir a UNIRIO a excelência que lhe cabe.

A todos os meus professores que, certamente, contribuíram para este momento e que sempre transpareceram o amor pela biblioteconomia. Em especial a minha orientadora, Prof. Simone Weitzel, pelas aulas que suscitaram não apenas o tema desta pesquisa, mas o interesse na interseção entre biblioteconomia e tecnologia.

A todos que por acaso não tenha lembrado nestas linhas, mas que estiveram junto a mim nesta caminhada.

*“O cérebro eletrônico comanda*

*Manda e desmanda*

*Ele é quem manda*

*Mas ele não anda”*

*(Gilberto Gil)*

## RESUMO

O presente trabalho tem por objetivo a identificação das boas práticas de preservação digital aplicada aos Repositórios Institucionais (RIs). Tem como objeto as instituições públicas de ensino e de pesquisa na região sudeste do Brasil, conforme critérios de edital de fomento para RIs. Ressalta os benefícios trazidos pelas Tecnologias da Informação e Comunicação, dentre eles, os RIs, e também o papel da política de preservação digital no estabelecimento das práticas de preservação, refletindo sobre a responsabilidade institucional, imputada por vezes, apenas à mudança de suporte para o digital. Apresenta a complexidade da preservação digital que, diferente do documento em suporte material, possui dimensões que envolvem a fragilidade do ambiente técnico, suportes e extensões, que devem ser considerados na política de preservação digital. Utiliza metodologias de pesquisa documental, realizada através de avaliação nos endereços digitais dos RIs selecionados, e questionário fechado, aplicados às instituições, realizando uma análise quantitativa dos resultados através de gráficos e quadros, destacando as divergências aferidas entre os resultados das pesquisas. Detalha as principais estratégias e iniciativas de preservação digital utilizadas. Estabelece, baseado na literatura e nos resultados obtidos, subsídios para o desenvolvimento de política de preservação digital.

**Palavras-chave:** Preservação digital. Repositório Institucional. DSpace. Estratégias de Preservação Digital.

## ABSTRACT

This study aims to identify good practices to digital preservation applied to Institutional Repositories (IRs). Focuses on the public institutions of education and research in the southeast of the Brazil, based on the funding public notice criteria for IRs. Highlights the benefits of Information and Communication Technology, including, IRs, and also the role of digital preservation policy in the establishment of conservation practices, considering the institutional responsibility, attributed sometimes just change of support to digital. Shows the complexity of digital preservation that, unlike the document in support material, has dimensions involving the fragile technical environment, support and extensions, which should be considered in digital preservation policy. Uses methods of documentary research, done through a direct assessment in digital addresses of selected ICs, and questionnaire in closed questionnaire format, applied to institutions, performing a quantitative analysis of the results using graphs and tables, highlighting the measured differences between the results of researchs. Details the main digital preservation strategies and initiatives used. Establishes, based on the technical literature and the results, subsidies for the development of digital preservation policy.

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

## LISTA DE GRÁFICOS

Gráfico 1	Pesquisa documental - Tamanho da coleção.....	35
Gráfico 2	Pesquisa documental - Creative Commons.....	37
Gráfico 3	Pesquisa documental - Política de preservação digital.....	37
Gráfico 4	Questionário - Localização geográfica.....	38
Gráfico 5	Questionário - Tamanho da coleção.....	39
Gráfico 6	Questionário - Política formalizada.....	40
Gráfico 7	Questionário - Redes de preservação digital.....	41
Gráfico 8	Questionário - Estratégias de preservação.....	41
Gráfico 9	Questionário - Identificação persistente.....	42
Gráfico 10	Questionário - Extensões de texto.....	44
Gráfico 11	Questionário - Extensões de imagem.....	45
Gráfico 12	Questionário - Extensões de vídeo.....	46
Gráfico 13	Resultados - Tamanho da coleção.....	47

## LISTA DE QUADROS

Quadro 1	Formatos de Texto - Alta adequação para preservação.....	22
Quadro 2	Formatos de Texto - Média adequação para preservação.....	22
Quadro 3	Formatos de Texto - Baixa adequação para preservação.....	23
Quadro 4	Formatos de Imagem - Alta adequação para preservação.....	23
Quadro 5	Formatos de Imagem - Média adequação para preservação.....	23
Quadro 6	Formatos de Imagem - Baixa adequação para preservação.....	24
Quadro 7	Formatos de Áudio - Alta adequação para preservação.....	24
Quadro 8	Formatos de Áudio - Média adequação para preservação.....	25
Quadro 9	Formatos de Áudio - Baixa adequação para preservação.....	25
Quadro 10	Formatos de Vídeo - Alta adequação para preservação.....	25
Quadro 11	Formatos de Vídeo - Média adequação para preservação.....	25
Quadro 12	Formatos de Vídeo - Baixa adequação para preservação.....	26
Quadro 13	Amostra de Repositórios Institucionais.....	33
Quadro 14	Média entre registro e comunidades.....	36
Quadro 15	Resultados - Tipos de arquivos - Texto.....	43
Quadro 16	Resultados - Tipos de arquivos - Imagem.....	44
Quadro 17	Resultados - Tipos de arquivo - Vídeo.....	45
Quadro 18	Resultado consolidado.....	46

## SUMÁRIO

1	<b>INTRODUÇÃO</b> .....	11
2	<b>PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS</b> .....	13
2.1	A PRESERVAÇÃO DIGITAL.....	15
2.2	ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL .....	17
2.3	INICIATIVAS DE PRESERVAÇÃO DIGITAL .....	28
2.3.1	LOCKSS .....	28
2.3.2	Rede Cariniana .....	29
2.4	OS REPOSITÓRIOS INSTITUCIONAIS.....	29
2.5	O DSpace .....	31
3	<b>ESTUDO DE CASO</b> .....	33
3.1	OS REPOSITÓRIOS SELECIONADOS .....	33
3.2	PESQUISA DOCUMENTAL.....	34
3.3	COLETA DE DADOS PELO QUESTIONÁRIO .....	34
3.4	TABULAÇÃO E ANÁLISE DOS DADOS.....	34
	<b>REFERÊNCIA</b> .....	53
	<b>ANEXO A - EDITAL FINEP/PCAL/XBDB 003/2009</b> .....	57
	<b>APÊNDICE A – ROTEIRO PARA PESQUISA DOCUMENTAL</b> .....	64
	<b>APÊNDICE B – QUESTIONÁRIO</b> .....	65

## 1 INTRODUÇÃO

Inegáveis são os benefícios trazidos pela introdução das Tecnologias da Informação e Comunicação (TIC) no cotidiano das pessoas. Estes benefícios alcançam também aos Repositórios Institucionais (RI) que contribuem para a disseminação e democratização do acesso à informação. Contudo, esta facilidade na divulgação de documentos proporcionada pela TIC através dos RIs podem esconder percalços próprios ao de um acervo, tal como o bibliográfico, no tocante a preservação e garantia do acesso.

Conforme aponta Cunha (1999) a preservação material do papel requer espaço físico, condições climáticas, cuidados no manuseio, controle de empréstimo, além de custos com a biblioteca (manutenção do prédio, água, luz). Parece certo que com o novo paradigma digital – ou *hiper-paradigm* como define Goicoechea (2007), onde o prefixo "hyper" representa os hipertextos – a concentração do acervo em um RI aparenta, em um primeiro momento, a solução para diversos destes problemas. Contudo, ela representa na verdade a modificação de alguns aspectos de preservação que não podem – ou não devem – ser ignorados, como o estabelecimento de protocolos gerais de backup dos arquivos de sistema ou até mesmo de um módulo híbrido de trabalho, no caso de uma interrupção parcial do serviço. Assim, os cuidados específicos com os documentos digitais são postos em segundo plano, por se tratarem de objetos intangíveis e a preocupação com o tamanho do espaço de armazenamento, por exemplo, é invisível. De fato pode incidir em redução dos custos com a biblioteca esta mudança de paradigma, porém isso não reflete a preservação digital, que carece de políticas e iniciativas para o funcionamento.

Portanto, torna-se imprescindível debruçar-se sobre este assunto para garantir que os RIs possam estar alinhados aos objetivos de salvaguardar os acervos e garantir o acesso para esta e outras gerações.

Pretende-se nessa pesquisa estudar as políticas de preservação digital em Repositórios Institucionais (RIs) situadas na região sudeste brasileiro que utilizam como base tecnológica o DSpace por meio de pesquisa documental e questionário fechado aplicados às instituições. Como preservação digital, foi compreendido o conjunto de estratégias e protocolos na unidade da informação que visam à manutenção do banco de dados e dos objetos digitais disponibilizados, devidamente formalizados. Este conjunto de estratégias e protocolos representa um compromisso de longo prazo com o acesso e preservação (SAYÃO, 2007a), sendo sua formalização imprescindível para a garantia da eficácia do sistema.

Para estabelecer uma amostra das instituições a serem pesquisadas, partiu-se dos mesmos critérios de seleção de instituições elegíveis adotados para um edital de apoio a projetos de implantação de repositórios institucionais. Neste caso, utilizaremos os critérios do Edital FINEP/PCAL/XBDB 003/2009 (Anexo A), a saber:

- a) Ser uma Instituição Pública de Ensino e de Pesquisa com atuação em território nacional;
- b) Manter de programas de pós-graduação reconhecidos pelo MEC/CAPES;
- c) Ter o compromisso com o modelo de interoperabilidade Open Archives.

Limitaremos à amostra a região sudeste devido às restrições de cobertura do tema.

Levou-se em consideração a seguinte questão problema: Quais são as estratégias de preservação digital inseridas nos processos de gestão dos RIs das instituições públicas de ensino e de pesquisa da região sudeste do Brasil e de que maneira estas ações estão formalizadas como política?

Este trabalho tem como objetivo principal a identificação das boas práticas de preservação digital aplicada aos Repositórios Institucionais das instituições públicas de ensino e de pesquisa na região sudeste do Brasil a fim de contribuir com subsídios para a elaboração de uma política específica para os RIs, para isso será necessário:

- a) Identificar as estratégias de preservação digital usada – isoladamente ou combinadas – nos RIs das instituições públicas de ensino e de pesquisa com atuação em território nacional, mantenedora de programas de pós-graduação reconhecidos pelo MEC/CAPES;
- b) Identificar se a política de preservação digital está formalizada;
- c) Identificar como esta política está estabelecida: se é uma política da unidade de informação; se é da organização ou de parte de uma rede de colaboração;
- d) Identificar a interação dos RIs com iniciativas de preservação digital;

Os resultados desta pesquisa apontam que a maioria dos RIs avaliados não apresentam políticas de preservação disponíveis nem para consulta, tão pouco para uso interno e concentram seus esforços apenas na preservação digital realizando backups dos objetos digitais, delineando assim um longo caminho a ser trilhado para a preservação digital no Brasil.

## 2 PRESERVAÇÃO DIGITAL EM REPOSITÓRIOS INSTITUCIONAIS

O conhecimento produzido, preservado e transmitido às gerações futuras pode ser considerado um dos grandes diferenciais da humanidade. Este conhecimento preservado possibilitou o avanço da ciência, que se caracteriza por ser “uma instituição social, dinâmica, contínua, cumulativa” (TARGINO, 2000, p. 2).

Devido a estas características, a preservação e o acesso do conhecimento científico produzido ao longo de gerações de pesquisadores impulsiona a criação de um novo conhecimento para as gerações futuras. A ressignificação de um objeto na ciência moderna pressupõe um avanço, confirmação ou refutação de uma pesquisa já realizada, conforme nos explica Abrantes e Martins:

A construção praxica [sic] do conhecimento nos remete, portanto, à realidade histórica a se conhecer, visto que os indivíduos se desenvolvem em relações de apropriação da história contida nos objetos produzidos pelo homem e nas relações estabelecidas entre eles na base de tais produções (ABRANTES; MARTINS, 2007, p. 315).

A necessidade de se salvaguardar o conhecimento registrado remonta às mais antigas civilizações. Vidal (2010, p. 146) destaca o fato de que “desde a antiguidade a Humanidade procurou, nos mais diversos suportes, guardar, conservar e preservar a sua memória”. Independente do suporte do registro da informação, a necessidade de preservá-los sempre foi um sentimento comum aos homens.

Durante os séculos as bibliotecas vêm enfrentando as modificações em diversos campos correlatos à sua atuação, como o volume sempre crescente da produção científica, a massificação dos recursos informacionais, a comunicação científica e o mercado editorial, inclusive ampliando a função para, além de organizar o conhecimento, a de sistematizar e prover acesso à informação em soluções qualitativas que visem à redução do tempo de busca e serviços mais precisos (SILVA, 2008).

Apesar da própria palavra biblioteca “designar um local onde se armazenam livros” (MORIGI; SOUTO, 2005, p. 190) e ainda povoar no imaginário popular, o papel das bibliotecas clássicas, definidas como “aquela onde a maioria dos itens do seu acervo é constituída de documentos em papel” (CUNHA, 1999, p. 258) pouco se imagina que as bibliotecas sempre atuaram na vanguarda da tecnologia, pois

A biblioteca possui, como instituição social, uma longa e complexa história pouco conhecida pela maioria das pessoas. As tecnologias da imprensa, máquina de escrever, telefone, telex, mimeógrafo, microfilme, cartão

perfurado nas margens, computador, disco ótico e redes eletrônicas afetaram e alteraram a biblioteca ao longo do tempo. Algumas dessas tecnologias, tais como o microfilme e o disco ótico, tiveram suas primeiras aplicações testadas dentro de uma biblioteca (CUNHA, 1999, p. 257).

Sendo assim, a necessidade em melhorar os processos e métodos, as bibliotecas acabaram sendo beneficiadas pelo desenvolvimento tecnológico, tornando o fazer bibliotecário mais proveitoso e eficaz, diante da crescente produção do conhecimento.

Atualmente, conforme aponta Sayão (2005), grande parte dos documentos gerados já nasce digitalmente, mesmo não havendo “estratégias consolidadas para garantir o acesso de longo prazo às informações digitais” (SAYÃO, 2005, p. 116). Isso se dá pela facilidade em produzir – o que Sayão chama de compulsão – e compartilhar estes documentos os quais são infinitamente superiores à capacidade de preservá-los. A sociedade contemporânea traz em si a urgência do acesso, da produção e do compartilhamento. Talvez seja esta uma hipótese para que os documentos atuais nasçam digitalmente e os que não nasceram digitalmente, sejam convertidos a este novo formato. Assim, o surgimento das redes, em especial à Internet, e a utilização dos microcomputadores proporcionaram à criação de diversas formas de acesso à informação, mesmo que por outro lado também tenham proporcionados novos suportes/formatos e o crescimento da produção intelectual registrada.

Esta massificação da produção gera problemas que antes estavam restritos à recuperação da informação nos catálogos das bibliotecas. Com a “Web se tornando algo como uma única e imensa biblioteca digital” (MARCONDES, 2005, p. 109), cresce mais ainda a responsabilidade e preocupação com a recuperação e preservação deste conhecimento que cresce vertiginosamente entre as bibliotecas que gerenciam acervos.

Uma série de políticas que visam à modificação da relação entre os produtores da informação e os usuários vem sendo discutidas, dentre elas o *Copyleft* e o Movimento de Acesso Aberto. Este último é responsável pela criação de serviços, protocolos e políticas que visam, entre outros objetivos, à interoperabilidade dos provedores de dados e serviços. Um destes sistemas são os Repositórios Institucionais que

Inseridos no coração do movimento mundial em favor do acesso aberto à informação científica, [...] constituem de fato inovação no sistema de comunicação da ciência e no modo como a informação – aquela que alimenta e resulta das atividades acadêmicas e científicas – é gerenciada. (LEITE et al, 2012, p. 7).

Os repositórios institucionais são, segundo Santos Junior (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”.

Ou seja, destaca-se entre as formas atuais de dinamização e democratização do acesso à informação, por suas características genéricas – um tipo de coleção que conta com diversos tipos de documentos e formatos divulgados na Web, além de ser gerenciado por uma instituição, que garante uma verificação e tratamento da informação prestada.

Estes repositórios nascem da necessidade da criação de um espaço único de acesso às informações produzidas e sobre uma determinada instituição, portanto multidisciplinar, sendo um ponto confiável de coleta de dados e informações, haja vista ser um ambiente controlado, hierarquizado e ordenado, aliando as vantagens dos acervos digitais, como acesso aberto, por exemplo. Além disso, ele segue a lógica da Iniciativa dos Arquivos Abertos, onde os resultados de pesquisa produzidos com financiamento público através de agências de fomento, por exemplo, devem ser disponibilizados gratuitamente na Web, por meio dos RIs (SANTOS JÚNIOR, 2010).

De maneira geral, estes repositórios são implementados utilizando 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, tradução nossa).

Sendo assim, a baseado na filosofia da ICA, o DSpace é desenvolvido e representa o mais completo software livre para a criação de RIs, sendo amplamente utilizado em mais de 100 países<sup>1</sup>.

## 2.1 A PRESERVAÇÃO DIGITAL

A preservação é um fator primordial, tanto para os documentos impressos entre outros materiais, quanto para os digitais. A garantia de acesso futuro não pode ser considerada um item inerente à automatização dos processos, mas assim como nos documentos em papel, deve surgir a partir da “adoção de ferramentas que protejam e garantam a sua manutenção” (MÁRDERO ARELLANO, 2004, p. 17).

Esta preservação digital deve ser exercida sobre um documento digital que, conforme Vidal (2010, p. 147), “é uma unidade de registo de informação codificada em dígitos binários, acessível por meio de sistemas computacionais”. Esta unidade de registro pode ser imagem fixas ou em movimento, texto, página da web, banco de dados e até multimídias.

---

<sup>1</sup> Quantidade obtida no site [www.dspace.org](http://www.dspace.org), na seção who's using Dspace? [Quem usa o Dspace?]

A preservação digital é um esforço que “exige políticas, estratégias e ações para garantir, no decorrer dos anos, acesso a conteúdo válido apesar dos desafios impostos pelos suportes (mídias) e mudanças tecnológicas” (CORRÊA, 2010, p. 23). Não é, portanto, fruto de uma ação isolada, mas de um esforço institucional de quem pretende exercê-la, para garantir o acesso futuro aos documentos digitais. Estas políticas devem existir, preferencialmente, por escrito, pois independente de serem utilizadas individualmente, entre instituições que trabalham por cooperação, elas

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).

Ao contrário dos documentos que apresentam suportes materiais como o papel ou pergaminho, por exemplo, que em condições específicas podem ficar anos, décadas, séculos preservados, prontos para transferir a informação ali depositada a próxima pessoa que o acessar, os documentos digitais necessitam de vários elementos externos para realizar esta mesma transferência. Sendo assim, “para muitos documentos analógicos, a preservação do suporte e sua forma física garantem a conservação de seu conteúdo” (CORRÊA, 2010, p. 24), o que não pode ser garantido com os objetos digitais. Um livro pode ficar décadas sobre uma mesa – em condições específicas – e ser relido nos dias futuros, quando dificilmente uma música em uma fita K7 será ouvida hoje em dia, em lares onde já existem rádios com entradas USB. A preservação digital deve ter como principal meta a de garantir recuperação e visibilidade do documento digital no futuro (SANTHANAGOPALAN; FOX; McMILLAN, 2006, p. 1).

Isso tende a não acontecer “devido ao fato de que a informação tem um fator de dependência do seu ambiente técnico” (SANTHANAGOPALAN; FOX; McMILLAN, 2006, p. 1). Garantir a existência de um arquivo/formato pode não significar automaticamente a garantia de acesso aos arquivos nas mesmas condições em que foram produzidos, pois o software reprodutor daquele conteúdo pode ter sido atualizado, imputando modificações consideráveis a arquivo. Esta obsolescência natural ou provocada (GRANGER, 2002) traz uma grande fragilidade aos documentos nascidos digitalmente.

Miguel Arellano (2004) critica esta ênfase dada na geração e aquisição de materiais digitais na era da informação, sustentando que é necessário manter a preservação e acesso aos acervos digitais atuais. Com isso, vale lembrar, que atualmente alguns documentos digitais

são considerados documentos válidos ou registros oficiais, como, por exemplo, Nota Fiscal Eletrônica e Diários Oficiais, documentos estes gerenciados segundo leis e padrões específicos de emissão, valendo como documentos fiscais ou comprobatórios. Contudo, a temática da preservação digital até para estes documentos parece não ser relevante.

Também é necessário o entendimento da “constante obsolescência dos equipamentos e programas informáticos” (CUNHA, 1999, p. 256), que demanda dos bibliotecários mecanismos para um constante monitoramento da estrutura do objeto digital, a fim de garantir o acesso à informação digital em um suporte que mantenha a interação conforme o objeto digital foi criado. Assim, à medida que os sistemas computacionais são alterados, também os formatos que registram a informação digital devem ser mudados, mas nunca se afastando das funcionalidades originais. Esta obsolescência, seja por uma natural velocidade das transformações nos softwares e hardwares, seja pelo interesse no ganho comercial de alguns (GRANGER, 2002), reflete a velocidade em que esta constante transformação acontece e que o profissional da informação deve estar apto a gerenciar.

Na literatura é possível identificar diversas estratégias de preservação digital, contudo estas estratégias de preservação apresentam problemas que, conforme Santhanagopalan, Fox e Mcmillan no Brasil podem ser sintetizados em:

1. Falta de atividades relacionadas às normas de preservação;
2. Dados científicos estão sendo armazenados em bancos de dados de natureza proprietária;
3. Práticas atuais são feitas por meio da técnica de preservação tradicional. (SANTHANAGOPALAN; FOX; McMILLAN, 2006, p. 2, tradução nossa).

Estes pontos, ressaltados por Santhanagopalan, Fox e Mcmillan, resumem os desafios que esta pesquisa pretende demonstrar com as pesquisa documental e de campo.

## 2.2 ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL

As diversas estratégias de preservação digital (PD) existentes visam contribuir para a preservação do objeto digital e garantir o acesso futuro. Estas estratégias não comportam todos os itens previstos para a PD, sendo sempre focados em um item específico do tema, evidenciando a complexidade do problema da preservação digital. Parece não haver uma estratégia única capaz dar conta de todos os processos de preservação digital (SAYÃO, 2005), ou pelo menos uma fórmula única que seja capaz de ser utilizada. O que há são soluções específicas para casos específicos.

Estas estratégias descrevem detalhadamente ações específicas que fazem parte da preservação digital, apesar de serem vistas como a própria ação de preservação digital. Os

requisitos da preservação digital ainda não estão totalmente identificados ou consolidados na literatura, mas, conforme crê Granger (2002, p. 2) “esses requisitos incluem ferramentas, técnicas ou componentes de hardware e software” (tradução nossa).

Bullock (1999, tradução nossa) nos traz um conjunto de objetivos genéricos que temos que ter em vista no momento de traçar um planejamento de preservação digital, tais como:

- a) Fixe o documento com o um todo;
- b) Preserve a presença física<sup>2</sup>;
- c) Preserve o conteúdo;
- d) Preserve a apresentação;
- e) Preserve a funcionalidade;
- f) Preserve a autenticidade;
- g) Localize e referencie o objeto digital ao longo do tempo;
- h) Preserve a proveniência;
- i) Preserve o contexto.

Assim, fica evidente que, ao pensarmos em preservação digital, devemos ter em vista tanto o objeto digital – o arquivo em .pdf ou em qualquer outra extensão – quanto termos a preocupação com o padrão e metadados inseridos, para garantirmos a recuperação da informação no futuro.

Sendo assim, é possível identificar pontos comuns para a preservação digital como a necessidade de unidade na construção de uma política de preservação. Por vezes a iniciativa da preservação digital fica a cargo apenas da unidade de informação. Trata-se de uma preocupação coletiva, principalmente para instituições que utilizam-se do autodepósito de publicações em seus repositórios. Além disso, por vezes estas políticas não estão devidamente registradas ou divulgadas, gerando um ruído desnecessário na comunicação com os responsáveis pelo povoamento do repositório, podendo também sobrecarregar a unidade de informação, na padronização de entradas e correções das mesmas.

Na literatura é possível identificar diversas estratégias de preservação digital, sendo as citadas abaixo as mais consolidadas pela literatura em PD, a saber:

---

<sup>2</sup> Bullock explica que a presença física corresponde, neste caso, aos códigos binários – a série de "1" e "0" – que são à base de um objeto digital.

- a) Preservação da tecnologia: baseia-se na preservação do equipamento tecnológico adequado/próprio para o acesso aos objetos digitais, que garantam a visualização do documento de acordo como foi desenvolvido (SAYÃO, 2010).
- b) Emulação: baseia-se na emulação (possibilitar o funcionamento do software original em uma plataforma tecnológica mais avançada). Assim o objeto digital pode ser visualizado como desenvolvido em um ambiente que reproduz as especificações técnicas de software e hardware originais (SAYÃO, 2010).
- c) Migração: é definida por Hedstrom (2001, p. 3) como "um conjunto de tarefas organizadas para a realização da transferência periódica de materiais digitais a partir de uma configuração de hardware/software de um computador para outro ou de uma geração de tecnologia de computador para a geração seguinte" (tradução nossa).
- d) Encapsulamento: trata-se de autodescrever os objetos digitais e encapsulados "em estruturas físicas ou lógicas com todas as informações necessárias para que seja decifrado e compreendido no futuro" (SAYÃO, 2010, p. 10).

Contudo, realizando uma pesquisa mais ampla, é possível identificar outras estratégias de PD, como, por exemplo, os metadados, que podem receber diversas nomenclaturas e serem classificados em várias classes também.

No Dspace estes metadados são classificados como metadados descritivos, estruturais e administrativos (MÁRDERO ARELLANO, 2008). Cada um destes níveis é responsável por um tipo de informação que garantirá a recuperação do objeto digital (no caso dos metadados descritivos), na forma como a apresentação do documento é preservada (metadados estruturais) e na garantia de fidelidade e autenticidade (metadados administrativos).

Sayão e Sales (2012) estabelecem a mesma classificação dos metadados no contexto da Curadoria Digital, conceituada pelos autores como o conjunto dos processos de PD presente no ciclo de vida dos objetos digitais, como na descrição e representação da informação, o planejamento da preservação, a participação e monitoramento da preservação dos objetos digitais (SAYÃO; SALES, 2012).

Assim, a Curadoria Digital seria uma estratégia que se utilizaria de diversos recursos, tanto dos metadados, quanto do armazenamento digital para assegurar a sustentabilidade dos dados. Este conjunto de ações não existe isoladamente, como uma ferramenta dentro dos RIs, mas como parte integrante deles.

Na Curadoria Digital, os metadados assumem o papel descritivo dos objetos digitais, ampliando o papel da descrição das informações nos documentos digitais. Sendo possível, então, imaginar novas dimensões para o uso destes metadados, avançando de uma visão técnica para gestão dos objetos digitais como um todo, como a inclusão do controle dos direitos, intercâmbio de informações, interoperabilidade técnica e semântica e reuso da informação (SAYÃO, 2010, p. 3).

Para preservar o objeto digital, Sayão (2010) consolidou três níveis de metadados: descritivos, estruturais e administrativos.

- e) Os metadados descritivos: representam os metadados mais conhecidos como título, autor e palavras-chave, que servem para descrevermos os objetos digitais.
- f) Os metadados estruturais: auxiliam-nos na estruturação dos documentos digitais, preservando a sua apresentação – ordem como os elementos são apresentados, por exemplo.
- g) Os metadados administrativos: ajudam-nos no processo de gestão no ciclo de vida dos objetos digitais, possibilitando a inclusão de informações de criação, especificidades e dependências técnicas – como qual o programa e versão dos objetos digitais foram criados para ser utilizado.

Também é possível detectar na literatura o uso dos identificadores persistentes como uma estratégia de PD. O identificador persistente possibilita a criação de um endereço fixo (*link*) ao objeto digital que “deverá perdurar por um período tão longo quanto seja necessário; mesmo que a organização que o atribuiu ao objeto não mais exista quando este for usado” (SAYÃO, 2007b, p. 67). O objetivo da criação destes identificadores persistentes é garantir o acesso ao objeto digital mesmo que este arquivo seja “movido, ou sua propriedade transferida” (SAYÃO, 2007b, p. 68) fazendo que eles continuem ativos na web.

Como os métodos mais comuns de atribuição de link persistente, podemos citar:

- h) URN: um link formado pelo <urn>:<Namespace Identifier (NID)>:<Namespace specific string (NSS)>, sendo o NID uma numeração atribuída pelo Internet Engineering Task Force (IETF) à instituição responsável pelo objeto digital. Devem obedecer a alguns critérios básicos como: validade e unicidade global e persistência. (SAYÃO, 2007b).

Exemplo: urn:lex:br:federal:lei:2008-06-19;11705

i) PURL: é um esquema desenvolvido pela Online Computer Library Center (OCLC), composto por: <protocolo>/<endereço do serviço de resolução>/<nome do recurso>. (SAYÃO, 2007b)

Exemplo: <http://purl.oclc.org/OCLC/PURL/FAQ>

j) HANDLE: criado sob as mesmas bases que o URN, trata-se de um sistema de informação criado pela Corporate for National Research Initiatives (CNRI) baseado em uma hierarquia de servidores que se comunicam através de protocolos abertos e que contam com servidores cache intermediários. Ele divide-se em Global Handle Registry – operado pela CNRI – e Servidores Locais de Handles (LHS - Local Handle Services). O seu endereço é consiste em “prefixo” e “sufixo”. O prefixo é usado para identifica à instituição autorizada pela CNRI para o uso do, e o sufixo o nome do recurso, ficando: <Handle>:=<Handle Name Authority>/<Handle Local Name>. (SAYÃO, 2007b).

Exemplo: <http://hdl.handle.net/10183/37473>

k) DOI: Digital Object Identifier (DOI) “é uma aplicação específica do Handle System, voltada para a identificação persistente de recursos digitais aos quais possam ser atribuídos direitos de propriedade intelectual, bem como para o intercâmbio de informações sobre essas propriedades em um ambiente de rede” (SAYÃO, 2007b, p. 73). O endereçamento é formado pelo <Código do Diretório>.<Código de Registro>/<Seqüência Sufixo DOI> (SAYÃO, 2007b).

Exemplo: 10.5965/1984723814262013199

Também é possível identificar o Backup como uma estratégia de preservação digital, bastante difundida no dia-a-dia, por, conforme afirma Frigo

Fazer backups de documentos digitais é garantir sua preservação, segurança e a confiabilidade da informação registrada. Esta ação garantirá que a informação do documento digital não se perca, para que seja possível sua utilização no presente e no futuro (FRIGO, 2012, p. 93).

Esta estratégia permite garantir o acesso ao objeto digital, caso ele seja danificado.

Outra estratégia de preservação digital disponível são as que utilizam redes P2P (Peer-to-Peer, em tradução livre Ponto-a-Ponto). Baseia-se na criação de infraestruturas geograficamente distantes que trabalham em colaboração, realizando backups em rede dos objetos digitais de cada ator da rede de distribuição, permitindo a sincronização no caso de um objeto digital seja danificado, conforme explica Reich sobre as

Redes de Distribuição de Preservação Digital tem infraestruturas geográficas e organizacionais intencionalmente distribuídas como componentes

essenciais de concepção e exigem a colaboração da comunidade e da cooperação para a implementação. Não é possível alcançar a robustez necessária para garantir a persistência a longo prazo de objetos digitais através de abordagens técnicas ou de organização centralizados (REICH, 2012, p. 76, tradução nossa).

Assim, assegura-se a que haverá sempre uma cópia dos objetos digitais em uma rede elétrica diferente – no caso de inacessibilidade por falta de energia – ou em um servidor diferente – no caso de inacessibilidade por avaria do equipamento (Hardware) do servidor.

Outra importante questão que merece ser levantada em estratégias de preservação digital é a escolha dos tipos de arquivos que povoarão os RIs. Arquivos em extensão que utilizam software proprietário ou que não possibilitam a edição de metadados, por exemplo, podem dificultar o desenvolvimento da estratégia de PD e não garantir o acesso futuro aos objetos digitais.

Para esta estratégia, Munhoz (2011) sinaliza os principais formatos, levando-se em consideração a natureza do objeto digital, assim estruturados:

Quadro 1 – Formatos de Texto - Alta adequação para preservação

Extensão	Formato
HTML, XHTML ou XML*	* Contendo referência para o DTD ou esquema de validação de formato
TXT*	* Textos planos usando codificação de caracteres UTF-8, USASCII ou UTF-16 com Byte Order Mark
PDF*	*PDF/A-1 segundo a norma ISO 19005-1
ODT	Open Document text

Fonte: Munhoz (2011, p. 29)

Quadro 2 – Formatos de Texto – Média adequação para preservação

Extensão	Formato
HTML, XHTML ou XML*	* Sem referência para o DTD ou esquema de validação de formato
CSS	Cascading Style Sheets
TXT*	* Texto plano com codificação de caracteres em ISO-8859-1

continua

Extensão	Formato (continuação)
PDF*	*com fontes embebidas
RTF	Rich Text Format 1.x
HTML*	* HTML 4.x incluindo declaração de DOCTYPE
SXW	Open Office Text Document
DOCX	Office Open XML
DTD	DTD
SGML	SGML

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

Quadro 3 – Formatos de Texto - Baixa adequação para preservação

Extensão	Formato
DOC	Microsoft Word
OS	Postscript
PDF*	*PDF cifrado
DVI	DVI

Fonte: Munhoz (2011, p. 30)

Quadro 4 – Formatos de Imagem - Alta adequação para preservação

Extensão	Formato
PNG	PNG
JP2*	*JPEG2000 usando compressão sem perda de informação
TIF*	*TIFF sem compressão
SVG	SVG

Fonte: Munhoz (2011, p. 30)

Quadro 5 – Formatos de Imagem - Média adequação para preservação

Extensão	Formato
JP2*	*JPEG2000 usando compressão com perdas de informações

continua

Extensão	Formato (continuação)
GIF	GIF
JPG	JPEG/JFIF
TIF*	*TIFF com compressão
BMP	BMP
DNG	Digital Negative
CGM	Computer Graphic Metafile e WebCGM

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

Quadro 6 – Formatos de Imagem - Baixa adequação para preservação

Extensão	Formato
SWF	Macromedia Flash
PSD	PhotoShop
JPF, JPX	JPEG 2000 Part 2
SID	MrSID
TIFF*	*TIFF em formato Planar
FPX	FlashPix
RAW*	*RAW (a extensão pode variar de acordo com o dispositivo de captura)
EPS	Encapsulated Postscript

Fonte: Munhoz (2011, p. 31)

Quadro 7 – Formatos de Áudio – Alta adequação para preservação

Extensão	Formato
AIF, AIFF	AIFF com Pulse-code modulation
WAV, BWF	WAV com Pulse-code modulation
OGG, OGA	Ogg Vorbis

Fonte: Munhoz (2011, p. 31)

Quadro 8 – Formatos de Áudio - Média adequação para preservação

Extensão	Formato
MP3	MP3 (MPEG-1/2, Layer 3)
FLAC	Free Lossless Audio Codec
AU	SUN Audio sem compressão
MID, MIDI	Standard MIDI
MP4, M4A, AAC	Advance Audio Coding

Fonte: Munhoz (2011, p. 31)

Quadro 9 – Formatos de Áudio – Baixa adequação para preservação

Extensão	Formato
RA, RM, RAM	RealNetworks 'Real Audio'
WMA	Windows Media Audio
WAV	WAV com compressão
AIFC	AIFC com compressão
SND	NeXT SND

Fonte: Munhoz (2011, p. 32)

Quadro 10 – Formatos de Vídeo – Alta adequação para preservação

Extensão	Formato
MOV	QuickTime Movie sem compressão
AVI	AVI sem compressão
MJ2	Motion JPEG 2000 (ISO/IEC 15444-4)
AVI, MOV	Motion JPEG

Fonte: Munhoz (2011, p. 32)

Quadro 11 – Formatos de Vídeo – Média adequação para preservação

Extensão	Formato
MPG, MPEG	MPEG-1, MPEG-2
MP4	MPEG-4
OGG, OGM, OGV	Ogg Theora

Fonte: Munhoz (2011, p. 32)

Quadro 12 – Formatos de Vídeo – Baixa adequação para preservação

Extensão	Formato
WMV	Windows Media Video
AVI	AVI com compressão
MOV	QuickTime Movie com compressão
RV, RM	RealNetworks 'Real Video'

Fonte: Munhoz (2011, p. 32)

O Modelo de Referência Open Archival Information System (OAIS) também pode ser indicado como uma estratégia de preservação digital, na medida em que estrutura um modelo para o fluxo do objeto digital dentro do RI, visando realizar o monitoramento do processo de preservação digital como um todo.

O modelo, composto por “cinco componentes [...]: planejamento de preservação, ingestão, gestão dos dados, acesso e administração” (CASTRO et al, 2009, p. 287), sinaliza que o objeto digital – denominado de *Submission Information Package* (SIP) em inglês e de Pacote de Informação de Submissão (PIS) em tradução livre – para ser inserido no repositório precisa ser submetido a uma instância de ingestão (*ingest*), que é uma interface que permite a incorporação do PIS ao repositório.

Esta ingestão pode ser feita diretamente (autodepósito), repassada ao gestor da RI ou feita diretamente com conferência dos dados posteriormente. Neste momento as informações descritivas (metadados) sobre este documento são geradas e, junto com o objeto digital, compõem o *Archival Information Package* (AIP), ou Pacote de Informação de Arquivo (PIA) em tradução livre.

A partir daí, o objeto digital passa a ser incorporado a um local específico para armazenamento (*Archival Storage*) e os dados também (*Data Management*), passando a incorporar um registro na área de acesso. Assim como a ingestão (*ingest*) é um componente de interface, o acesso (*access*) também é, mas este regula a busca do usuário através de estratégias de busca nos metadados atribuídos ao AIP, conforme explica Castro (et al, 2009, p. 287): “este componente é responsável pela recuperação da metainformação e do material preservado, agora denominado Pacote de Informação de Disseminação (PID) [*Dissemination Information Package* (DIP)] que será consumido pela entidade externa”.

Este modelo pode ser visualizado através da figura abaixo:

Figura 1 - Modelo de referência OAIS



Fonte: CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (2002, p. 4-1).

Estas estratégias devem ser escolhidas levando em conta quesitos técnicos, financeiros e estruturais, mas sempre de acordo com a política de preservação.

Em relação às estratégias de preservação digital que ocorrem no nível dos RIs Miguel Arellano e Fernando Leite conceituam como “*small-p preservation* porque eles permitem o acesso imediato hoje, amanhã e no futuro, com o objetivo exclusivo de aumentar o uso e o impacto hoje, amanhã e no futuro” (MÁRDERO ARELLANO; LEITE, 2009, p. 8).

Em contraponto está à “*large-P Preservation*”, que seria a “preservação digital como um fim em si mesma” (MÁRDERO ARELLANO; LEITE, 2009, p. 7) que é de responsabilidade dos editores, bibliotecas assinantes ou bibliotecas depositárias, que devem garantir a perenidade dos produtos que venderam ou que adquiriram (MÁRDERO ARELLANO; LEITE, 2009).

A diferença entre estas formas de preservação é o auto arquivamento dos documentos. Enquanto no “*small-p preservation*” o que é salvaguardado são as cópias dos documentos digitais publicados em revistas, jornais e afins, na “*large-P preservation*” as políticas são

aplicadas aos originais. Há então um conflito de interesses, pois “o propósito do autoarquivamento para o acesso aberto não é preservar o artigo do editor nem substituí-lo [...] [mas sim] prover um suplemento dele, uma outra versão para aqueles que poderiam ser usuários e que não podem pagar pelo acesso” (MÁRDERO ARELLANO; LEITE, 2009, p. 9).

### 2.3 INICIATIVAS DE PRESERVAÇÃO DIGITAL

As iniciativas são ações conjuntas que visam produzir uma maneira racional e colaborativa de preservação digital. O esforço pela preservação digital envolve diversos fatores, sendo possível que fatores financeiros possam ser sinalizados como obstáculos execução de uma política de preservação digital. As iniciativas de preservação digital surgem como um auxílio na tarefa extremamente trabalhosa e custosa para apenas uma instituição desenvolver.

#### 2.3.1 LOCKSS

Dentre estas iniciativas, podemos citar o LOCKSS (*Lots of Copies Keeps Stuff Safe* – em inglês, Muitas Cópias Mantém os Materiais a Salvo, em tradução livre) que é um software de código aberto que permite a criação de Redes Privadas Locais, em tradução livre para *Private LOCKSS Network* em inglês (PLN). Estas PLNs “preservam cópias de conteúdos digitais criados localmente por organizações científicas e culturais”. (MÁRDERO ARELLANO, 2012, p. 85).

O LOCKSS atua como um compartilhador de documentos “*peer-to-peer*” entre as instituições que se associam, formando as PLN. Uma vez dentro da PLN, as instituições devem configurar sua LOCKSS Box em seu servidor, que é onde deverão ser alocados os arquivos digitais a serem coletados, guardados e preservados localmente. Isto feito, o LOCKSS cria backups destes arquivos em todas as instituições que formam a PLN, gerenciando estes arquivos, substituindo-o em caso de perda do objeto digital ou atualização. Assim como as demais instituições recebem os objetos digitais de uma unidade da PLN, esta unidade também recebe os arquivos das demais unidades, criando uma rede com diversos pontos de backup e sincronização.

O LOCKSS atende à metodologia de *Distributed Digital Preservation* – DDP (Preservação Digital Distribuída, em tradução livre), que visa estabelecer critérios mínimos para a preservação digital distribuída. Este tipo de PD leva em consideração a dispersão

geográfica de backups, a fim de garantir acesso e recuperação destes dados, independente de qualquer falha ou avaria física de uma instituição. Para isso, Skinner e Schultz (2010) trazem uma sequência de melhores práticas para este tipo de preservação, das quais destacamos:

- a) Sites preservando o mesmo conteúdo não devem estar dentro de um raio de 75 a 125 quilômetros um do outro;
- b) Locais de preservação devem ser distribuídos em diferentes redes de energia;
- c) Conteúdo preservado em locais diferentes deve estar online e devem ser verificados regularmente para evitar *bit-rot*<sup>3</sup> e outros problemas;

### 2.3.2 Rede Cariniana

A Rede Cariniana é uma iniciativa que tem o “objetivo de propor a construção e customização de uma rede nacional de serviços de preservação digital” (MÁRDERO ARELLANO, 2012, p. 87), contando com a parceria de instituições de ensino e pesquisa e com a colaboração de especialistas brasileiros, nesta construção. Sua atuação reforça a necessidade da criação de iniciativas de preservação que garantam a salvaguarda dos registros digitais em âmbito nacional.

Trata-se de uma iniciativa que busca orientar as instituições na prática da preservação digital de publicações de acesso aberto, oferecendo “pacotes de softwares, aplicações e ambientes multimídia para a implementação e desenvolvimento de documentos digitais preserváveis” (MÁRDERO ARELLANO, 2012, p. 87).

A operacionalização da rede é baseada no LOCKSS, adotando o modelo de DDP com pretensões de atingir redes de âmbito nacional. Proporciona um espaço para a busca de melhores metodologias e práticas em preservação digital, de forma colaborativa.

## 2.4 OS REPOSITÓRIOS INSTITUCIONAIS

Os Repositórios Institucionais (RIs), tal como conhecemos hoje, são sistemas criados a partir do “final do ano 2000” (MÁRDERO ARELLANO, 2008, p. 125), que possui a “capacidade de manter e gerenciar materiais por longos períodos de tempo e prover o seu acesso apropriado” (MÁRDERO ARELLANO, 2008, p. 124).

---

<sup>3</sup> Trata-se de um fenômeno de mídia de armazenamento que pode deteriorar os *bits* gradualmente ao longo do tempo

Conforme Miguel Arellano (2008) os RIs podem ser divididos em 3 grupos:

- a) Temático: cobre uma área do conhecimento;
- b) Institucionais: São repositórios utilizados para armazenar, preservar, divulgar à produção de instituições em formato digital;
- c) Centrais: “provedores de serviços nacionais e internacionais que permitem a reunião de dados coletados tanto de bibliotecas digitais, quanto de repositórios temáticos e repositórios institucionais” (MÁRDERO ARELLANO, 2008, p. 124).

Sendo os dois primeiros entendidos como provedores de dados e o terceiro como provedor de serviços.

A partir da década de 1990, diferentes organizações começaram a armazenar eletronicamente trabalhos publicados entre outros documentos, delimitando os primeiros passos do Movimento do Acesso Aberto. Com o passar do tempo, o Movimento do Acesso Aberto objetivou estes repositórios como estratégia para promover o acesso aberto por meio das seguintes ações:

[...] para facilitar o autodepósito das versões dos trabalhos (avaliados ou não) dos autores, permitindo um acesso rápido para aqueles que não possuem condições de pagar uma assinatura. (MÁRDERO ARELLANO, 2008, p. 34)

Hoje vivemos em meio a diversos documentos digitais, e a própria necessidade de ordená-lo pressionou ao desenvolvimento dos repositórios digitais, que são recursos que visam auxiliar na dinamização e organização deste vasto conteúdo. Além disso, de acordo com Rezende e Abreu, nos dias atuais os sistemas estão orientados pela analogia “objetivo-entrada-transformação-saída-feedback” (REZENDE; ABREU, 2003, p. 33), diferente a analogia “entrada-processamento-saída”.

Cabe ressaltar também que as ferramentas para RIs não são destinadas exatamente para a finalidade de preservação digital (MÁRDERO ARELLANO; LEITE, 2009). Estas novas ferramentas possibilitam, também, a preservação digital, mas não foram criadas para este propósito em especial. Ou seja, os conteúdos podem até ser preservados, mas a atividade fim desta ferramenta não é ser uma única estratégia de preservação digital, mas parte de uma política de preservação, a ser desenvolvida.

De acordo com Miguel Arellano (2008), “os repositórios institucionais podem ser pensados como parte de uma ação dentro da gestão documental, que inclui sua criação, tratamento, transmissão e acesso”. Assim, pensar nos RIs como um fim sob a perspectiva da preservação digital pode tornar a tarefa da preservação deficiente e falha. De fato os RIs

constituem um importante espaço de apoio à produção científica, provendo acesso rápido e gratuito aos documentos, resultados de pesquisas, teses e trabalhos realizados por uma comunidade, favorecendo as discussões e ao processo de comunicação científica, mas não representam, em si, uma estratégia de preservação digital completa.

## 2.5 O DSpace

O DSpace é um software de código aberto desenvolvido em 2002 pela Massachusetts Institute of Technology (MIT) e a Hewlett-Packard (HP) com o objetivo de “prover acesso livre, recuperável e coerente aos materiais de ensino do MIT” (MÁRDERO ARELLANO, 2008, p. 139 apud DSPACE, 2002) e hoje se consolida como uma ferramenta flexível na criação de Repositórios Institucionais ou multidisciplinares para bibliotecas, arquivos e centros de pesquisa e informação.

Estas instituições devem gerenciar um grande fluxo de informações geradas diariamente, em diferentes formatos e ao mesmo tempo prover acesso às estas informações na Web, encontrando no sistema DSpace uma ferramenta consistente para esta operação, conforme pontam Blattmann e Weber quando dizem que

O sistema DSpace possibilita gerenciar os diferentes e complexos recursos no formato digital (padrões de documentos como TIFF, AIFF, XML, ou publicados como especificações PDF, RIFF) para coletar, preservar, indexar e distribuir os itens digitais de ambientes acadêmicos (departamentos, laboratórios, centros, escolas, programas) (BLATTMANN; WEBER, 2008, p. 467).

Estas características que são evidenciadas por Miguel Arellano. Para o autor, o DSpace suporta “uma variedade maior de tipos de materiais [...] [também] possui mais opções de preservação de objetos digitais” (MÁRDERO ARELLANO, 2008, p. 138).

A possibilidade de atender os diversos tipos de unidades de informação é graças à possibilidade de inclusão de grande uma variedade de tipos de documentos, associados a uma ampla possibilidade de metadados descritivos, estruturais e principalmente administrativos.

O software é constituído sob os princípios da Iniciativa de Código Aberto, operando com “padrões já aceitos internacionalmente, como o *Open Archive Initiative* (OAI), o *Open Archival Information System Reference Model* e o *Dublin Core*” (MÁRDERO ARELLANO, 2008, p. 140) colaborando não apenas com a preservação digital dos documentos inseridos em seu banco de dados, mas também com a interoperabilidade para importação ou exportação de seus dados.

No DSpace são encontrados os metadados descritivos, estruturais e administrativos, sendo cada categoria baseada em algum conjunto específico de metadados. Os metadados descritivos gerais são baseados na *Library Application Profile for Elements and Qualifiers*, mas é configurável a vários outros padrões, se necessário. Ele se subdivide em dois outros níveis, os metadados descritivos para as comunidades e outro para as coleções, que utilizam metadados derivados do Dublin Core (MÁRDERO ARELLANO, 2008).

No DSpace a preservação digital ocorre no nível dos *bytes*. Os *bitstreams* – sequências de *bytes* dos arquivos digitais – são armazenados e identificados através de seus metadados com o objetivo de possibilitar o acesso a longo prazo, conforme apontam Blattmann e Weber (2008)

A importância dos metadados é justificada pela necessidade dos documentos disponíveis estarem codificados em formatos e linguagens de marcação padronizadas, de modo a facilitar a interoperabilidade, o acesso e a manutenção, a recuperação e localização eficaz do acervo digital para o usuário (BLATTMANN; WEBER, 2008, p. 473).

De uma maneira geral, o DSpace consegue unir vantagens de gerenciamento e identificação dos objetos digitais, ainda aplicando processos de preservação e gerenciamento, conforme aponta Miguel Arellano que

As instituições que adotam o DSpace conseguem, além de identificar quais são os arquivos ou coleções que precisam de maior proteção, realizar um planejamento sobre a manutenção de cópias de software e hardware e sua documentação necessária para identificar os tipos de arquivo indispensáveis para o uso dos objetos digitais no futuro. (MÁRDERO ARELLANO, 2008, p. 143).

A criação destas coleções se dá pela face mutável do software, que permite ao usuário do sistema criar e editar comunidades e coleções, classes de documentos digitais, atribuir metadados aos níveis de operação e, também, gerenciar a preservação digital de seus documentos.

### 3 ESTUDO DE CASO

Para atingir o objetivo de identificar as melhores práticas em Preservação Digital (PD) nos Repositórios Institucionais (RIs) das instituições públicas de ensino e de pesquisa na região sudeste e de que maneira estas ações estão formalizadas como política, empregaremos duas metodologias para o estudo de caso: pesquisa documental, para a coleta de dados nos RIs sobre as informações básicas referente às políticas e demais aspectos gerais da estratégia de PD e questionário com questões fechadas para coletar dados mais específicos. Depois, será realizada uma análise quantitativa das respostas para identificar as boas práticas em relação à PD em RIs.

#### 3.1 OS REPOSITÓRIOS SELECIONADOS

A lista de RIs selecionados para a amostra da pesquisa foi baseada nos critérios de instituições elegíveis do Edital FINEP/PCAL/XBDB 003/2009, presentes na região sudeste do Brasil. Estes critérios delimitam um escopo do perfil das instituições selecionadas pela amostra, tendo em vista a diversidade de assuntos e atividades dos Repositórios Institucionais existentes, garantindo a homogeneidade das instituições selecionadas e o uso de uma ferramenta comum: o DSpace.

Os nove repositórios selecionados para a amostra estão assim distribuídos:

Quadro 13 – Amostra de Repositórios Institucionais

RIs	UF	Tamanho da Coleção
1	SP	43.350
2	MG	2588
3	SP	338
4	SP	1.374
5	RJ	425
6	MG	4.237
7	SP	126
8	SP	231
9	RJ	225

Fonte: O autor (2014)

Inicialmente o universo da amostra seria de 12 (doze) RIs, localizado em busca pela Internet que atendiam aos critérios do edital de fomento aos RIs, contudo 03 (três) não estavam acessíveis no momento da pesquisa, todos do estado de Minas Gerais.

### 3.2 PESQUISA DOCUMENTAL

A pesquisa documental baseou-se na visita ao endereço eletrônico dos 9 (nove) Repositórios Institucionais selecionados e listado acima para avaliar a disponibilidade de recursos e informações básicas aos usuários dos sistemas. Foram selecionados 6 (seis) itens baseados nos critérios do edital de fomento aos RIs e na literatura técnica referente ao assunto, com o objetivo de identificar a coleção, as política e estratégias de preservação digital. (Apêndice A).

### 3.3 COLETA DE DADOS PELO QUESTIONÁRIO

Foi desenvolvido um questionário com 12 (doze) questões fechadas, elaboradas com o objetivo de identificar as práticas das instituições com relação à preservação digital.

O objetivo do questionário é evidenciar as práticas referentes à preservação digital, com as questões agrupadas em três blocos: identificação da instituição, política de preservação digital e estratégias de preservação digital. (Apêndice B).

O questionário foi desenvolvido através de um formulário online, utilizando recursos online – Google.docs – para a tabulação dos resultados, facilitando a coleta de dados e permitindo o acompanhamento das respostas em tempo real.

O questionário foi aplicado no período de um mês. Apesar da receptividade dos bibliotecários em relação ao questionário, em alguns casos a resposta foi enviada após troca de diversas mensagens sobre a necessidade do levantamento para a conclusão desta pesquisa. Relato a falta de informações na Internet para opções ao atendimento online, como contato telefone via telefone.

### 3.4 TABULAÇÃO E ANÁLISE DOS DADOS

A tabulação dos dados coletados foi realizada rapidamente, devido ao formulário digital utilizado na coleta dos dados.

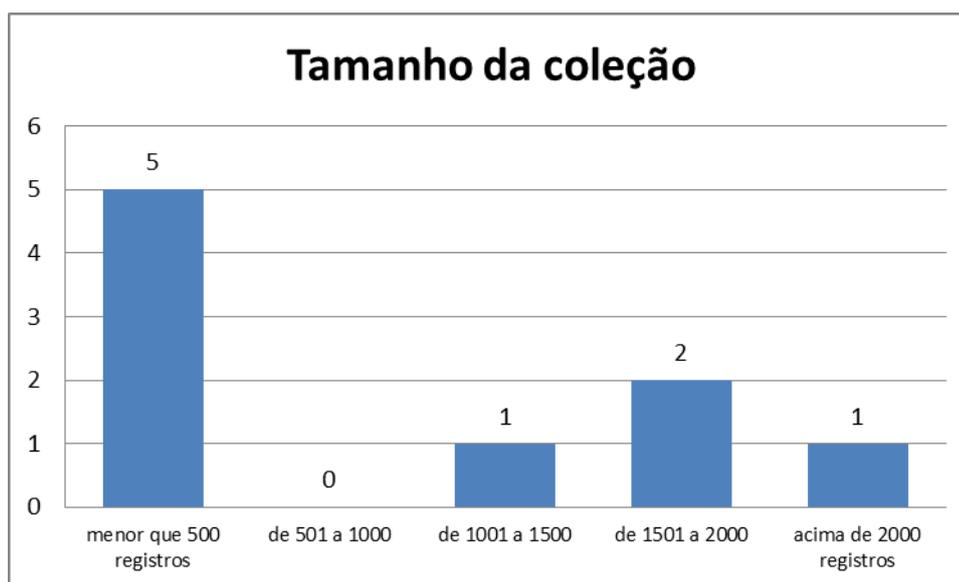
Os resultados apresentados correspondem a uma amostra total de 9 (nove) repositórios da região sudeste, de instituições públicas de ensino e de pesquisa com atuação em território

nacional, mantenedora de programas de pós-graduação reconhecidos pelo MEC/CAPES, conforme o critério de instituições elegíveis do Edital FINEP/PCAL/XBDB 003/2009. Os resultados serão apresentados em duas seções: pesquisa documental e questionário.

O primeiro ponto a ser avaliado na pesquisa documental diz respeito ao tamanho da coleção. Foi possível identificar um grande número de RIs que apresentam uma coleção menor que 500 registros, divergindo do resultado aferido no questionário – respondido pelos gestores dos RIs – conforme veremos mais a frente.

Estes resultados foram obtidos através do somatório dos valores encontrados nos endereços eletrônicos dos repositórios selecionados na amostra. Cada comunidade inserida no repositório traz o total de registros cadastrados em sua categoria, formando assim o número total de registros daquele repositório. Através deste somatório manual foi possível estimar a quantidade total de registros e comunidades, dados responsáveis pela criação do quadro 14, que veremos a seguir.

Gráfico 1 - Pesquisa documental - Tamanho da coleção



Fonte: O autor (2014)

O DSpace permite a criação de comunidades, possibilitando uma organização dos diversos temas que uma instituição pode possuir. Apenas 2 (duas) médias entre o número de registros e comunidade ultrapassam a marca de 500 registros, levando-se em consideração o tamanho da coleção e comunidades aferidos na pesquisa documental. Assim, verifica-se que a produção média por comunidade depositada nos RIs encontra espaço para evolução, haja vista

que o banco de dados do DSpace está preparado para receber uma quantidade bem maior de registros.

Quadro 14 – Média entre registro e comunidades

RI	UF	Tamanho da coleção	Comunidades	Média
1	SP	43.350	55	788
2	MG	2588	14	185
3	SP	338	11	31
4	SP	1.374	6	229
5	RJ	425	1	425
6	MG	4.237	8	530
7	SP	126	1	126
8	SP	231	1	231
9	RJ	225	9	25

Fonte: O autor (2014)

O segundo item avaliado refere-se à indicação das licenças do *Creative Commons*. Como o critério 1.4 (item a) para as instituições elegíveis do Edital FINEP/PCAL/XBDB 003/2009 solicita que a instituição manifeste o compromisso de promover e viabilizar o modelo de interoperabilidade *Open Archives* é desejável que haja estas licenças. Entretanto, 67% dos repositórios visitados não indicam usar as licenças do *Creative Commons*, conforme pode ser percebido pelo gráfico abaixo. É importante ressaltar que não é possível afirmar que as licenças são de fato utilizadas ou não, pois não fica explícito.

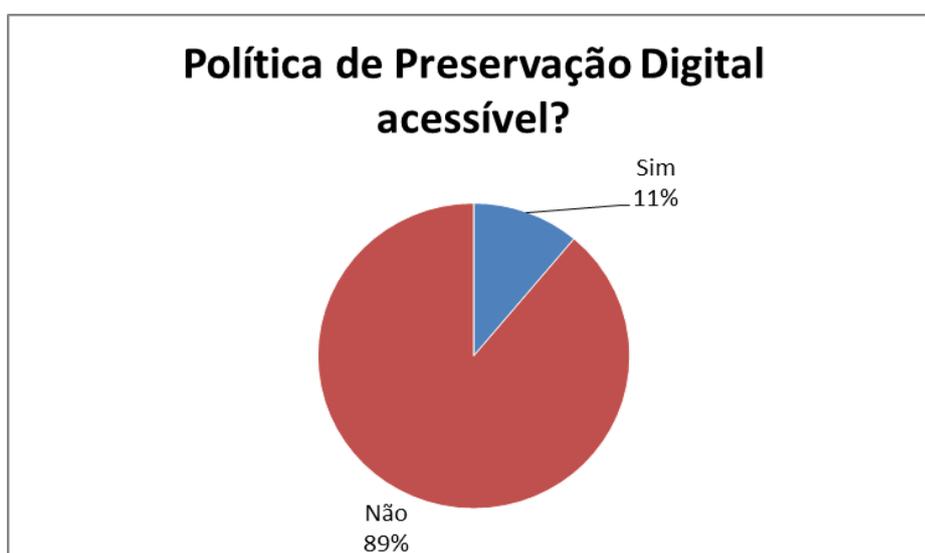
Gráfico 2 - Pesquisa documental – Creative Commons



Fonte: O autor (2014)

Outro item observado na pesquisa documental foi a publicidade da política de preservação digital. 89% dos repositórios visitados não tornam acessíveis suas políticas de preservação digital. Assim como no resultado anterior, vale ressaltar que este resultado demonstra que não é possível acessar a política de preservação. Contudo, baseado também no resultado do questionário referente a este assunto, não é possível descartar a possibilidade que esta política de fato não esteja formalizada. O único RI que declarou possuir uma política de preservação digital, sinalizou que esta política é institucionalizada.

Gráfico 3 - Pesquisa documental – Política de preservação digital



Fonte: O autor (2014)

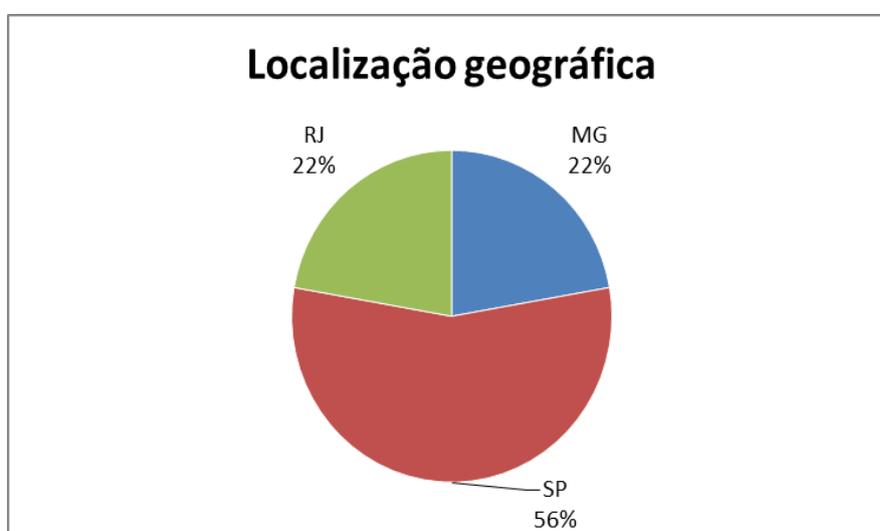
O quinto item analisado corresponde à indicação aos formatos autorizados para inclusão no RI. Não foi possível identificar em nenhum dos RIs pesquisados quais são, de fato, os formatos de arquivos disponíveis e passíveis de povoamento naqueles repositórios, sendo possível mensurar estes apenas no questionário.

Do mesmo modo, foi verificado no sexto item avaliado, a não identificação a participação em redes de preservação distribuída na página inicial em 100% dos RIs pesquisados. Mesmo com 20% dos RIs pesquisados declarando realizar este modo de preservação digital, como veremos a seguir, não é possível identificar de maneira clara esta parceria no endereço eletrônico dos RIs pesquisados.

O último item analisado na pesquisa documental foi à apresentação de link persistente nos endereços dos objetos digitais, onde foi possível identificar um uso de 100% do Handle System.

Os dados apresentados a seguir foram captados através de questionário fechado com 9 (nove) RIs, durante o estudo de caso, no momento da pesquisa documental. Dos quatro Estados da federação que compõe a região sudeste do Brasil, apenas o Espírito Santo não possui RI contemplado na amostra. Então, a distribuição geográfica dos RIs pesquisados é dividida entre: 02 unidades no Rio de Janeiro, 02 em Minas Gerais e 05 no Estado de São Paulo, sendo distribuídos proporcionalmente, conforme o gráfico abaixo.

Gráfico 4 - Questionário - Localização geográfica

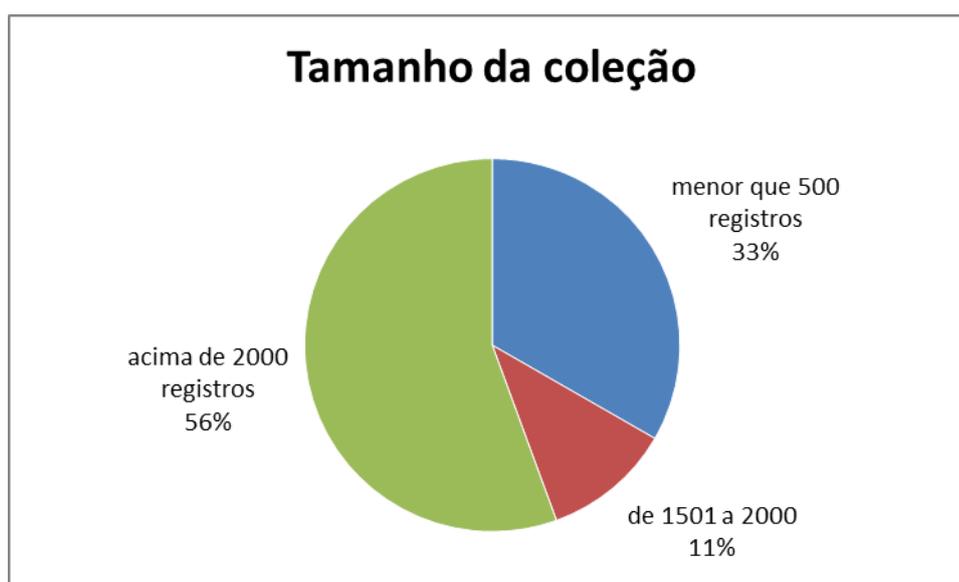


Fonte: O autor (2014)

Os 56% indicam uma concentração de RIs no estado do São Paulo. Este percentual é composto também pelas unidades no interior do estado, com 22% para os estados de Minas Gerais e Rio de Janeiro.

O questionário revelou que 56% dos RIs declararam possuir um acervo superior à 2.000 registros, sendo um total divergente com o resultado aferido na pesquisa documental. Levando-se em consideração o número de RIs que declararam possuir um acervo maior que 1000 registros, o total é de 67%.

Gráfico 5 - Questionário - Tamanho da coleção



Fonte: O autor (2014)

A amostra revelou que 89% dos RIs não possuem uma política de preservação digital formalizada, conforme ilustra o gráfico abaixo.

Gráfico 6 - Questionário - Política formalizada



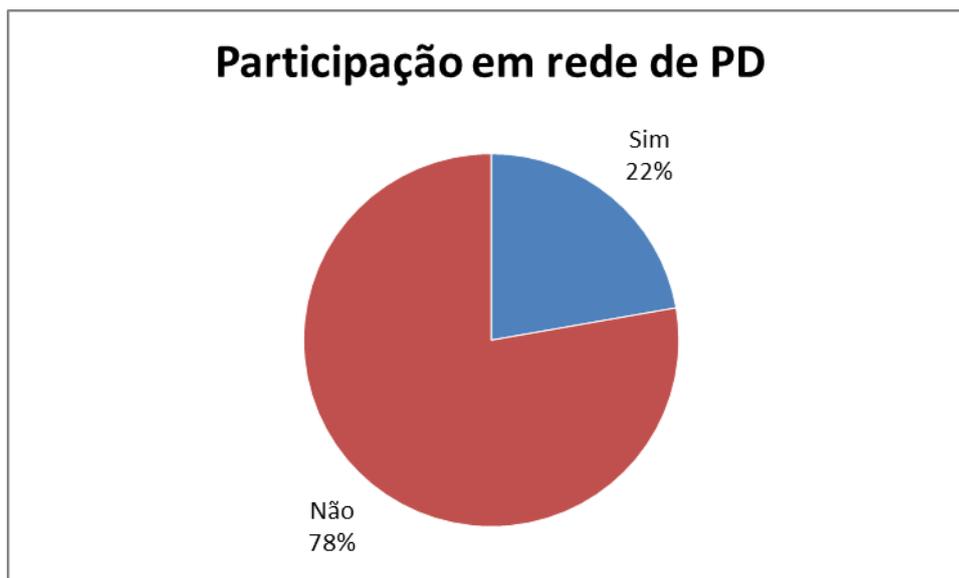
Fonte: O autor (2014)

A falta deste documento, formalizado – escrito de fato – pode evidenciar problemas em relação à padronização de critérios, métodos ou periodicidade definida para o desenvolvimento do processo de preservação digital. O conhecimento tácito não deve ser a única fonte de conhecimento para ações de política de preservação digital, conforme foi evidenciado pelos resultados obtidos. Somente 1 de 10 RI possui uma política de preservação digital institucionalizada.

A política de preservação digital, quando formalizada e pública, ajuda a sedimentar um posicionamento transparente do repositório em relação à preservação digital, explicando ao autor (depositantes) os procedimentos adotados e reforçando a garantia de perpetuação e acesso.

A participação em iniciativas em preservação digital em rede demonstrou um percentual negativo de 78%. Destes 22%, apenas a iniciativa da Rede Cariniana foi citada, 2 (duas) vezes, sendo uma delas mencionada como “IBICT”.

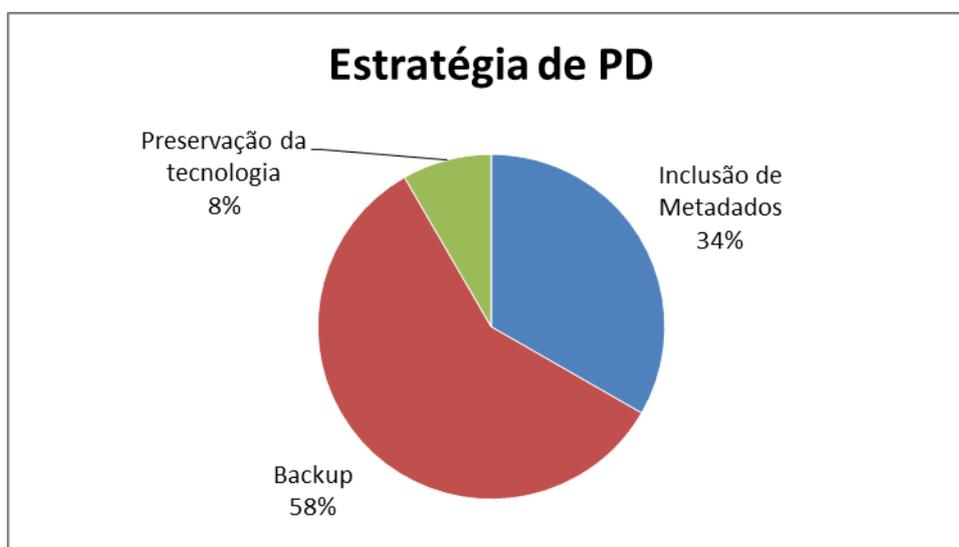
Gráfico 7 - Questionário - Redes de preservação digital



Fonte: O autor (2014)

Em relação as estratégias de preservação digital utilizadas, um número significativo – 58% – declararam utilizar o backup com o estratégia de preservação digital, seguido de 34% de menções à inclusão de metadados, sendo que apenas 02 unidades declararam realizar as duas estratégias ao mesmo tempo. Assim, é possível identificar que as estratégias de backup e inclusão de metadados estão profundamente difundidas nos RIs, representando, juntas, 92% das respostas colhidas.

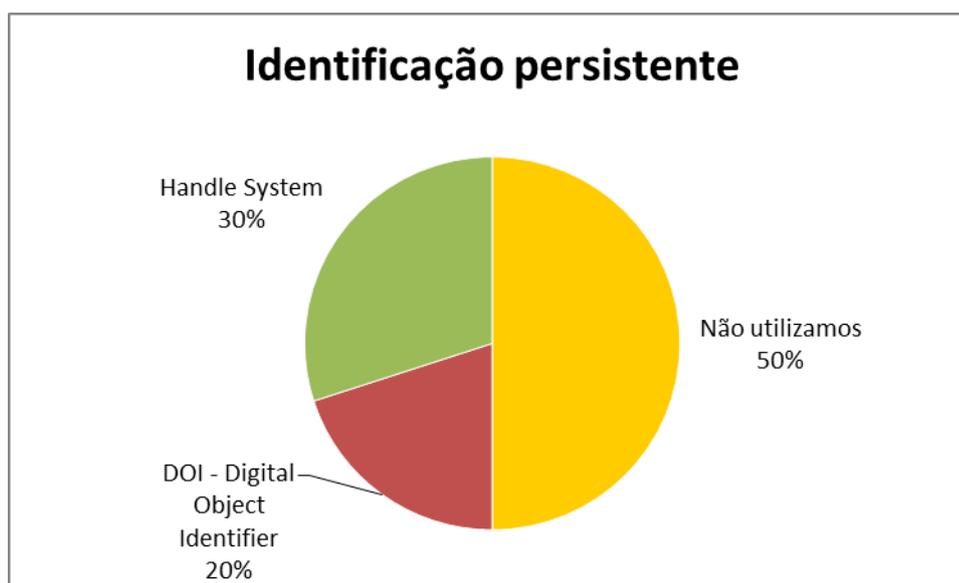
Gráfico 8 - Questionário - Estratégias de preservação



Fonte: O autor (2014)

A identificação persistente revelou um resultado divergente ao apresentado na pesquisa documental. Houve a declaração de uso de, preferencialmente, dois sistemas de identificação: o Handle System (30%) e o DOI (20%). Porém, também revelou um número demasiadamente grande – 50% – em relação aos RIs que declararam não utilizar nenhum tipo de identificação persistente, em contraponto ao revelado durante a pesquisa documental.

Gráfico 9 - Questionário - Identificação persistente



Fonte: O autor (2014)

Para os tipos de arquivos identificados, a apresentação dos resultados precisou de um arranjo diferenciado, devido a grande opção de respostas. Para isso, foram selecionados apenas as extensões que receberam alguma menção na resposta do questionário. A partir daí, os resultados foram separados em quadros onde aparecem subdivididos 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) – classificação baseada na tabela de adequação para a preservação de Munhoz (2011).

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. Se as mesma extensões fossem salvas sem a referência para o

DTD com esquema de validação de formato, estava classificado como média adequação de preservação.

Sendo assim, os resultados obtidos correspondem às tabelas a seguir:

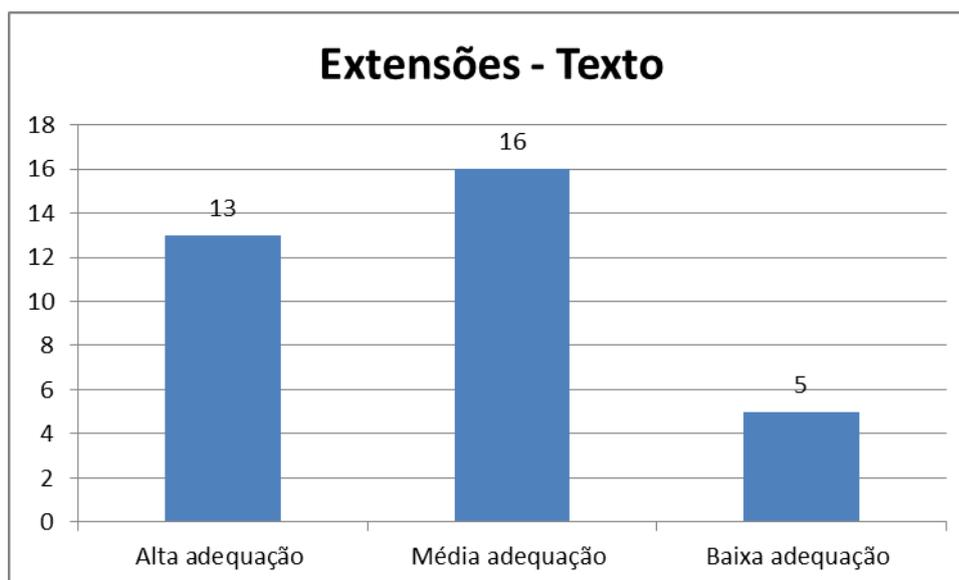
Quadro 15 – Resultados – Tipos de arquivos - Texto

	Extensão	Uso
Alta	PDF/A	6
	TXT	4
	HTML, XHTML ou XML	2
	ODT	1
Média	PDF	9
	DOCX	4
	RTF	2
	CSS	1
Baixa	DOC	4
	Todos os formatos	1

Fonte: O autor (2014)

É possível identificar uma forte tendência na utilização dos arquivos em PDF/A e PDF, mesmo sendo de classificação de adequação para preservação diferentes, alta e média, respectivamente. Também é possível perceber um aparente equilíbrio entre os formatos considerados de alta e de média adequação, de acordo com o gráfico abaixo. Apesar deste números, o gráfico abaixo demonstra que a distribuição dos usos é mais homogênea entre as extensões de alta adequação, levando em consideração o total dos grupos do quadro acima.

Gráfico 10 - Questionário - Extensões de texto



Fonte: O autor (2014)

A análise de extensões de imagens levou em consideração os 5 (cinco) RIs que não aceitam imagens (sendo os RIs 1, 5, 7, 8 e 9), não classificando seu resultado em nenhuma categoria. Também levou-se em consideração os RIs que declararam aceitar qualquer tipo de imagem, classificando esse resultado como baixa adequação para a preservação digital, por ser uma prática nociva ao próprio RI.

Quadro 16 – Resultados – Tipos de arquivos - Imagem

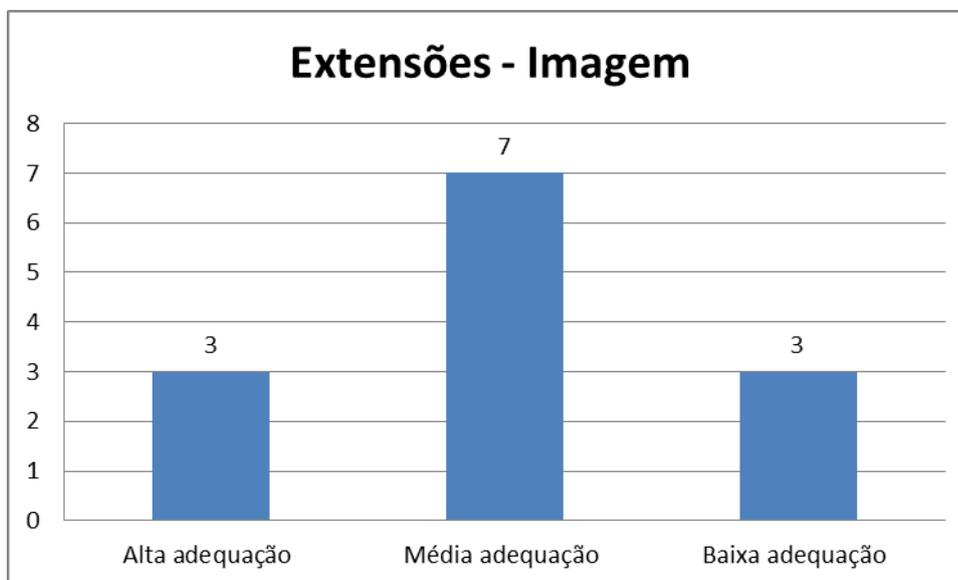
	Extensão	Uso
Alta	TIF	2
	PNG	1
Média	JPG	4
	GIF	2
	BMP	1
Baixa	PSD	1
	TIFF	1
	Todos os formatos	1

Fonte: O autor (2014)

Assim, é possível identificar uma disparidade entre o uso das extensões com alta e média adequação. Esta disparidade cresce, levando-se em consideração a soma dos resultados

de média e baixa adequação. Além disso, das 4 opções de extensão disponíveis para alta adequação, conforme Munhoz (2011), apenas 2 foram citadas.

Gráfico 11 - Questionário - Extensões de imagem



Fonte: O autor (2014)

A análise dos resultados para os tipos de extensões de vídeo seguiram o mesmo critério da análise de imagens, ou seja, foram considerados os RIs que declararam não aceitar este tipo de mídia (sendo os RIs 5, 7, 8 e 9) e este resultado não foi inserido em nenhuma categoria. Os resultados demonstram uma maior tendência no uso de arquivos de média adequação, apesar do, também, alto índice de alta adequação.

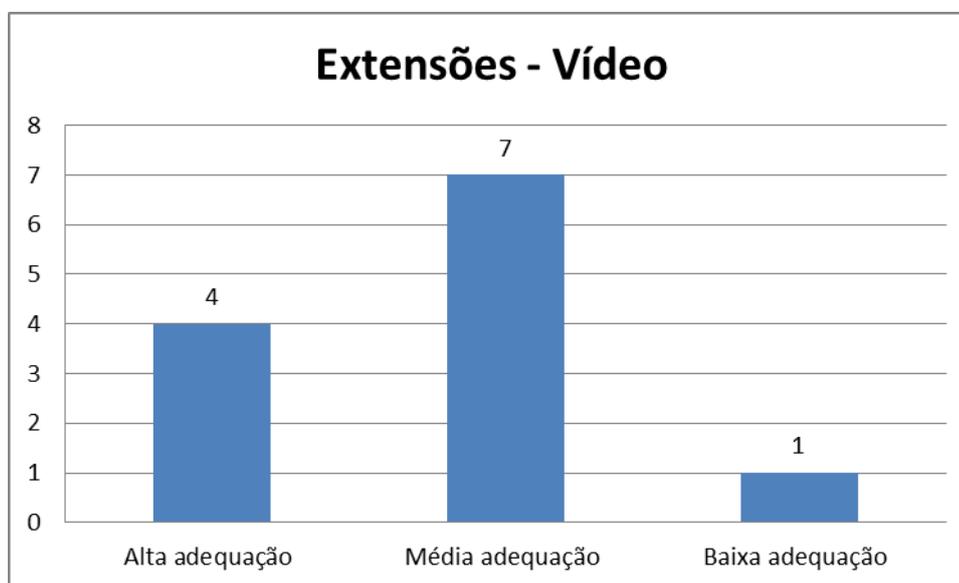
Quadro 17 – Resultados – Tipos de arquivo - Vídeo

	Extensão	Uso
Alta	MOV	2
	AVI	2
Média	MPEG	4
	MP4	3
Baixa	Todos os formatos	1

Fonte: O autor (2014)

O resultado também demonstra que apenas 4 dos 10 formatos disponíveis (sem levar em consideração as variações), conforme Munhoz (2011), foram citados nas respostas, o que sinaliza a baixa adesão deste formato nos RIs.

Gráfico 12 - Questionário - Extensões de vídeo



Fonte: O autor (2014)

Grande parte das extensões de vídeos declaradas em resposta ao questionário foi de média adequação.

Assim, os formatos citados durante o questionário correspondem ao quadro a seguir apresentado, culminando na amostra consolidada dos formatos utilizados nas instituições públicas de ensino e de pesquisa com atuação em território nacional, mantenedora de programas de pós-graduação reconhecidos pelo MEC/CAPES, conforme critério de edital de fomento a RIs.

Quadro 18 – Resultado consolidado

	Texto	Imagem	Vídeo
Alta	HTML, XHTML ou XML TXT PDF/A ODT	PNG TIF	MOV AVI

continua

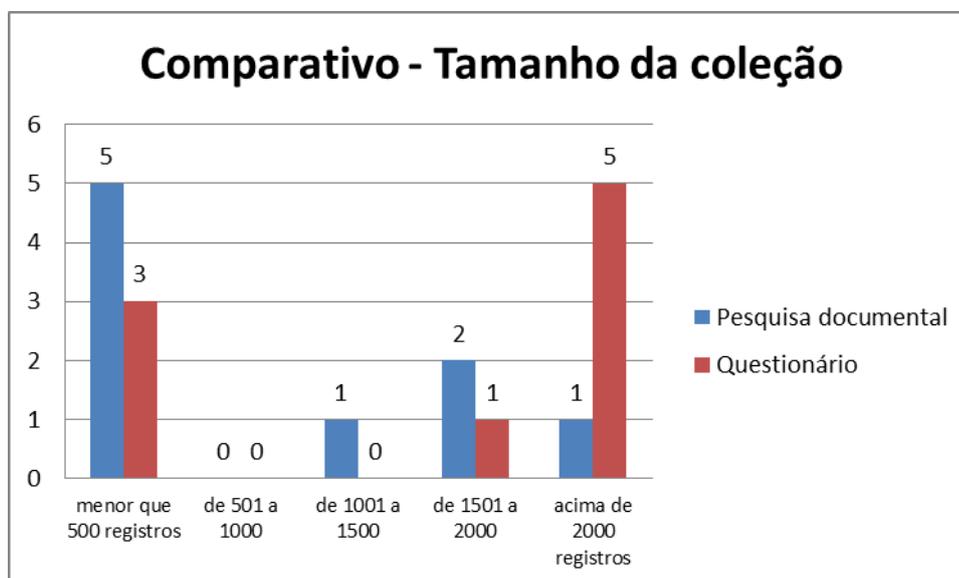
	Texto	Imagem	Vídeo (continuação)
Média	PDF CSS RTF DOCX	GIF JPG BMP	MPEG MP4
Baixa	DOC	PSD TIFF	---

Fonte: O autor (2014)

Dentre os resultados obtidos durante o estudo de caso, é importante ressaltar a discrepância em 2 (dois) em especial: tamanho da coleção e identificador persistente.

Através da pesquisa documental e dos resultados do questionário, é possível identificar uma polarização inversa entre os resultados do quesito tamanho da coleção. Enquanto na pesquisa documental é revelado o maior índice de repositórios com uma coleção menor que 500 registros, com os outros três resultados apontando uma média de 1,5, os resultados do questionário revelam o maior índice de repositórios com uma coleção acima de 2.000 registros, com os outros dois resultados apontando uma média 2, conforme gráfico abaixo.

Gráfico 13 - Resultados - Tamanho da coleção



Fonte: O autor (2014)

Através do gráfico comparativo é possível identificar uma acentuada divergência em relação ao que os gestores declaram como sendo o acervo total que povoa os RIs do total que está de fato disponível para consulta

A identificação persistente também revelou um resultado divergente em comparação entre a pesquisa documental e o questionário. Confrontando com os resultados obtidos na resposta ao questionário onde o Handle System obteve 30% e o DOI 20% das citações de resposta, enquanto os outros 50% declararam não utilizar nenhum tipo de identificação persistente, com os resultados da pesquisa documental, verificou-se uma grande diferença. Durante a pesquisa documental foi possível identificar 100% de uso do Handle System como método de identificação persistente aos objetos digitais, pelo uso do endereço “.../handle/” nas URLs dos objetos digitais depositados no RIs selecionados pela amostra.

#### 4 CONSIDERAÇÕES FINAIS

A pesquisa realizada buscou ressaltar aspectos importantes das estratégias de preservação digital utilizadas pelos RIs com o objetivo de mapear, da maneira mais fiel possível, as práticas de preservação digital utilizadas no dia-a-dia, evidenciando possíveis contribuições para a formalização de um subsídio para políticas de preservação em repositórios instrucionais.

Destacam-se a colaboração de todos os bibliotecários responsáveis pelos RIs na construção desta pesquisa e no esforço contínuo em, cada vez mais, possibilitar o acesso à informação de forma ágil, completa e irrestrita.

Os resultados obtidos através da pesquisa documental e o questionário demonstraram ser suficientes na identificação dos objetivos da pesquisa, ajudando a delimitar as estratégias de preservação digital das unidades selecionadas.

Foi possível também identificar um percentual de 89% de RIs pesquisados que não possuem uma política de preservação digital formalizada, e apenas 01 (um) RI pesquisado apresenta uma política de indexação institucionalizada. Este percentual evidencia um uso inexpressivo de um importante recurso como a política de preservação digital, estando aquém dos resultados desejados, demonstrando o distanciamento desta política com os demais objetivos da universidade/instituição a qual ela presta serviço, a qual, certamente, tem interesse na preservação do conhecimento por ela produzido. O baixo índice no uso de extensões de arquivos de alta adequação de preservação digital pode refletir a falta de políticas na escolha destes arquivos.

A política de preservação digital é um importante instrumento para a garantia do cumprimento desta estratégia, contudo, ficou evidente que este instrumento não está presente em grande parte dos RIs pesquisados. É possível que haja atividades de preservação digital nas unidades pesquisadas. Contudo, não é possível mensurar sob quais circunstâncias ela se dá ou se realmente estas ações são cumpridas. Um instrumento formalizado, discutido e baseado nas melhores práticas em preservação digital ajudaria a consolidar uma identidade positiva inerente à preservação digital: a garantia de acesso futuro aos documentos ali depositados.

Também apurou-se um alto o percentual de RIs que não participam de redes de preservação digital, com 78%. A política de preservação em redes oferece benefícios como à

garantia perene de acesso aos objetos digitais, uma vez que os mesmos encontram-se salvaguardados em mais de uma unidade participante da rede. Este isolamento verificado pode ser explicado pela falta de política de preservação digital, pois, pressupõem-se que este nível de preservação digital seja aplicadas a coleções que ofereçam o mínimo de esquematização, com uma política de preservação digital definida e em pleno funcionamento.

Também ficou evidente nos resultados da amostra a falta de diversificação nas estratégias de preservação digital. Apenas 03 (três) estratégias são citadas: 58% Backup; 34% Inclusão de metadados e 8% Preservação da tecnologia. Além disso, apenas 2 (dois) RIs pesquisados declararam utilizar mais de uma estratégia como medidas complementares.

Considerando o baixo número de políticas institucionais formalizadas, a baixa adesão às estratégias de preservação em rede e a falta de diversificação das estratégias de preservação digital, os resultados apresentados evidenciam a fragilidade do papel da política de preservação digital, corroborando com os problemas apontados por Santhanagopalan, Fox e Mcmillan (2006, p. 2, tradução nossa) para a preservação digital no Brasil, principalmente os itens 1 - “falta de atividades relacionadas às normas de preservação” e 3 - “práticas atuais são feitas por meio da técnica de preservação tradicional”.

É necessário um avanço no estudo desta temática para o estabelecimento de diretrizes sólidas para os critérios de criação de uma política de preservação digital. Este trabalho teve por objetivo identificar as boas práticas de preservação digital e baseando-se nos resultados obtidos e no conjunto de objetivos genéricos apresentados por Bullock (1999), define-se que esta política de preservação digital prescindiria de 3 (três) grupos básicos de esquematização: o objeto digital, os metadados e a política.

O objeto digital que faz parte de um RI precisa ter a garantia de sua preservação a longo prazo sem a perda de conteúdo ou modificação na sua apresentação. Para tanto, é necessário estabelecer critérios durante o processo de povoamento destes objetos, com a escolha de formatos que possibilitem um bom grau de adequação as práticas de preservação.

Os metadados devem ter uma participação ainda maior no processo de preservação digital, ampliando seu uso para além dos metadados descritivos. É necessário o uso de metadados estruturais e administrativos, complementando a função descritiva do objeto digital no contexto do RI.

A política de preservação formalizada deve ser o instrumento a sedimentar as práticas de preservação digital, sendo institucionalizadas ou não. Uma vez formalizada, ela explicitará

para a comunidade atendida pelo RI os procedimentos de preservação, e poderá ainda garantir a autenticidade e proveniência, através de termos de publicação e outros instrumentos que podem ser associados a ela.

Dessa forma a presente pesquisa identificou os seguintes critérios que devem ser considerados nas propostas de políticas de preservação digital:

- Escolha de formatos de alta e média adequação (visando minimizar a necessidade de migração);
- Uso de estratégias de preservação como, por exemplo, a inclusão de metadados descritivos e backup (se possível usar as estratégias de formas combinadas);
- Uso de metadados estruturais e administrativos;
- Uso de identificação persistente;
- Adesão a uma rede distribuída para preservação digital;
- Uso de uma política de preservação formalizada;

Essas práticas podem ser utilizadas como um subsídio de política de preservação digital.

Cabe ressaltar que mesmo no Edital FINEP/PCAL/XBDB 003/2009 (Anexo A), usado como parâmetro no critério de instituições elegíveis para a amostra desta pesquisa, o tema da preservação digital não é abordado ou mesmo uma garantia de manutenção mínima do conteúdo, mesmo sendo um edital de apoio a projetos de implantação de RIs. É certo que o mesmo edital condiciona compromissos como de conexão permanente na Internet do RI e no estabelecimento de uma política institucional de informação, que vise garantir a alimentação do RI por parte de seus pesquisadores das instituições. Contudo, estas medidas garantem apenas o acesso e o povoamento, respectivamente, mas não a preservação digital e por consequência a manutenção do conhecimento produzido para as gerações futuras.

A necessidade em produzir – por parte dos autores – e de organizar e promover acesso a essa produção – por partes dos RIs – deve encontrar o equilíbrio com a preservação digital, para garantir que o esforço dispensado hoje esteja acessível no futuro.

Os recursos eletrônicos representam, de fato, um avanço na comunicação científica, possibilitando acesso rápido e contínuo aos acervos dos RIs. Contudo, é necessário lembrar que os recursos tecnológicos não estão imunes a problemas, tanto no que se refere aos softwares quanto aos hardwares. Estes problemas podem comprometer o acesso parcial ou total ao acervo, limitando o uso. É necessária a criação estratégias que visem à garantia da

permanência de acesso aos objetos digitais, assim como foram desenvolvidos, e também aos registros dos bancos de dados dos RIs, para que, havendo qualquer falha o restabelecimento possa se dar da forma mais rápida possível.

As estratégias apresentadas neste trabalho não correspondem, necessariamente, a um elevado investimento financeiro para que a preservação digital seja realizada a contento, mas às estratégias dependem de políticas formalizadas e sedimentadas nas práticas das instituições que detenham um RI. O homem imputa ao computador a responsabilidade de preservação, como se a modificação do suporte para o digital já representasse este fim. É necessária a criação de espaços de discussão e troca de experiências para reforçar a importância das práticas de preservação digital.

## REFERÊNCIA

ABRANTES, Angelo Antônio; MARTINS, Lígia Márcia. A produção do conhecimento científico: relação sujeito-objeto e desenvolvimento do pensamento. **Interface: comunicação, saúde e educação**, Botucatu, SP, v. 11, n. 22, p. 313-325, maio/ago. 2007.

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.

BULLOCK, Alison. Preservation of digital information: issues and current status. **Network Notes**, Ottawa, Canadá, n. 60, apr. 1999. Disponível em: <<http://epe.lac-bac.gc.ca/100/202/301/netnotes/netnotes-h/notes60.htm>>. Acesso em: 01 maio. 2014.

CASTRO, Cristiane Yanase Hirabara de. et al. Repositórios institucionais confiáveis: repositório institucional como ferramenta para a preservação digital. In. SAYÃO, Luis. et al (Org.). **Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação**. Salvador: EDUFBA, 2009. p. 283-304. Disponível em: <<http://www.repositorio.ufba.br/ri/handle/ufba/473>>. Acesso em: 12 set. 2014.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS - CCSDS. **Reference model for an Open Archival Information System (OAIS)**: Magenta Book. Washington: National Aeronautics and Space Administration. jun. 2002. Disponível em: <<http://public.ccsds.org/publications/archive/650x0m2.pdf>>. Acesso em: 12 set. 2014.

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: 12 abr. 2014.

CUNHA, Murilo Bastos da. Desafios na construção de uma biblioteca digital. **Revista Ciência da Informação**, Brasília, v. 28, n. 3, p. 257-268, set./dez. 1999.

FRIGO, Denise. **Preservação digital: um subsídio para o centro de artes e letras da UFSM**. 2012. 155 f. Dissertação (Mestrado Profissionalizante em Patrimônio Cultural) - Centro de Ciências Sociais e Humanas, Universidade Federal de Santa Maria, Santa Maria, RS, 2012. Disponível em: <<http://coral.ufsm.br/ppgppc/images/dissertacaodenisefrigo.pdf>>. Acesso em: 10 set. 2014.

GOICOECHEA, Maria. Hyper-paradigm. In. SANZ, Amelia; ROMERO, Dolores (Edit.). **Literatures in the digital age: theory and praxis**. Newcastle, UK: Cambridge Scholars Publishing, 2007. p. 18-21.

GRANGER, Stewart. Digital preservation and deep infrastructure. **D-Lib Magazine**, v. 8, n. 2, fev. 2002. Disponível em: <<http://www.dlib.org/dlib/february02/granger/02granger.html>>. Acesso em: 10 fev. 2014.

HEDSTROM, Margaret. Digital preservation: problems and prospects. **Digital Libraries**, n. 20, mar. 2001. Disponível em: <[http://www.dl.slis.tsukuba.ac.jp/DLjournal/No\\_20/1-hedstrom/1-hedstrom.html](http://www.dl.slis.tsukuba.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html)>. Acesso em: 12 abr. 2014.

LEITE, Fernando et al. **Boas práticas para a construção de repositórios institucionais da produção científica**. Brasília: Ibict. 2012. p. 34. Disponível em: <<http://livroaberto.ibict.br/handle/1/703>>. Acesso em: 24 abr. 2014.

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

MÁRDERO ARELLANO, Miguel Ángel. Preservação de documentos digitais. **Revista Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27. maio/ago. 2004.

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://bdtd.bce.unb.br/tesesimplificado/tde\\_busca/arquivo.php?codArquivo=4547](http://bdtd.bce.unb.br/tesesimplificado/tde_busca/arquivo.php?codArquivo=4547)>. Acesso em: 25 fev. 2014.

MÁRDERO ARELLANO, Miguel Ángel. Cariniana: uma rede nacional de preservação digital. **Revista Ciência da Informação**, Brasília, v. 41, n. 1, p. 83-91. jan./abr. 2012. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/2127>>. Acesso em: 23 jul. 2014.

MÁRDERO ARELLANO, Miguel Ángel; LEITE, Fernando César Lima. Acesso aberto à informação científica e o problema da preservação digital. **Biblios: Revista de Bibliotecología y Ciencias de la Información**, Brasília. n. 35, p. 1-11. mar./jun. 2009. Disponível em: <<http://repositorio.unb.br/handle/10482/4937>>. Acesso em: 23 abr. 2014.

MORIGI, Valdir José; SOUTO, Luzane Ruscher. Entre o passado e o presente: as visões de biblioteca no mundo contemporâneo. **Revista ACB: Biblioteconomia em Santa Catarina**,

Florianópolis, v. 10, n. 2, p. 189-206. jan./dez. 2005. Disponível em:  
<<http://revistaacb.emnuvens.com.br/racb/article/view/432>>. Acesso em: 26 abr. 2014.

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://hdl.handle.net/10183/37473>>. Acesso em: 10 set. 2014.

OPEN SOURCE INITIATIVE. 2014. Disponível em: <<http://opensource.org/>>. Acesso em: 23 jul. 2014.

REICH, Victoria Ann. LOCKSS: ensuring access through time. **Revista Ciência da Informação**, Brasília, v. 41, n. 1, p. 74-82. jan./abr. 2012. Disponível em:  
<<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/2125>>. Acesso em: 23 jul. 2014.

REZENDE, Denis Alcides; ABREU, Aline França. **Tecnologia da informação aplicada a sistemas de informação empresariais**: o papel estratégico da informação e dos sistemas de informação nas empresas. 3. ed. São Paulo: Atlas, 2003.

SANTHANAGOPALAN, Kamini; FOX, Edward A; McMILLAN, Gail. A prototype for preservation and harvesting of international ETDs using LOCKSS and OAI-PMH. In. INTERNATIONAL SYMPOSIUM ON ELECTRONIC THESES AND DISSERTATIONS, 9., 2006. ,Québec, Canadá. **Anais eletrônicos...** Québec: NDLTD. 2006. Disponível em:  
<[http://scholar.lib.vt.edu/theses/presentations/ETD\\_2006\\_Kamini\\_Santhanagopalan.doc](http://scholar.lib.vt.edu/theses/presentations/ETD_2006_Kamini_Santhanagopalan.doc)>. Acesso em: 12 abr. 2014.

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](http://repositorio.unb.br/bitstream/10482/5343/6/2010_ErnaniRufinodosSantosJunior.pdf)>. Acesso em: 12 abr. 2014.

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; Brasília: UFBA; IBICT. 2005. p. 115-146.

SAYÃO, Luis Fernando. Conservação de documentos eletrônicos. In. GRANATO, Marcus; SANTOS, Claudia Penha dos; ROCHA, Cláudia Regina Alves da (Org.). **Conservação de acervos**. Rio de Janeiro: MAST. 2007a. p. 181-204. (MAST Colloquia; 9). Disponível em:  
<[http://www.mast.br/livros/mast\\_colloquia\\_9.pdf](http://www.mast.br/livros/mast_colloquia_9.pdf)>. Acesso em: 10 fev. 2014.

SAYÃO, Luis Fernando. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes - URN, PURL, DOI, Handle System, CrossRef e OpenURL. **TransInformação**, Campinas: v. 19, n. 1, p. 65-82. jan./abr. 2007b. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/621>>. Acesso em: 10 set. 2014.

SAYÃO, Luis Fernando. Uma outra face dos metadados: informações para a gestão da preservação digital. **Revista Encontros Bibli**, Florianópolis, v. 15, n. 30, p. 1-31. 2010.

SAYÃO, Luis Fernando; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Revista Informação & Sociedade**, João Pessoa, v. 22, n. 3, p. 179-191. set./dez. 2012.

SILVA, Patrícia Maria. Sistemas de informação em bibliotecas: o comportamento dos usuários e bibliotecários frente às novas tecnologias de informação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 5, n. 2, p. 1-24. jan./jun. 2008.

TARGINO, Maria das Graças. Comunicação científica: uma revisão de seus elementos básicos. **Revista Informação & Sociedade**, João Pessoa, v. 10, n. 2, p. 1-27. 2000. Disponível em: <<http://www.ies.ufpb.br/ojs/index.php/ies/article/view/326>>. Acesso em: 26 abr. 2014.

VIDAL, Alexandra. A conservação e a preservação de documentos digitais: um desafio na era da sociedade de informação. **Revista da Faculdade de Ciências Humanas e Sociais**. Porto, Portugal, v. 7, p. 144-154. 2010. Disponível em: <<http://hdl.handle.net/10284/2809>>. Acesso em: 09 abr. 2014.

## **ANEXO A - EDITAL FINEP/PCAL/XBDB 003/2009**



### **EDITAL DE CHAMADA FINEP/PCAL/XBDB Nº 003/2009**

A FUNDAÇÃO DE CIÊNCIA, APLICAÇÕES E TECNOLOGIA ESPACIAIS – FUNCATE inscrita no CNPJ sob o número 51.619.104/0001-10, entidade de direito privado, com sede em São José dos Campos, Estado de São Paulo, Av. Dr. João Guilhermino, 429, 11º andar, Centro CEP 12.210-131, na forma do Convênio 23.02.0267.00 firmado com a FINEP – Financiadora de Estudos e Projetos, tendo o IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia como executor, torna público presente Edital e convoca interessados a apresentarem propostas para concessão de recursos tecnológicos às instituições públicas (federais, estaduais e municipais) de ensino e pesquisa, no País, em conformidade com a Lei nº 8.666, de 21 de junho de 1993, e suas alterações, que regulam a realização de eventos científicos.

#### **1 – Informações Gerais**

##### **1.1 – Objetivo**

Apoiar projetos de implantação de repositórios institucionais (RI) nas instituições públicas (federais, estaduais e municipais) de ensino e pesquisa e sua integração ao Portal Oásis.Br, com vistas a possibilitar o registro e a disseminação da produção científica dessas instituições e proporcionar maior visibilidade à sua produção científica.

##### **1.2 - Linha de apoio**

O apoio é específico às instituições públicas (federais, estaduais e municipais) de ensino e pesquisa que possuem programas de pós-graduação reconhecidos pelo

Ministério da Educação/Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (MEC/CAPES), para a implantação de seus RI. Para tanto, serão distribuídos *kits* tecnológicos à cada proposta vencedora, em regime de comodato, composto de :

- 1 (um) servidor, pré-formatado e configurado com:
  - sistema operacional baseado na plataforma UNIX/LINUX;
  - software básicos: Apache, MySQL e PHP;
  - software: Dspace e SEER.

### 1.3 – Cronograma

Lançamento da chamada no DOU e na página do IBICT	26 de abril de 2010
Data limite para submissão das propostas	23 de julho de 2010
Análise, julgamento e seleção	26 a 27 de julho de 2010
Divulgação dos resultados	28 a 30 de julho de 2010

### 1.4 – Instituições Elegíveis

Instituição Pública de Ensino e Pesquisa com atuação em território nacional, mantenedora de programas de pós-graduação reconhecidos pelo MEC/CAPES e que atendam os seguintes quesitos obrigatórios.

1.4.1 - Os quesitos obrigatórios indicados a seguir são válidos para o presente Edital. O atendimento às mesmas é considerado imprescindível para o exame da proposta. A ausência ou insuficiência de informações sobre quaisquer delas resultará em não-enquadramento da proposta.

**a)** – manifestar, formalmente, o compromisso institucional de promover e viabilizar as ações necessárias para a implantação do RI compatível com o modelo de interoperabilidade *Open Archives*;

**b)** - manifestar, formalmente, o compromisso institucional de estabelecer uma política institucional de informação (PII) que vise garantir a alimentação do RI por parte de seus pesquisadores, em um prazo não superior a três meses. Caso, a instituição já possua a sua política que garante a alimentação dos RI por parte dos seus pesquisadores, basta apresentá-la;

**c)** – manifestar, formalmente, o compromisso institucional de constituir e manter uma equipe composta por técnicos de informação e informática que se responsabilizará pela implantação e operação do RI;

**d)** – manifestar, formalmente, o compromisso institucional de manter conexão permanente com a Internet para garantir o acesso ao RI, a fim de possibilitar a coleta automática periódica de metadados a ser realizada pelo IBICT, com vistas a alimentar o Portal Oásis.Br.

1.4.2 – Além de satisfazer os requisitos solicitados, a instituição deverá encaminhar ao IBICT as seguintes informações adicionais:

**a)** Relação dos cursos de pós-graduação da instituição;

**b)** Série histórica de produção científica da instituição nos últimos 5 (cinco) anos (basta indicar a quantidade publicada por ano);

**c)** Número de pesquisadores existentes na instituição, distribuídos por titulação, indicando a sua classificação junto ao CNPq (exemplo: 1A, 1B, 1C etc);

**d)** Relação das revistas científicas com revisão por pares com a indicação das suas respectivas classificações no Qualis da Capes.

**e)** Metodologia de desenvolvimento e implantação do RI;

**f)** Cronograma das etapas de desenvolvimento e implantação do RI;

**g)** Características de conexão com a Internet (necessariamente dedicada, com IP fixo):

– rede institucional padrão T (cabramento em par trançado), preferencialmente 100BaseT:

– servidor deve estar, preferencialmente, em um domínio totalmente qualificado.

## **2 – Apresentação e Envio das Propostas**

2.1 - As propostas contendo toda a documentação e informações estabelecidas em 1.4.1 e 1.4.2, deverão ser encaminhadas via correio, juntamente com formulário próprio, disponível no endereço eletrônico do IBICT <http://www.ibict.br/>, totalmente preenchido e assinado pelo dirigente máximo da instituição de ensino

e pesquisa, devendo ser postado nos correios, no endereço abaixo, até 19 de julho de 2010 .

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT  
SAS – QUADRA 5 – LOTE 6 – BLOCO H – 2º. ANDAR  
70070-912 BRASÍLIA – D.F.

2.2 - Qualquer proposta enviada fora do prazo e dos padrões definidos pelo presente Edital será desconsiderada.

2.3 – A ausência do envio da documentação prevista em 2.1, ou o envio incompleto, acarretará a desabilitação da proposta.

### **3 – Da Análise, Julgamento e Seleção.**

3.1 – As propostas enviadas ao endereço mencionado no item 2.1 dentro do prazo estipulado, serão analisadas e avaliadas comparativamente.

3.1.1 – A análise preliminar será realizada por uma equipe técnica, designada pelo IBICT, quanto ao enquadramento das propostas às condições e exigências do presente Edital.

3.1.2 Após a análise preliminar, as propostas serão encaminhadas para uma **Comissão Julgadora**, instituído pelo IBICT, que fará criteriosa análise do mérito, procedendo assim ao julgamento e à seleção das propostas contempladas em conformidade com as determinações deste Edital.

3.1.2.1 – A Comissão Julgadora será constituída pelo IBICT e composta por 1 (um) representante do IBICT, 1 (um) da CAPES e 1 (um) representante da FINEP, 1 (um) representante da Comissão Brasileira das Bibliotecas Universitárias e 1 (um) representante da Associação Brasileira de Editores Científicos.

3.2 – A análise, julgamento e seleção das propostas são soberanas, não cabendo, pois, qualquer possibilidade de recursos.

### **4 – Critérios de Julgamento**

4.1 – Serão selecionados os projetos que apresentarem toda a documentação solicitada nos itens 1.4 e 1.4.1, assim como demonstrarem atendimento aos requisitos técnicos constantes no item 1.4.2.

4.2 – A Comissão Julgadora fará a seleção de instituições que demonstrarem efetivamente, em seus projetos, terem condições técnicas para a construção e manutenção do RI.

4.3 - Serão priorizadas as instituições que se dispuserem a facilitar o uso de seus recursos a outras instituições do mesmo Estado, ou de Estados de sua Região, que não se qualificarem a concorrer pelos termos do presente edital;

4.4 - Serão excluídas do processo as instituições que já possuam RI.

## **5 - Recursos Financeiros**

Os recursos para aquisição dos kits tecnológicos são oriundos do Convênio nº 23.02.0267.00 firmado com a Financiadora de Estudos e Projetos – FINEP.

## **6 - Prazo e Execução do Projeto**

6.1 – O processo de distribuição dos kits tecnológicos obedecerá o prazo de três meses após a divulgação dos resultados do presente processo licitatório.

6.1.1 – Este prazo de três meses será concedido para que as instituições possam discutir e estabelecer as suas PII de forma a garantir que os seus respectivos pesquisadores depositem uma cópia de seus trabalhos publicados em publicações com revisão pelos pares.

6.1.2 - As instituições cujas propostas forem vencedoras no presente processo licitatório somente receberão o seu kit tecnológico mediante a comprovação de estabelecimento de sua respectiva PII.

6.2 – Uma vez entregues os kits tecnológicos, as instituições terão três meses, contados à partir da data de entrega do referido kit, de prazo para o efetivo desenvolvimento e implantação do seu respectivo RI.

6.2.1 – Findo este prazo de três meses, o RI deverá estar apto para ser coletado pelo Portal Oásis.Br.

6.2.2 - O não-cumprimento do prazo definido no Item 6.2 deverá ser devidamente justificado pela instituição.

6.2.2.1 – Caso a Comissão Julgadora deste processo de seleção não acate a justificativa, a concessão do kit tecnológico será cancelada e a instituição será considerada inadimplente, devendo devolver o referido kit tecnológico ao Instituto, o qual será objeto de um novo processo seletivo para sua redistribuição a outras instituições.

## **7 - Contrapartida**

Será considerada como contrapartida da instituição receptora:

7.1 - o compromisso de retirar o equipamento no IBICT com todas as despesas pagas pela instituição contemplada;

7.2 – o desenvolvimento, atualização, manutenção e administração do seu RI, com recursos próprios e a sua manutenção em bom estado de funcionamento.

7.2.1 A não atualização do RI por um período de tempo igual ou maior do que 3 (três) meses acarretará no cancelamento da concessão do kit tecnológico e, obviamente, na sua devolução ao Ibict, tornando a respectiva instituição inadimplente.

7.2.2 O uso do kit tecnológico deve ser dedicado à hospedagem do RI, podendo ser utilizado também para hospedar revistas científicas compatíveis com o modelo de interoperabilidade *Open Archives*, e/ou repositório de teses e dissertações, sendo vedado o seu uso para qualquer outro propósito.

## 8 – Cláusula de Reserva

O Comitê indicado reserva-se o direito de resolver os casos omissos e as situações não previstas no presente Edital.

## 9 – Considerações finais

- i. O resultado final será divulgado na página do IBICT e publicado no Diário Oficial da União (DOU).
- ii. A metodologia de desenvolvimento e implantação do RI deve ter como objetivo final a completa implantação do repositório institucional. Isto quer dizer que não há necessidade de o repositório ser implantado de uma única vez contendo a produção científica de toda a instituição. A metodologia poderá contemplar uma implantação gradual, de departamento por departamento, ou de faculdade por faculdade, ou de escola por escola. Essa implantação gradual deverá esta espelhada no cronograma para o seu acompanhamento.
- iii. A política institucional de informação deverá ser aplicável a toda a instituição e não apenas a uma única unidade ou a parte de suas unidades, razão pela qual se deve comprovar que essa política tenha sido estabelecida e aprovada pela instituição. Entretanto, aplicação da política institucional de informação poderá ser realizada de forma gradual.

No contexto deste edital deve-se entender por **instituição** o todo compreendido pela universidade ou pelo instituto ou centro de pesquisa. Para efeito deste edital não se considera departamento, faculdade ou escola como instituição, uma vez que essas unidades constituem parte do todo, ou seja da instituição.

- iv. Os compromissos formais, a que se refere o item 1.4.1 e seus subitens deverão ser assinados pelo dirigente máximo da instituição de ensino e pesquisa.

Esclarecimentos acerca do conteúdo deste Edital de Chamada poderão ser obtidos através dos telefones (61) 32176249 e e-mails: [bianca@ibict.br](mailto:bianca@ibict.br).

São José dos Campos,

Fundação de Ciência, Aplicações e Tecnologia Espaciais

Lauro Eduardo de Souza Pinto

Hélio Kuramoto

Coordenador do Projeto PCAL

Bianca Amaro de Melo

Coordenadora do Projeto XBDB

## APÊNDICE A – ROTEIRO PARA PESQUISA DOCUMENTAL

1 Tamanho da coleção

menor que 500 registros

de 501 a 1000

de 1001 a 1500

de 1501 a 2000

acima de 2000 registros

2 Indicação Creative Commons

Sim  Não

3 Política de Preservação Digital acessível?

Sim  Não

4 Relação dos formatos de arquivos aceitos?

Sim  Não

5 Identificação a participação em redes de preservação distribuída

Sim  Não

6 Apresenta identificação persistente no link de endereço aos objetos digitais?

Sim  Não

## APÊNDICE B – QUESTIONÁRIO

### Pesquisa - Preservação Digital em Repositório Institucional

Pesquisa realizada pelo discente Marco Aurélio Alencar, para seu Trabalho de Conclusão de Curso apresentado como requisito para a obtenção do título de Bacharel em Biblioteconomia, pela Escola de Biblioteconomia da Universidade Federal do Estado do Rio de Janeiro - UNIRIO

\* Obrigatório

1 - Instituição

\_\_\_\_\_.

2 - Estado \*

Sigla da UF: \_\_\_\_\_.

3 - Tamanho da coleção \*

Total de registros no Repositório. Pode ser um valor aproximado.

( ) menor que 500 registros

( ) de 501 a 1000

( ) de 1001 a 1500

( ) de 1501 a 2000

( ) acima de 2000 registros

Sobre a Política de Preservação Digital

4 - A política de preservação digital está formalizada? \*

A política está documentada, escrita em algum local para consulta?

( ) Sim \* se possível encaminhar via email

( ) Não

Caso negativo, você será redirecionado à pergunta nº 6

5 - A política de preservação digital é institucionalizada? \*

As ações, procedimentos ou políticas de preservação digital fazem parte da norma institucional ou equivalente (resolução ou portaria, por exemplo) da instituição, da unidade de informação ou da rede responsável pelo Repositório Institucional?

Há uma política institucional (para toda Universidade/Instituição)

Há uma política local (da unidade de informação apenas)

Há uma política em rede de preservação digital interinstitucional

Outro: \_\_\_\_\_

Sobre a Política de Preservação Digital em Rede

6 - A instituição ou RI faz parte de alguma rede que tenha finalidade da preservação digital? \*

Exemplo: Rede Cariniana ou LOCKSS

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: \_\_\_\_\_

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: \_\_\_\_\_

9 - Utiliza algum método de identificação persistente para os objetos digitais? \*

Se necessário, marque mais de um

- Não utilizamos
- URN - Uniform Resource Name
- RUL - Uniform Resource Locator
- PURL - Persistent URL
- DOI - Digital Object Identifier
- Handle System
- Outro: \_\_\_\_\_

10 - Quais são os tipos de arquivos (extensão) autorizados para inclusão no RI para documentos de TEXTO \*

As extensões de arquivo são os sufixos após o nome do objeto digital, como um arquivo em Word é exemplo.doc e em Excel é exemplo.xls

- |                                |   |
|--------------------------------|---|
| <input type="checkbox"/> PDF   | <input type="checkbox"/> HTML, XHTML ou XML |
| <input type="checkbox"/> PDF/A | <input type="checkbox"/> CSS                |
| <input type="checkbox"/> DOC   | <input type="checkbox"/> SXW                |
| <input type="checkbox"/> DOCX  | <input type="checkbox"/> DTD                |
| <input type="checkbox"/> TXT   | <input type="checkbox"/> PS                 |
| <input type="checkbox"/> ODT   | <input type="checkbox"/> Outro: _____       |
| <input type="checkbox"/> RTF   | <input type="checkbox"/> Qualquer formato   |

11 - Quais são os tipos de arquivos (extensão) autorizados para inclusão no RI para documentos de IMAGEM \*

- |                                      |   |
|--------------------------------------|---|
| <input type="checkbox"/> Sem imagens | <input type="checkbox"/> BMP              |
| <input type="checkbox"/> JPG         | <input type="checkbox"/> DNG              |
| <input type="checkbox"/> JPX         | <input type="checkbox"/> PSD              |
| <input type="checkbox"/> JP2         | <input type="checkbox"/> RAW              |
| <input type="checkbox"/> JPF         | <input type="checkbox"/> SVG              |
| <input type="checkbox"/> GIF         | <input type="checkbox"/> SWF              |
| <input type="checkbox"/> TIFF        | <input type="checkbox"/> Outro: _____     |
| <input type="checkbox"/> TIF         | <input type="checkbox"/> Qualquer formato |
| <input type="checkbox"/> PNG         |   |

12 - Quais são os tipos de arquivos (extensão) autorizados para inclusão no RI para documentos de VÍDEO (imagem em movimento) \*

- |                                     |   |
|-------------------------------------|---|
| <input type="checkbox"/> Sem vídeos | <input type="checkbox"/> MJ2              |
| <input type="checkbox"/> MOV        | <input type="checkbox"/> OGG              |
| <input type="checkbox"/> MPG        | <input type="checkbox"/> OGM              |
| <input type="checkbox"/> MPEG       | <input type="checkbox"/> OGV              |
| <input type="checkbox"/> AVI        | <input type="checkbox"/> Outro: _____     |
| <input type="checkbox"/> WMV        | <input type="checkbox"/> Qualquer formato |
| <input type="checkbox"/> MP4        |   |