

PRESERVAÇÃO DIGITAL: tendências atuais dos conceitos e técnicas

DIGITAL PRESERVATION: current trends of the concepts and techniques

Jacqueline de Araújo Cunha

jacquelinecunh@gmail.com

Em 20007, mestranda do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba (UFPB)

Marcos Galindo Lima

<http://lattes.cnpq.br/7413464711814360>

<https://orcid.org/0000-0001-5611-9586>

Doutor e professor da Universidade Federal de Pernambuco (UFPE)

Submetido: 2007

Publicado: 28 dez. 2023

RESUMO

Discute as atuais visões acerca da preservação digital, à luz da literatura publicada acerca do tema. Objetivou-se delinear um estado da arte sobre a preservação digital, seus conceitos e técnicas, bem como apresentar alguns dos organismos hoje empenhados nesta área, tanto nacionais como internacionais. Utilizou-se como metodologia pesquisas de cunho bibliográfico, onde foram utilizados textos nacionais e internacionais, de livros e periódicos em meio digital. Apresenta ainda um quadro elaborado a partir da pesquisa, onde serão abordadas as principais técnicas de preservação, suas vantagens e desvantagens. Concluiu-se que muitas são as ações neste sentido, mas que ainda não existem consensos no que diz respeito às técnicas e metodologias utilizadas, e portanto fundamenta-se abrindo um novo e amplo campo de estudos onde se produzam contribuições a fim de garantir a acessibilidade e preservação dos objetos digitais para a posteridade.

PALAVRAS-CHAVE: Preservação digital. Técnicas de preservação. Acessibilidade.

ABSTRACT

It argues the current views concerning the digital preservation, to the light of the literature published concerning the subject. It goals to delineate a state of the art on the digital preservation, its concepts and techniques, as well as presenting some of the organisms today stressing in this area, nationally as well internationally. It was used as methodology a bibliographic research, where national and international texts had been used, of books and periodics in digital media. It still presents a picture elaborated from the research, where the main techniques of preservation, its advantages and disadvantages will be approached. It concluded that many are the actions in this direction, but that not yet consensuses exist about the used techniques and methodologies, and therefore it is based opening a new and ample field of studies where they will produce contributions in order to guarantee the accessibility and preservation of digital objects for the posterity.

KEY WORDS: Digital preservation. Techniques of preservation. Accessibility.

1 INTRODUÇÃO

Desde que surgiram os primeiros registros informacionais, o homem sentiu a necessidade de preservar o conhecimento para a posteridade. Araújo (2002, p.32) quando afirma que “[...] a informação é um recurso para a ação política do sujeito social que transforma estruturas mentais e sociais, pois possibilita aos sujeitos sociais a criação de novos estados de conhecimento [...]. Este grande potencial da informação para transformação é um aspecto de fundamental importância para a sociedade, pois possibilita a renovação ou a reinvenção das estruturas e consequentemente da realidade”.

Portanto, preservar os suportes informacionais, seu acesso e uso no meio digital, levanta-se como condição de extrema importância para preservar-se o insumo fundamental do processo de construção do conhecimento, a informação. Percebe-se ainda, nesta mesma perspectiva, que a preservação permitiu às sociedades ao redor do mundo o acúmulo de conhecimento e a utilização do legado informacional daqueles que, durante anos e séculos, construíram paulatinamente seus saberes. Há, portanto, a necessidade de garantir que as informações que são produzidas hoje, estejam acessíveis na posteridade, pois configuram-se um rico patrimônio humano, fruto de sua produção cultural, social e ou científica.

Porém, o acelerado desenvolvimento dos suportes midiáticos, em especial os digitais, trazem novos desafios no que concerne a sua preservação a longo prazo, dado o desconhecimento do tempo de vida útil que estas mídias podem alcançar em perfeito estado de uso. Além disso, a obsolescência dos software configura-se como um desafio a mais neste processo.

Os estudos que tratam da problemática da preservação digital ganham relevância por vivermos numa sociedade cuja economia encontra-se pautada no conhecimento e na informação. Somado a isso, percebemos a utilização cada vez maior e em massa das tecnologias digitais na produção e veiculação dos conteúdos informacionais.

Deste modo, pretendeu-se delinear um quadro geral sobre conceitos, o histórico e a análise das técnicas, bem como os projetos de preservação digital nacionais e internacionais. A partir da discussão de alguns conceitos, pode-se verificar diferentes perspectivas nesta área. Algumas considerações acerca das técnicas de preservação, até hoje desenvolvidas, mostram os avanços obtidos, porém denunciam a necessidade de elaboração de políticas de preservação que vislumbrem a preservação como parte do processo de gerenciamento de coleções digitais. São ainda elencadas algumas das principais iniciativas de Preservação Digital em âmbito internacional e nacional.

Os limites deste artigo não nos permitem realizar uma análise profunda deste fenômeno, nem esgotar todas as fontes, mas oferecer alguns elementos que possibilitem conhecer um pouco das estratégias e ações que vem sendo empreendidas no âmbito da preservação de conteúdos digitais. Para tanto, foi utilizada a pesquisa bibliográfica, como método, com a leitura e análise de literatura científica já publicada acerca da preservação digital, encontrado em fontes informacionais como: livros, periódicos e na rede mundial de computadores.

A seguir apresentaremos um apanhado conceitual acerca da preservação digital onde serão elencados os principais conceitos encontrados nos textos analisados.

2 PRESERVAÇÃO DIGITAL: uma abordagem conceitual

Se partirmos do conceito de preservação no contexto analógico, encontraremos em Ferreira (2001), o conceito de que a preservação é “o ato de preservar” e este, por sua vez, significa “livrar de algum mal ou dano”. Assim sendo, podemos inferir que a preservação da informação no contexto analógico significa ‘garantir a integridade física do suporte’. Porém, quando se trata de suportes digitais a integridade física não parece suficiente, visto que se fazem necessários dispositivos que tornem acessíveis os conteúdos para o acesso humano (os discos rígidos, cd’s, disquetes, *etc*). Isto leva a uma necessidade de preservação também dos software, bem como dos equipamentos necessários à utilização dos mesmos. Este aspecto das mídias digitais traz novos desafios a preservação da informação.

Muitas definições de preservação digital são utilizadas hoje. As mais utilizadas são as de organizações com grandes projetos nesta área, tais como a *Online Computer Library Center* (OCLC), *Association for Information and Image Management* (AIIM), a *United Kingdom Office for Library Networking* (UKLON), dentre outras. Para visualizar melhor esta diferenciação alguns dos conceitos encontrados na literatura serão listados a seguir.

Para a *Association for Information and Image Management* (AIIM) a preservação digital consiste na “[...] habilidade de manter documentos digitais e arquivos acessíveis por períodos de tempo que transcendam avanços tecnológicos sem afetar por alteração ou perda da legibilidade” (CHAPMAN, 2001, tradução nossa).

No conceito proposto, da AIIM, a preservação digital consiste numa “habilidade” que tem por pretensão dar cabo de problemas de ordem tecnológica no que concerne à legibilidade dos conteúdos digitais ao longo do tempo. Ferreira (2006) considera que “[...] a preservação digital consiste na capacidade de garantir que a informação digital permaneça acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação” (FERREIRA, 2006, p. 20).

Nesta segunda definição, uma outra variável se acrescenta, a autenticidade. Pois, em se tratando de documentos digitais, nota-se que existe uma certa vulnerabilidade no que diz respeito a alterações, diferentemente das mídias analógicas onde, uma vez registrado e impresso, a modificação do documento torna-se inviável a menos que uma nova impressão seja feita. Tomemos por exemplo um arquivo registrado numa mídia digital (*pen drive*, cd, disquete), para modificá-lo basta abrir o arquivo com o software adequado e fazer a alteração, em seguida “salvar” a alteração, ao passo que num livro ou periódico, por exemplo, seria necessária uma nova impressão, ou seja, um outro documento, pois o impresso original não é passível de modificações. Portanto, quando são documentos de valor institucional, histórico ou legislativo, ao serem preservados, torna-se fundamental a sua autenticidade, posto que sem esta característica, a preservação será inútil.

Outra definição para a preservação digital bastante citada nos documentos que abordam o tema é a da *Research Library Group/Online Computer Library Center Report* (RLG/OCLC), a qual afirma que a “preservação digital refere-se a uma série de atividades gerenciadas necessárias para assegurar o acesso contínuo e preservação de materiais digitais” (CHAPMAN, 2001, tradução nossa).

Esta terceira definição acrescenta uma outra característica, que é o processo através do qual a preservação digital torna-se possível, ou seja, não se trata apenas de uma ‘habilidade’ ou ‘capacidade’ ou melhor, ‘técnica’ de preservar, mas sim um “gerenciamento de atividades”, cujo objetivo é a preservação digital.

Vale mencionar ainda a definição proposta por Hedstrom *apud* Boeres e Arellano, o qual conceitua a preservação digital como sendo “[...] um processo distribuído que envolve o planejamento, alocação de recursos e aplicação de métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável” (HEDSTROM *apud* BOERES; ARELLANO 2005, p. 2)

Esta última definição parece ser a mais completa, uma vez que privilegia a questão do “valor” da informação digital a ser preservada. Talvez por perceber os elevados custos que esta atividade demanda e por isso vislumbra a preservação como um processo que demanda planejamento e alocação de recursos.

Sobre as definições de preservação digital, acima relacionadas, percebe-se algumas nuances que as diferenciam, muito embora, de um modo geral, todas mencionem a preservação da informação digital a partir da preservação do suporte e do acesso ao seu conteúdo de forma legível. Tais definições ajudam a configurar um entendimento de qual é o sentido em que estão sendo direcionados, em alguns contextos, os esforços de organismos internacionais quanto à preservação de documentos digitais no mundo. Percebe-se que a questão da preservação da informação, antes do aparecimento do digital, estava relacionada, quase que exclusivamente, a durabilidade dos suportes.

Hoje, se vê também relacionada à rápida obsolescência tecnológica dos documentos legíveis por computador (quer seja em relação ao software ou hardware). A digitalização oportunizou uma preservação e conservação de suportes analógicos, bem como tornou possível o acesso aos conteúdos informacionais através da *web*. Porém, novamente questões de durabilidade são postas. A velocidade com que as tecnologias se desenvolvem e são rapidamente suplantadas, traz consigo o risco de perderem-se documentos em meio digital por falta de equipamentos ou de software capazes de fazer a leitura desses suportes.

Stephen Chapman, bibliotecário de preservação para iniciativas digitais do Centro de Preservação Weissman da Biblioteca da Universidade de Harvard, aponta alguns problemas que existem com as definições propostas de preservação digital supracitadas. Um dos pontos, diz respeito à proposta de que os arquivos estarão “mantidos e disponíveis” ao por longos períodos, e também a noção de que os arquivos serão distribuídos e reproduzidos “sem perda” (CHAPMAN, 2001).

A partir do exposto percebemos que a preservação digital apresenta-se como uma atividade relevante a qual demanda algumas estratégias, sobre as quais abordaremos a seguir.

3 ALGUMAS ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL

Por informação digital, o objeto de aplicação da Preservação Digital, ou material digital, entendem-se tanto os documentos que foram convertidos do suporte analógico para o digital (digitalizado), quanto os que já foram produzidos neste meio.

Ferreira (2006) e a *Digital Preservation Coalition* (2007) listam algumas das

estratégias de preservação digital, dentre as quais podemos citar:

- a) **Preservação de tecnologia** - conservação e manutenção de todo software e hardware necessários a correta apresentação dos objetos digitais;
- b) **Refrescamento** - transferência de informação de um suporte físico para outro mais atual;
- c) **Emulação** - uma forma de superar a obsolescência de software e hardwares através do desenvolvimento de tecnologias para imitar sistemas obsoletos em gerações futuras de computadores;
- d) **Migração para suportes analógicos** - consiste na reprodução de um objeto digital em suportes analógicos tais como papel, microfilme ou qualquer outro suporte analógico de longa duração;
- e) **Atualização de versões** - consiste em atualizar materiais digitais produzidos por um determinado software através de regravação em uma versão mais atual do mesmo;
- f) **Conversão para formatos concorrentes** - trata-se de converter um objeto digital para um formato que necessariamente não tenha sido desenvolvido pela mesma empresa que elaborou o software proprietário no qual este foi produzido. Também está restrito a alguns tipos de objetos. Pretende resguardar conteúdos da descontinuidade dos software, ou seja, quando o software não passar por versões atuais;
- g) **Adesão a padrões (inclui a técnica de normalização)** - adesão a padrões abertos estáveis e largamente utilizados ao criar e arquivar recursos digitais. Eles não estão presos a plataformas específicas de hardwares e software o que resguarda por algum tempo a mais o recurso digital da obsolescência tecnológica. Pode ainda ser auto imposto por instituições que geram recursos digitais ou impostas por agências que os recebam;
- h) **Migração a pedido** - esta técnica foi proposta para evitar a deformação de objetos digitais originais. Então toda migração feita de um formato para outro partirá sempre do original, e não de uma versão que já foi atualizada;
- i) **Migração distribuída** - trata-se do desenvolvimento e distribuição de conversores através da *net* que podem ser utilizados através de aplicações cliente. De acordo com Ferreira (2006) O *Lister Hill National Center for Biomedical Communications* possui um serviço web que converte objetos digitais de 50 formatos distintos para pdf. A Universidade do Minho está desenvolvendo serviço para disponibilização de várias centenas de serviços de conversão;
- j) **Encapsulamento** - reunir em conjunto com o recurso digital e o que quer que seja necessário para manter o acesso a ele. Isto pode incluir metadados, software visualizador e arquivos específicos constituintes do recurso digital;
- k) **Identificadores permanentes** - um meio de localizar um objeto digital mesmo quando sua localização muda;
- l) **Arqueologia digital** - é resgatar recursos digitais os quais tornaram-se inacessíveis pelo resultado da obsolescência tecnológica e/ou degradação da mídia, não é tanto uma estratégia em si mesma, mas uma substituta para quando materiais digitais ficaram fora de um programa de preservação sistemática;
- m) **Pedra da Roseta Digital** - o modelo baseia-se em três momentos diferentes os quais sejam: Processo de preservação do conhecimento [1] registro da codificação do

formato de arquivo e do conteúdo em binários, [2] recuperação dos dados e [3] reconstrução dos documentos a partir das especificações construídas na primeira etapa. Denomina-se Pedra da Roseta Digital por tratar-se de uma técnica que pretende traduzir para novos software os arquivos digitais advindos de tecnologias já obsoletas, com parâmetros que permitam uma tradução, assim como a Pedra da Roseta, descoberta por soldados franceses no ano de 1799 e que permitiu a tradução dos hieróglifos egípcios;

- n) **Metadados de preservação** - a metainformação de preservação é responsável por reunir, junto do material custodiado, informação detalhada sobre a sua proveniência, autenticidade, atividades de preservação, ambiente tecnológico e condicionantes legais “[...] tem como objectivo descrever e documentar os processos e actividades relacionadas com a preservação de materiais digitais. Ou seja, a metainformação de preservação é responsável por reunir, junto do material custodiado, informação detalhada sobre a sua proveniência, autenticidade, actividades de preservação, ambiente tecnológico e condicionantes legais” (FERREIRA, 2006, p. 54).

Vários modelos de metadados de preservação têm sido propostos por organizações internacionais. A *Open Archival Information System* (OAIS) constitui-se hoje um modelo de referência. De acordo com Arellano (2004), a estrutura conceitual da OAIS para metadados de preservação vem sendo usada por inúmeras organizações com o propósito de identificar seus elementos de metadados específicos, como por exemplo *Cornell University*, as Bibliotecas Nacionais da Austrália e da Nova Zelândia, *Online Computer Library Center* (OCLC), *Michigan Institute of Technology* (MIT) entre outras. Por tratar-se de um modelo referencial, não se constitui de uma implementação específica, mas delibera uma lista de condições de elementos que devem ser considerados no estabelecimento de um projeto de preservação de qualquer tipo de documento, seja digital ou analógico.

A não existência de um consenso quanto ao uso das técnicas e procedimentos, bem como a multiplicidade de formatos justifica o fato de que cada vez mais cresce o número de organismos preocupados com a preservação de acervos digitais, principalmente no que diz respeito à acessibilidade.

Uma outra questão pode ser colocada no que diz respeito a alguns conteúdos veiculados na rede tais como as páginas e sites. Estes ficam, na maioria das vezes, de fora das políticas de preservação, bem como ainda não se tem na literatura conteúdos que abordem o tema sem que seja de forma ainda bastante superficial. Sobre esse tipo de documento é que trataremos a seguir.

4 CONSERVAÇÃO DE DOCUMENTOS EM HTML E SITES, UMA QUESTÃO A RESOLVER

Não se menciona nestas estratégias citadas a questão dos documentos em HTML, ou das páginas de sites na *web*, que também se configuram como documentos digitais. Entretanto, como veremos mais adiante, alguns projetos desenvolvidos principalmente nos Estados Unidos, já trabalham com a preservação deste tipo de documento.

Neste sentido convém citar o Projeto Minerva, sobre o qual abordaremos mais adiante, o qual é liderado pela Biblioteca do Congresso Americano. A idéia do projeto

por enquanto é servir de repositório das páginas ou sites da Web. O que se faz no momento é apenas a guarda enquanto novas estratégias neste sentido vão sendo estudadas. Ainda não se tem conhecimentos de outras metodologias neste sentido.

Sobre a preservação deste tipo de documentos em especial a literatura é extremamente escassa. Porém sobre documentos digitais mais convencionais é possível encontrar diversas estratégias que vem sendo empreendidas.

Assim, faremos a seguir uma breve revisão acerca de técnicas de preservação digital mais mencionadas.

5 ANALISANDO AS TÉCNICAS DE PRESERVAÇÃO DIGITAL

Apesar de todas as técnicas já desenvolvidas, nenhuma parece ser completa o suficiente para garantir a preservação e acessibilidade dos documentos digitais, talvez o uso concomitante de algumas delas. Baseados no levantamento bibliográfico construímos o seguinte quadro das estratégias de preservação digital, suas descrições, vantagens e desvantagens:

Quadro 1: Principais estratégias de preservação digital, conceitos, vantagens e desvantagens

NOME	VANTAGENS	DESVANTAGENS
Preservação de tecnologia	▪ Conservação mantém a funcionalidade e interface do original; ▪ Conservação retarda o tempo quando outras estratégias de preservação serão exigidas; ▪ Fidedignidade ao objeto digital preservado;	▪ Necessidade de um espaço físico; ▪ Alto custo com a manutenção de equipamentos e software já obsoletos; ▪ Restrição quanto ao acesso; ▪ Só é viável como estratégia de curto e médio prazo; ▪ Suporte técnico e a facilidade de acesso vão aos poucos desaparecendo em um relativo curto espaço de tempo;
Refrescamento	▪ Garantia da integridade física do suporte do objeto digital; ▪ Pré-requisito para o sucesso de outras estratégias de preservação;	▪ Não garante acessibilidade por não tratar da obsolescência do software;
Emulação	▪ Recriar a funcionalidade e interface do original; ▪ Evita os repetidos custos com a migração; ▪ Importante na preservação de software;	▪ Necessidade de desenvolvimento de software emuladores cada vez mais potentes; ▪ São capazes de reproduzir parcialmente as funcionalidades e interfaces dos sistemas originais; ▪ Problemas de direitos autorais de software proprietários; ▪ Pressupõe que futuros usuários conheçam o funcionamento dos software já obsoletos;
Atualização de versões	▪ Adia a necessidade de utilizar estratégias mais elaboradas de preservação digital; ▪ As novas versões são oferecidas por um crescente número de fornecedores;	▪ Não é rotineiramente oferecido por todos os fornecedores; ▪ Só tem valor de curto e médio prazo; ▪ Mesmo quando existe as versões atualizadas não se pode presumir que dure indefinidamente; ▪ Sua disponibilidade continuada depende de demandas do mercado, as quais são

		notoriamente voláteis. E consequentemente pode cessar a sua viabilidade sem aviso prévio;
Adesão a padrões (inclui a técnica de normalização)	▪ Adiamento do tempo em que as estratégias de preservação mais onerosas serão necessárias; ▪ Simplificar e diminuir os custos das estratégias de preservação de longo prazo; ▪ Pode simplificar a migração e adquirir economias em escala na migração de itens similares; ▪ Pode beneficiar criadores tanto quanto a preservação a longo prazo. Ajuda a compartilhar alguns dos esforços sobre o ciclo de vida dos recursos; ▪ A adesão a padrões facilitará todas as outras estratégias de preservação digital; ▪ A padronização de formatos abertos promove a interoperabilidade entre sistemas de hardware e software distintos;	▪ Depende da disposição dos criadores e de suas habilidades para respectivamente compartilhar ou mudar a versão do arquivo; ▪ Padrões estáveis não são viáveis para alguns formatos; ▪ Mesmo quando padrões estáveis existem, eles próprios estão sujeitos a serem alterados nas suas atualizações; ▪ Alguns produtores de software não documentam muito bem o seu próprio padrão e ou prejudicam a conversão para um padrão aberto;
Migração para suportes analógicos	▪ Material não fica mais vulnerável a obsolescência tecnológica; ▪ O custo da conversão se dá uma única vez; ▪ Garantia de acessibilidade por longos períodos de tempo devido a durabilidade do suporte analógico; ▪ Funciona como uma estratégia interina de preservação enquanto são desenvolvidas infraestruturas para preservação digital mais apropriadas; ▪ Restrição quanto ao acesso; ▪ Necessidade de espaços físicos adequados;	▪ Somente possível para objetos digitais que possuam uma representação aproximada em suportes analógicos. Ex. textos ou imagens; ▪ Perda de funcionalidade da fonte digital; ▪ Viável apenas para documentos que não utiliza grandes recursos e funcionalidade da tecnologia digital; ▪ Não viável para fontes digitais mais complexas onde a perda da funcionalidade poderia diminuir ou destruir a usabilidade ou integridade da fonte; ▪ Perda das vantagens da tecnologia digital principalmente no que diz respeito a eficiência do uso de espaço; ▪ Custos com conversão para padrões de arquivo e armazenamento em condições de arquivamento (o último custo será recorrente e o custo cumulativo será significativo através do tempo);
Conversão para formatos concorrentes	▪ Quando os fabricantes mantêm uma compatibilidade numa versão atual;	▪ Muitos fabricantes descontinuem um software impossibilitando o uso da técnica;
Migração a pedido	▪ Uma vez criado o módulo capaz de ler as propriedades do formato de origem é necessário apenas desenvolver os codificadores específicos para cada formato de saída;	▪ Será necessário manter por longo período um largo conjunto de conversores para garantir a capacidade de conversão;

Identificadores permanentes	▪ É criticamente importante para ajudar a estabelecer a autenticidade de um recurso; ▪ Provê acesso a um recurso mesmo tendo mudado a sua localização; ▪ Supera problemas causados pela natureza instável das URL's; ▪ Permite interoperabilidade entre coleções;	▪ Não existe um sistema simples aceito por todos; ▪ Os altos custos de estabelecer ou utilizar uma solução deste tipo; ▪ É dependente da continuidade de manutenção do sistema de identificador permanente;
Migração distribuída	▪ Deixa transparente ao usuário as especificidades de conversor e plataforma; ▪ Redundância (original <i>versus</i> arquivo convertido) garante a fidedignidade; ▪ Vários caminhos de migração não restringe ao uso de conversores; ▪ Esta técnica pode ser associada a outras técnicas (normalização e migração a pedido); ▪ A criação de uma rede global de conversores poderá conduzir a uma redução generalizada dos custos de preservação;	▪ Poderá não ser adequada a todos os contextos; ▪ Multiplicidade de arquivos implica largo uso de espaço em <i>bytes</i>; ▪ Transferência de arquivos através da internet de volume alto de informações implica em custos altos; ▪ Exige largura de banda, segurança de dados e utiliza muito tempo em transferência de dados;
Encapsulamento	▪ Garantir todo o suporte informacional requerido para o acesso e manutenção como enquanto uma só entidade; ▪ Pode potencialmente superar algumas das maiores desvantagens de estratégias alternativas; ▪ Provê meios úteis de focar a atenção naqueles elementos que são necessários ao acesso;	▪ Pode produzir grandes arquivos com duplicação (i.e. visualizadores) através da coleção a menos que estejam linkados; ▪ O software encapsulado está ainda aberto à rápida obsolescência tecnológica;
Arqueologia digital	▪ Existe um crescente número de especialistas terceirizados oferecendo este serviço. Tem-se mostrado tecnicamente possível para recuperar uma ampla faixa de informação de mídias danificadas ou obsoletas (entretanto necessariamente no mesmo formato);	▪ Muito mais caro a longo prazo que utilizar estratégias de preservação digitais; ▪ Não é interessante se a informação a recuperar não tiver uma relação custo-benefício que a justifique; ▪ Há possibilidade de materiais serem perdidos por não justificarem os custos a serem investidos; ▪ Há risco de que algum material não seja recuperado com sucesso;
Pedra da Roseta Digital	▪ Tecnologia que independe de plataforma de hardware e software;	▪ Tecnologia ainda não desenvolvida completamente;
Metadados de preservação	▪ Existe um modelo que possui grande aceitabilidade em vários continentes (OAIS).	▪ Não preserva o próprio documento, mas descrições a respeito do mesmo.

Fonte: Construído pelos autores, baseado na *Digital Preservation Coalition* (2007) e Ferreira (2006).

A seguir faremos um breve apanhado dos projetos e ações a níveis mundial e nacional de preservação digital.

5.1 PROJETOS E AÇÕES DE PRESERVAÇÃO DIGITAL

Tão logo o computador passou a fazer parte do cotidiano das organizações e ainda a utilização de novos suportes midiáticos começou a se massificar, as problemáticas de preservação começaram a tomar novas formas e, portanto, a preocupar profissionais envolvidos na gestão de documentos. Além das iniciativas já mencionadas no início deste artigo, outras aqui serão mencionadas.

5.1.1 Tudo começo em Paris

Segundo Thomaz (2005), o início das preocupações acerca da preservação digital surgiu na comunidade arquivística em meados dos anos 60, em Paris, no *International Congress on Archives*. Porém na década seguinte é que as discussões acerca do tema ganharam mais espaço em consequência da crescente utilização de computadores na produção e gerenciamento de arquivos.

5.1.2 As experiências em preservação digital na Austrália

No âmbito australiano dois projetos merecem destaque que são o *Victorian Electronic Records Strategy* (VERS) e o *Preserving and Accessing Networked Documentary Resources of Austrália* (PANDORA). O primeiro projeto tem como objetivo produzir um padrão de gerenciamento e preservação de registros eletrônicos, ao passo que o segundo objetiva criar um repositório com acesso a longo prazo a publicações on-line.

5.1.3 O Reino Unido e a Irlanda em preservação digital

No Reino Unido e Irlanda vale citar o projeto liderado pelo *Consortium of University Research Libraries* (CURL) denominado CEDARS (*CURL Exemplars in Digital ARchives*). O projeto baseia-se no Modelo OAIS (*Open Archives Information System*) e propõe-se a criação de um modelo de arquivo para um sistema de arquivo e não explicitamente inclui um módulo de preservação. Assim como o CEDARS, outros projetos investigam o uso do modelo OAIS, como é o caso do PANDORA e também do *Networked European Deposit Library* (NEDLIB), na Europa.

5.1.4 O Reino Unido e os EUA se aliam para preservação digital

Outro projeto que vale ser mencionado é o *Creative Archiving at Michigan and Leeds: Emulating the Old and the New* (CAMILEON) o qual é financiado pelo *Joint Information Systems Committee* (JISC) no Reino Unido e o *National Science Foundation* (NSF) nos Estados Unidos. Tal projeto examina resultados acerca da implementação da técnica de emulação como uma estratégia de preservação digital. O projeto reconhece o potencial da emulação para a retenção da funcionalidade e interface dos objetos digitais. Ele objetiva desenvolver ferramentas, manuais, e custos para a emulação comparando-a a outras opções de preservação digital.

5.1.5 Experiências Estado-Unidenses em preservação

Thomaz (2005, p. 11) discute iniciativas dos EUA, dessa forma, “[...] considera-se o marco inicial das discussões sobre os desafios do documento eletrônico para os arquivos, o relatório preparado pelo professor Robert Henri Bautier para o *International Concil on Archives* – ICA, apresentado na *Conférence Internationale de la Table Ronde des Archives* – CITRA, realizada em Bonn, Alemanha, em 1971”.

A partir de então, seguiram-se a criação de grupos de estudos e organismos preocupados em aprofundar os conhecimentos nesta área. Assim sendo, vale citar alguns projetos e estudos de caso realizados nos contextos da Europa, Austrália e Estados Unidos.

Nos Estados Unidos, também vale mencionar o projeto do *National Archives and Records Administration* (NARA), o qual é conduzido pelo Centro de Supercomputação de San Diego e fundado pelo NARA. Visa desenvolver um arquivo permanente para suportar a inserção, armazenamento de arquivo, recuperação de informação e preservação de coleções digitais. Seu principal objetivo é preservar não somente os dados originais, mas também o contexto que permita que o dado seja interpretado. O projeto propõe uma abordagem para manutenção de dados digitais por centenas de anos através do desenvolvimento de um ambiente que suporte a migração de coleções para novos sistemas de software.

Ainda nos EUA encontramos uma iniciativa de pesquisa multinacional na qual escolas de arquivologia, engenharia computacional, instituições de arquivo nacional, e representantes de indústrias privadas colaboram para desenvolver conhecimento teórico e metodológico requerido para preservação permanente de registros autênticos usando sistemas eletrônicos. Trata-se do projeto “InterPARES – expressão latina para “entre pares” e acrônimo do *International Research on Permanent Authentic Records in Electronic Systems*” (THOMAZ, 2006, p. 121). O projeto consistiu de duas fases, onde na primeira, “[...] concluída em dezembro de 2001, examinou documentos gerados em bancos de dados e sistemas gerenciadores de documentos, cuja preservação seja determinada por requisitos administrativos e legais das organizações, enfatizando a questão da autenticidade ao longo do ciclo de vida dos documentos, e produziu como principais resultados o gabarito para análise de documento eletrônico (*Template for analysis*), os requisitos para avaliação e manutenção da autenticidade de documentos eletrônicos (*Requirements for assessing and maintaining the authenticity of electronic records*), os modelos funcionais para seleção e preservação de documentos eletrônicos e o glossário do projeto (*Glossary Committee*)” (THOMAZ, 2006, p. 121).

A segunda fase do projeto InterPARES, iniciada em janeiro de 2002 e concluída em 2006, concentrou-se em documentos digitais produzidos em sistemas dinâmicos, interativos e experienciais no curso de atividades artísticas, científicas e de governo eletrônico, enfatizando além da questão da autenticidade, as questões de fidedignidade e precisão (THOMAZ, 2006).

PRISM é um outro projeto cuja sigla vem das palavras americanas *Preservation, Reliability, Interoperability, Security, Metadata* que em português significam, respectivamente, preservação, confiabilidade, interoperabilidade, segurança e metadados. O projeto é liderado pela Universidade de Cornell e é financiado pela *Digital Library Initiative* para investigar e desenvolver políticas e mecanismos necessários à

integridade das informações nas bibliotecas digitais. O projeto foca na sobrevivência a longo prazo da informação digital, confiabilidade dos recursos e serviços de informação, interoperabilidade, segurança e metadados. De acordo com Lee et al (2002), o projeto a partir de 2002 passou a direcionar o projeto me torno do desenvolvimento de técnicas para monitoramento da integridade dos recursos informacionais distribuídos pela web e também reforçando políticas de preservação reservadas aos usuários e proprietários de coleções.

Outra iniciativa de preservação na *Web* é válido citar os esforços empreendidos pela Biblioteca do Congresso americano. Segundo Lee et al (2002) o *Computer Science and Telecommunications Board* (CSTB) do *National Academies* formou convênio com o Comitê de Estratégia em Tecnologia da Informação para fazer recomendações de preservação digital para a *Library of Congress*. Foi então produzido um relatório pelo comitê incluindo recomendações específicas para melhorar a infraestrutura tecnológica, particularmente na área das redes, bases de dados, e segurança em tecnologia de informação. A partir do site da referida biblioteca é possível visualizar alguns projetos em desenvolvimento, como por exemplo o MINERVA – *Mapping the Internet Electronic Resources Virtual Archive*. (Mapeamento de Arquivos Virtuais de Recursos Eletrônicos na Internet). De acordo com Deborah M. Thomas dos *US Library of Congress*, o projeto surgiu do entendimento da biblioteca de que os documentos criados apenas no ambiente *web* (*Born digital*) têm uma durabilidade muito curta na rede, porém precisam ser preservados para gerações futuras. (BIBLIOTECA NACIONAL, 2003).

Outra iniciativa de preservação digital que merece ser mencionada é a de preservação do *National Institute of Standards and Technology* (NIST) americano tem como foco as mídias digitais. Conforme relata Lee et al (2002), o primeiro trabalho sobre preservação de dados no NIST pode ser identificada como sendo pesquisa sobre a medição da vida útil de um *cd*, realizada por Podio nos anos 80. Seu trabalho forneceu uma base para uma metodologia padrão para a medição da vida útil de discos ópticos. Com o aumento do uso do armazenamento digital nas bibliotecas e nos arquivos das agências governamentais, o instituto percebeu a importância da preservação digital. Então se iniciou vários projetos para investigar diversos aspectos os quais são: prova de longevidade , prova de intercambialidade e interoperabilidade de discos ópticos para uso em sistemas de armazenamento de alta densidade, e desenvolvimento de um potente sistema de codificação o qual objetiva desenvolver um métodos para encontrar e recuperar informações úteis em discos danificados.

5.1.6 A Europa como conglomerado para preservação digital

O NEDLIB, mencionado anteriormente, trata-se de um projeto iniciado por um comitê permanente da *Conference of European National Libraries* (CENL) com financiamento do Programa de Aplicação Telemática da Comissão Européia. Oito bibliotecas nacionais na Europa, um arquivo nacional, e as principais editoras fazem parte do projeto. A Biblioteca Nacional dos Países Baixos lideram o projeto. O NEDLIB pretende desenvolver uma estrutura arquitetônica comum e ferramentas básicas para construção de sistemas de repositório para publicações eletrônicas. Porém seu principal objetivo é promover um melhor entendimento entre os pontos fortes e fracos de diferentes estratégias de preservação a longo prazo.

5.1.7 O Canadá e a preservação digital

Dentre os projetos canadenses, destaca-se o projeto de *E*-preservação, o qual tem sido desenvolvido através de um esforço cooperativo entre a Biblioteca Nacional do Canadá e a *Canadian Initiative on Digital Libraries* (CIDL). A *E*-preservação visa suprir os canadenses com fácil acesso a políticas e efetuar pesquisas sobre a criação, uso, e preservação de coleções digitais. O projeto inclui manual sobre os vários aspectos da preservação digital incluindo aquisição de materiais digitais, formatos e metadados. (LEE *et al*, 2002).

Essas são algumas das iniciativas de projetos da área a nível mundial. Em si tratando de Brasil as ações ainda parecem bastante incipientes se comparadas ao resto do mundo. Sobre o panorama da preservação digital no âmbito brasileiro, discorreremos a seguir.

5.2 RESERVAÇÃO DIGITAL NO BRASIL

Como vimos anteriormente, as preocupações acerca da preservação digital no mundo tiveram sua primeira expressividade no contexto mundial partiu de um trabalho realizado no *International Concil on Archives* (ICA), em 1970. O Brasil apesar de contar com representantes no *Commitee on Automation* do ICA, na área de Arquivologia, os eventos da área na mesma década pouco abordaram a temática e quando aconteceu, foi através de profissionais de outras áreas.

A falta de discussões na área é apontada no trabalho como sendo uma consequência da pouca utilização dos recursos eletrônicos nos arquivos de então. “Pelo diagnóstico da Comissão Especial de Preservação do Acervo Documental - CEPAD, iniciativa do Arquivo Nacional do Brasil, foi possível saber que, na época, apenas 0,02% dos arquivos existentes em Brasília usavam recursos informáticos. O diagnóstico restringiu-se ao Distrito Federal, mas se pode inferir que a situação do restante do país não seria muito diferente” (THOMAZ, 2006, p. 116).

Somente a partir dos anos 1990, começam a surgir publicações que abordavam o tema da tecnologia da informação e dos documentos eletrônicos. Na mesma época, o Governo Federal iniciou um processo crescente de informatização dos serviços públicos, acompanhado pelas demais instâncias políticas. O fato mais significativo no período, foi a instituição do programa Sociedade da Informação no Brasil, através do Decreto no. 3.294 de 15 de dezembro de 1999 e, a criação do Comitê Executivo do Governo Eletrônico - CEGE, como desdobramento do mesmo, (através do Decreto s/n de 18 de outubro de 2000), que passou a atuar com base no documento ‘Proposta de Política de Governo Eletrônico para o Poder Executivo Federal’, publicado no mesmo ano. (THOMAZ, 2006). No mesmo ensejo, foi publicado o ‘Livro Verde’, como uma proposta de inserir o Brasil no novo modelo de Sociedade que, então, consolidava-se ao redor do mundo.

A ação mais significativa no que concerne à problemática da Preservação de Acervos Digitais a nível de Brasil foi a reformulação da Câmara Técnica de Documentos Eletrônicos – CTDE, do Conselho Nacional de Arquivos - CONARQ, criada em 1994 e em funcionamento a partir de 2002. Trata-se de um grupo de trabalho multidisciplinar que tem por objetivo “[...] definir e apresentar ao Conselho Nacional de Arquivos normas,

diretrizes, procedimentos técnicos e instrumentos legais sobre gestão arquivística e preservação dos documentos digitais, em conformidade com os padrões nacionais e internacionais” (CONARQ, 2006).

Os trabalhos técnicos do CTDE estão organizados em dois grupos: Gestão e Preservação. O primeiro grupo produziu um glossário de termos referente a documentos eletrônicos, diversos artigos e apresentações e encontra-se em fase final de elaboração uma especificação de requisitos funcionais para sistemas eletrônicos de gestão arquivística de documentos convencionais e eletrônicos. Já para o segundo, a câmara também desenvolveu diversos artigos e apresentações e especialmente a Carta de preservação do patrimônio arquivístico digital adaptada à realidade brasileira, em sintonia com o programa *Memory of the World* da UNESCO. Essa carta foi aprovada pelo CONARQ em sua 34^a reunião plenária, realizada no dia 6 de julho de 2004 e tem sido objeto de ampla divulgação em simpósios, seminários, congressos e organizações governamentais e privadas.

Os trabalhos da Câmara atualmente estão direcionados para a elaboração de uma especificação de metadados de preservação e uma resolução sobre recolhimento de arquivos digitais (THOMAZ, 2006).

Nos dois tópicos que se seguem mencionaremos algumas das pesquisas brasileiras sobre a preservação digital no âmbito das ciências documentais.

5.2.1 Pesquisas brasileiras em preservação digital na área da Ciência da Informação

Dentre os esforços brasileiros cabe aqui citar também o Laboratório Líber - Laboratório de Tecnologia da Informação da UFPE. De acordo com Lima (2005) o projeto do laboratório “[...] ocupou-se desde cedo do interesse da ciência da Informação, com a meta de avançar no campo das bibliotecas digitais além da fronteira tradicional estabelecida pelo controle bibliográfico”, e que atualmente o laboratório consolidou-se como espaço integrador de pesquisa e desenvolvimento tecnológico, tomando como objeto de observação a hipermídia, investiga numa perspectiva transdisciplinar, a instalação social do ciberespaço.

O laboratório oferece bolsas de iniciação científica para alunos de graduação em Biblioteconomia, Informática e design. Hoje existem alguns desses alunos já ingressos em programas de pós desenvolvendo estudos nas áreas de Preservação Digital, Inteligência Artificial e Banco de Dados (GALINDO, 2005).

5.2.2 Preservação digital, saberes em confluência

Enquanto iniciativas brasileiras no sentido de produzir saberes que auxiliem na gestão e preservação dos documentos digitais, vale citar também os eventos estão promovendo discussões acerca do tema.

Podemos citar o II Seminário Internacional de Arquivos de Tradição Ibérica, organizado pelo Arquivo Nacional do Brasil, na cidade do Rio de Janeiro no período de 18 a 22 de novembro de 2002, inteiramente dedicado ao tema documentos arquivísticos eletrônicos: gestão e preservação e o XIV Congresso Brasileiro de Arquivologia, organizado pela Associação dos Arquivistas Brasileiro – AAB, também na cidade do Rio

de Janeiro, no período de 23 a 28 de abril de 2006, com grande ênfase na infra-estrutura arquivística para preservação de documentos digitais.

A Biblioteca Nacional organizou em 2003 um Curso Regional de Preservação Digital, o qual foi patrocinado pela UNESCO. A ideia é de preparar a Biblioteca Nacional para capacitar outras bibliotecas e arquivos a nível nacional para a preservação dos acervos digitais. Os cursos contaram com a presença de consultores professores internacionais um representante Francês da Unesco, o da *University of Califórnia, Library of Congress* e ainda nacionais, entre eles professores da PUC-RJ e ainda da Fundação biblioteca Nacional, entre outros.

6 ALGUMAS CONSIDERAÇÕES

Discutimos os conceitos e sua evolução de características da compreensão de uma prática baseada apenas de habilidades individuais, passando à responsabilidade e compromisso com a autenticidade dos objetos digitais preservados e, então, de gerenciamento de objetos digitais. Entendemos com a discussão dos conceitos que se trata, portanto, de preservar não só a informação em si e os seus suportes, mas o seu acesso para que não se percam seus valores sociais intrínsecos dentre outros, científicos e tecnológicos. O que justifica a preocupação de Chapman (2001) a respeito das estratégias e técnicas que são apontadas na literatura como sendo as viáveis para preservação digital, é que o que se percebe é que a manutenção da disponibilidade e a reprodução sem perda de informações ainda não são viáveis nas técnicas disponíveis.

A partir da literatura utilizada, percebemos que a prática da preservação digital constitui-se ainda em parte de um ideal que se pretende alcançar, visto que as técnicas hoje utilizadas ainda não foram testadas, mesmo porque ainda não existiu o tempo hábil para tal para consolidação da preservação digital enquanto realidade. Porém a importância cada vez maior que a informação adquire nas sociedades, principalmente no atual contexto da sociedade da informação, justifica os esforços empreendidos neste sentido.

Os países desenvolvidos têm desenvolvido esforços desde 1960, como é o caso da França, Austrália, Reino Unido, Irlanda, Canadá e principalmente e fortemente os EUA. Embora que as iniciativas ainda se baseiam em tecnologias desenvolvidas em contextos bastante diferentes do nosso, onde os recursos para pesquisa e desenvolvimento de tecnologias não são tão escassos e a consciência da importância da preservação já se apresenta de forma consolidada. Entretanto, apesar dos esforços empreendidos até agora, os avanços permanecem ainda parcialmente no campo teórico, visto que ainda não se passou tempo suficiente para comprovar a real eficácia das estratégias e dos projetos até hoje realizados.

No âmbito nacional, começamos tarde, em 1970. E timidamente, com maior participação das instituições ligadas à CTDE e ao CONARQ. Poucas pesquisas em Ciência da Informação têm se destacado neste aspecto, além do LIBER. Alguns eventos e um Curso como tentativa de multiplicação não se compara ao investimento realizado internacionalmente neste esforço. A pesquisa nessa área no Brasil necessita de apontar para soluções mais adequadas a nossa realidade. Parcerias precisam acontecer entre instituições de ensino, Governo, editoras, dentre outros para que se possa criar uma estrutura que viabilize os avanços nesta área.

Um outro ponto importante a ser observado é a necessidade, pelo menos a nível nacional, do estabelecimento de políticas e diretrizes eficazes no suporte à preservação digital, principalmente a respeito das questões legais e as relacionadas a financiamentos para a atividade. Percebe-se que muitas das técnicas de preservação digital abarcam soluções bastante pontuais, as quais demandam recursos, tanto humanos quanto financeiros, de custo elevado. O que leva a inferir que o conceito de preservação na prática não é algo fácil de se alcançar, tanto do ponto de vista dos recursos, quanto até mesmo dos potenciais das técnicas e procedimentos já elaborados.

Apesar de que as discussões acontecidas no Brasil tenham dado seus primeiros passos no âmbito da Arquivologia (ARELLANO, 2004), aos poucos, os profissionais bibliotecários estão se inteirando da problemática. Eventos da área, tais como o já mencionado curso Regional de Preservação Digital em 2003 e também o Simpósio Internacional de Bibliotecas Digitais em 2005, apontam para uma preocupação e também uma crescente discussão acerca do tema. Outros estudos deverão ser realizados a fim de aprofundar outras questões acerca desse tema. Existe a necessidade premente de que os esforços neste sentido persistam, visto que se trata de uma preservação cujos reflexos podem se dar sobre a preservação da memória cultural e científica de organizações, países, e porque não dizer, povos.

REFERÊNCIAS

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

BARRETO, A. A. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 122-127, maio/ago. 1998.

BARRETO, A. A. Os destinos da Ciência da Informação: entre o cristal e a chama. **Datagramazero**, n. 0, p. 1-8, dez. 1999. Disponível em: http://www.dgz.org.br/dez/F_I_aut.htm. Acesso em: 20 maio 2005.

BOERES, S. A. de A.; ARELLANO, M. A. M. Políticas e estratégias de preservação de documentos digitais. In: CIFORM: Encontro Nacional de Ciência da Informação, 6., 2005, Salvador, 2005. **Anais** [...]. Disponível em: http://www.cinform.ufba.br/vi_anais/docs/SoniaMiguelPreservacaoDigital.pdf. Acesso em: 13 dez. 2006.

BRASIL. Biblioteca Nacional. **Digital preservation regional course**. Rio de Janeiro: BN, 2003. CD-ROM.

CHAPMAN, S. **What is digital preservation?** [S.l.]: OCLC, [2007?]. Disponível em: <http://www.oclc.org/news/events/presentations/2001/preservation/chapman.htm>. Acesso em: 28 abr. 2007.

FERREIRA, A. B. de H. **Miniaurélio século XXI escolar**: o minidicionário da língua portuguesa. Rio de Janeiro: Nova Fronteira, 2001.

FERREIRA, M. **Introdução à preservação digital**: conceitos, estratégias e atuais consensos. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006.

HEMINGER, A. R.; ROBERTSON, S. B. **Digital Rosetta Stone**: a conceptual model for maintaining long-term access to digital documents. [S.l.]: ERCIM, [2007?]. Disponível em: <http://www.ercim.org/publication/ws-proceedings/DELOS6/rosetta.pdf>. Acesso em: 10 jul. 2007.

LAVOIE, B.; DEMPSEY, L. Thirteen ways of looking at...Digital Preservation. **D-Lib Magazine**, v. 10, n.7/8, p. 1-13, jul./ago., 2004. Disponível em: <http://www.dlib.org/dlib/july04/lavoie/o7lavoie.html>. Acesso em: 20 mar. 2007.

LEE, K.-H.; SLATTERY, O.; LU, R.; TANG, X.; MCCRARY, V. The state of the art and practice in digital preservation. **Journal Res. Natl. Standard Technology**, v. 107, n. 1, p. 93-106, jan./fev., 2002.

LEVY, P. **Cibercultura**. Trad. Carlos Irineu da Costa. São Paulo: 34, 1999.

LIMA, M. G. de. Experimentando novos modelos de investigação em Ciência da Informação: o caso Liber. In: CONFERÊNCIA SOBRE CIÊNCIA DA INFORMAÇÃO. **Anais** [...]. Porto-Portugal, 2005.

LIBER Laboratório em Tecnologia da Informação. Recife: UFPE, [2007?]. Disponível em: <http://www.liber.ufpe.br/portal/index.html>. Acesso em: 20 jun. 2007.

SOCIETY of American Archivists (SAA). **Digital preservation management**: implementing short-term strategies for long-term problems. [S.l.]: [s.n.], [2007?]. Disponível em: http://www.library.cornell.edu/iris/tutorial/dpm/eng_index.html. Acesso em: 24 jul. 2007.

THOMAZ, K. P. Gestão e preservação de documentos eletrônicos de arquivo: revisão de literatura – parte 1. **Arquivística .net**, Rio de Janeiro, v. 2, n. 1, p. 8-30, jan./jun. 2005.

THOMAZ, K. P. Gestão e preservação de documentos eletrônicos de arquivo: revisão de literatura – parte 2. **Arquivística .net**, Rio de Janeiro, v. 2, n. 1, p. 114-131, jan./jun. 2006.

UK Office for Library Networking (UKOLN). **Good practice guide for developers of cultural heritage web services**: digital preservation. [S.l.]: [s.n.], [2007?]. Disponível em: <http://www.ukoln.ac.uk/interop-focus/gpg/Preservation/>. Acesso em: 25 jul. 2007.