

Estratégias, critérios e políticas para preservação de documentos digitais em arquivos

Claudia Carmem Baggio

Mestranda em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC) – Florianópolis, SC – Brasil.

Especialização em Gestão em Arquivos pela Universidade Federal de Santa Maria (UFSM) - Santa Maria, RS – Brasil.

Diretora da Biblioteca do Centro de Ciências Rurais (CCR) da Universidade Federal de Santa Maria (UFSM) - Santa Maria, RS – Brasil.

<http://lattes.cnpq.br/0021582155166044>

E-mail: ccbaggio@gmail.com

Daniel Flores

Pós-Doutorado pela Fundación Carolina/Universidad de Salamanca (USal), Espanha. Doutor em Ciência da

Informação pela Universidade Federal do Rio de Janeiro (UFRJ) - Rio de Janeiro, RJ - Brasil. Doutor em Met y Líneas de

Inv En Biblio y Documentación pela Universidad de Salamanca - Salamanca - Espanha. Professor da Universidade Federal de Santa Maria (UFSM) - Santa Maria, RS – Brasil.

<http://lattes.cnpq.br/9640543272532398>

E-mail: dfloresbr@gmail.com

Resumo

Apresenta a preservação de documentos digitais sob a ótica da arquivística, identificando as estratégias, bem como os critérios e políticas da preservação digital. A pesquisa relatada neste artigo visa, de maneira geral, contribuir para as discussões e questionamentos em torno da preservação de documentos digitais em arquivos.

Palavras-chave: Arquivística. Preservação digital. Estratégias. Critérios. Políticas.

Strategies, policies and criteria for preservation of digital documents in archives

Abstract

This work show the preservation of digital documents from the viewpoint of archival, identifying strategies, as well as the requirements and policies of digital preservation. The research reported in this article aims, in general, contribute to discussions and questions about the preservation of digital documents in files.

Keywords: *Archival. Digital preservation. Strategies. Criteria. Policies.*

Estrategias, criterios y políticas para preservación de documentos digitales en archivos

Resumen

Presenta la preservación de documentos digitales bajo la óptica de la archivística, identificando las estrategias, bien como los criterios y políticas de preservación digital. La investigación relatada en este artículo visa, de manera general, contribuir para las discusiones y preguntas alrededor de la preservación de documentos digitales en archivos.

Palabras clave: *Archivística. Preservación digital. Estrategias. Criterios. Políticas.*

INTRODUÇÃO

A preservação digital intensificou-se no início do século XXI, como alternativa à preservação em material impresso, tendo em vista a importância da manutenção de informações para futuras gerações.

Os documentos se tornam digitais rapidamente devido à vulnerabilidade do ambiente digital. Estes são frágeis e não existe segurança suficiente para garantir sua preservação digital de longo prazo. A fragilidade pode estar em sua rápida degradação física, na obsolescência tecnológica, na complexidade e nos custos. Como respostas a este desafio, surgem estratégias de preservação digital que procuram incorporar todos os aspectos relacionados a esta problemática: custos, legislação, gestão, acesso, políticas e critérios. As estratégias reúnem soluções parciais ante um problema complexo, estando envolvidos, entre outros itens, a migração, emulação, arqueologia digital, criptografia, metadados, tecnologias livres, ou abertas como XML, HTML, Dublin Core.

Com o objetivo de resolver tal problemática, têm-se hoje diversas propostas para a preservação digital, porém ainda não existe uma proposta totalmente viável e abrangente, e assim alguns métodos necessitam ser adotados com o objetivo de reduzir o perigo da perda dos documentos digitais.

REVISÃO DA LITERATURA

Segundo as resoluções da United Nations Educational, Scientific and Cultural Organization (Unesco), como a “Carta sobre a Preservação Digital” (2004), pode-se afirmar que existe uma preocupação muito grande em salvar o patrimônio cultural e científico digital, garantindo assim, o acesso a esses recursos às funcionalidades dos registros eletrônicos autênticos. Essa carta foi aprovada pelo Conselho Nacional de Arquivos (Conarq), em sua 34ª reunião plenária, realizada no dia 6 de julho de 2004.

Segundo Cunha e Lima (2007, p. 15), as preocupações acerca da preservação digital partiram de um trabalho realizado no International Council on Archives (ICA), em 1970. No Brasil, apesar de contar com representantes no Committee on Automation do ICA, na área de arquivologia, os eventos da área na mesma década pouco abordaram a temática. Somente a partir da década de 90, começaram a surgir publicações que abordavam o tema da tecnologia da informação e dos documentos eletrônicos no Brasil. Cresce a produção de documentos digitais: planilhas eletrônicas, sistemas de *workflow*, Gerenciamento Eletrônico de Documentos (GED), bases de dados, entre outros. Daí a necessidade de instalar na comunidade arquivística brasileira discussões a respeito dos documentos digitais, a fim de buscar participação ativa dos arquivistas na gestão dos documentos e definição de estratégias de preservação.

A literatura da área, em sua ampla maioria, tem sido produzida por pesquisadores preocupados com a herança cultural e artística da humanidade. Estudos realizados enfocam alguns métodos de preservação de documentos digitais e mostram sua limitação ante o enorme crescimento desses registros. Na área da arquivística, existe ainda o desafio de estabelecer padrões, políticas e estratégias para as melhores práticas de arquivamento digital. Este artigo tem como objetivos específicos: verificar quais as estratégias atuais de preservação dos documentos arquivísticos em formato digital; identificar na literatura os critérios e políticas de preservação digital em arquivos; apresentar as estratégias de preservação de documentos digitais, considerando que estejam armazenados para garantir a perenidade de seus conteúdos; reunir conhecimentos sobre as necessidades de preservação digital de documentos arquivísticos.

Não existe segurança suficiente nos sistemas digitais para garantir a preservação digital de longo prazo. A fragilidade dos documentos digitais deve estar em sua rápida degradação física, na obsolescência tecnológica, na complexidade e nos custos.

O problema da preservação digital está no contexto dos objetos digitais, nas informações armazenadas e na maneira como foram armazenadas. Não são apenas seqüência de zeros e uns ou cadeias de bits (*bitstream*), mas o conjunto de decisões que definiram a formação básica do objeto informacional como um objeto único. Isso envolve a descrição de documentos que podem ser representados de diferentes formas, mas o seu conteúdo será sempre interpretado do mesmo modo, no momento da sua recuperação, ou seja, será a mesma cadeia de bits com seu significado original. (ARELLANO, 2008, p. 22).

A Câmara Técnica de Documentos Eletrônicos (CTDE) do Conarq, órgão encarregado de preparar diretrizes para a gestão e a preservação de documentos arquivísticos no Brasil, é a instituição que tem pesquisado mais profundamente a gestão e a preservação por longo tempo de documentos digitais, com abordagem arquivística. O Grupo de Documentos Eletrônicos (GDAE) participa com dois membros nessa Câmara, tendo auxiliado na produção da Carta de Preservação Digital, Glossário de Termos, duas resoluções e do Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos (e-ARQ). (CONARQ, 2004).

O e-ARQ Brasil especifica todas as atividades e operações técnicas da gestão arquivística de documentos, desde a produção, tramitação, utilização e arquivamento até sua destinação final. Todas essas atividades poderão ser desempenhadas pelo Sigad, o qual, tendo sido desenvolvido em conformidade com alguns requisitos, conferirá credibilidade à produção e à manutenção de documentos arquivísticos.

A Carta para a Preservação do Patrimônio Arquivístico Digital é um instrumento oficial sobre a importância e necessidade de ações para a preservação de documentos digitais. CONARQ/CTDE (2004)

A Carta apresenta três proposições principais:

1. Elaboração de estratégias e políticas
 - gestão arquivística de documentos;
 - instrumentalização das organizações arquivísticas;

- governo eletrônico;
 - ações cooperativas.
2. Estabelecimento de normas
 - padrões e protocolos;
 - requisitos funcionais;
 - metadados de preservação;
 - segurança da informação digital.
 3. Promoção do conhecimento
 - agenda de pesquisa científica;
 - ensino e formação de recursos humanos;
 - disseminação do conhecimento.

É fundamental também o uso de padrões internacionais de arquivamento de metadados para proporcionar uma gestão mais efetiva de estratégias de preservação de documentos digitais. A preservação de documentos eletrônicos no Brasil vem chamando a atenção de diversas entidades, entre as quais o Arquivo Nacional, que está conduzindo estudos sobre criação, manutenção e recuperação de arquivos a longo prazo, com base no modelo Open Archive Information System (OAIS).

O Modelo Referencial Open Archival Information System (OAIS), desenvolvido pelo Consultative Committee for Space Data Systems (CCSDS), foi aprovado para publicação como padrão ISO 14721:2003 em fevereiro de 2003, que descreve um enquadramento conceitual para um repositório digital genérico, aberto a todas as comunidades com garantias de confiabilidade. O modelo OAIS consiste em uma estrutura conceitual que disciplina e orienta um sistema de arquivo dedicado a preservar e manter o acesso à informação digital por longo prazo.

O modelo de referência OAIS é reconhecido como o mais importante documento conceitual voltado para a preservação digital. O objetivo é aumentar o grau de consciência e compreensão dos conceitos relevantes para o arquivamento de objetos digitais, especialmente entre instituições não arquivísticas. (SAYÃO, 2012, p. 28-29).

Segundo Innarelli (2006, p. 10), a aplicação do OAIS em arquivos consiste na organização de pessoas e sistemas, tendo como responsabilidade a preservação e o acesso da informação à comunidade interessada e

como foco principal a informação digital, as formas primárias de armazenamento e suporte da informação para os materiais de arquivos digitais e físicos.

A definição de uma política de preservação envolve, geralmente, todas as facetas de um arquivo. Implica a criação de políticas de avaliação e seleção de materiais, a identificação de esquemas de metainformação apropriados (metainformação descritiva, técnica, de disseminação, estrutural e de preservação), a definição de estratégias de preservação adequadas a cada classe de objetos digitais, a criação de planos de sucessão para a eventualidade da organização detentora da informação interromper a sua atividade, a utilização de modelos sustentáveis de financiamento, entre outros. (FERREIRA, 2006, p. 66).

Cada vez mais os recursos culturais e educacionais mundiais estão sendo distribuídos e acessados em forma digital em substituição ao papel. A informação digital está sujeita à obsolescência técnica e à degradação física, correndo risco de perdas. Para enfrentar os novos desafios criados pelo mundo digital, algumas estratégias têm sido propostas para a preservação digital. Os principais

métodos recomendados para a preservação podem ser agrupados em dois tipos: os estruturais e os operacionais. Os estruturais tratam da parte de investimentos, enquanto os operacionais tratam das medidas concretas acerca da preservação digital, sendo suas principais práticas: a migração de suporte e o refrescamento do meio (preservação física), a conversão dos formatos, a emulação (preservação lógica) e a preservação do conteúdo (intelectual). (ARELLANO, 2006).

Alguns manuais de preservação de documentos digitais foram finalizados pelo Arquivo Nacional, como a versão impressa em português dos manuais “Diretrizes do produtor” e “Diretrizes do preservador”, que foram um esforço conjunto com a Câmara dos Deputados no âmbito do convênio com o Projeto InterPARES. Os livretos impressos serão distribuídos para os órgãos da administração pública federal e as instituições parceiras do NA.

“Diretrizes do produtor” (figura 1) apresenta um conjunto de recomendações e boas práticas

Figura 1 – Diretrizes do produtor



Fonte: CONARQ 2011

Figura 2 – Diretrizes do preservador

InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems*

Informações para contato

Projeto InterPARES
School of Library, Archival and Information Studies
University of British Columbia
Vancouver, BC V6T 1Z3 Canadá
Tel: +1 (604) 822-2894
Fax: +1 (604) 822-1200

Dr. Luciana Dornett, Diretora do Projeto
+1 (604) 822-2507
luciana.dornett@ubc.ca

Randy Preston, Coordenador do Projeto
+1 (604) 822-2894
interpares.project@ubc.ca

A maior parte do financiamento para o Projeto InterPARES foi fornecida pelo Social Sciences and Humanities Research Council do Canadá, e pelas National Historical Publications and Records Commission e National Science Foundation, dos Estados Unidos. O financiamento complementar foi fornecido pela Hamilton Fund Research Grant, pelo Vice President Research Development Fund, pela Decaria de Artes e pela Escola de Biblioteconomia, Arquivologia e Ciência da Informação da Universidade de British Columbia.

Para mais informações, acesse nosso site: www.interpares.org

Tradução e revisão: Arquivo Nacional e Câmara dos Deputados
Edição: Câmara dos Deputados

* IPI Pesquisa Internacional sobre Documentos Arquivísticos Autênticos Permanentes em Sistemas Eletrônicos.

Dpv
Diretrizes do preservador
A PRESERVAÇÃO DE DOCUMENTOS ARQUIVÍSTICOS DIGITAIS: DIRETRIZES PARA ORGANIZAÇÕES

Elementos de preservação

6 Mg Migração	8 Vb Visibilidade				
4 Rr Requisitos de Integridade	9 Id Identificação	10 Pr Preservação	11 Ac Acuidade	12 Ar Armazenamento	13 Au Autenticidade
16 De Descrição	17 Ge Gestão	18 Ob Obsolescência	20 Mo Monitoramento		
25 Cc Custódia Contínua	26 Sa Salva	27 Rb Registo de base	29 Av Avaliação	30 Tr Transferência	
34 Re Recolha em este	37 Do Documentação				

Fonte: CONARQ 2011

voltadas para organizações que elaboram e mantêm documentos arquivísticos digitais. (CONARQ, 2011).

“Diretrizes do preservador” (figura 2) visa a orientar as instituições quanto aos melhores procedimentos para a preservação dos documentos arquivísticos digitais sob sua guarda. O trabalho de tradução foi feito a partir dos originais em inglês, desenvolvidos pelos pesquisadores da Universidade de British Columbia, no Canadá. O conteúdo integral dos manuais está disponível para o público em geral no site: http://www.interpares.org/ip3/ip3_documents.cfm?team=4. (CONARQ, 2011).

CRITÉRIOS PARA PRESERVAÇÃO DIGITAL

Planejamento da preservação é um documento estratégico que determina que procedimentos

devem ser praticados para evitar obsolescência tecnológica e danos físicos que possam acarretar perdas de informação de documentos digitais. O objetivo é garantir as condições técnicas, gerenciais e materiais para preservar informações digitais durante o período que a organização dela precisar.

Para a preservação digital de documentos em arquivos são necessárias ações como a preservação tecnológica, a migração, a emulação, o encapsulamento, adoção de padrões e protocolos, adoção de política de gestão documental e tecnológica, controle público de legitimidade, além de uma política pública que inclua pesquisa científica, assim como ações por partes de arquivos, em todos os níveis. Isso dependerá de legislação apropriada.

Alguns esboços de critérios encontravam-se presentes em propostas de preservação digital, que focalizavam:

1. o controle físico e intelectual dos materiais para garantir sua preservação a longo prazo, com credenciamento de formatos de arquivos requeridos;
2. a padronização de estratégias de migração, validação de dados e especificações normativas para mídia física;
3. a padronização no uso de metadados de preservação.

A maioria das propostas de critérios menciona a influência do modelo de referência OAIS/SAAI, Sistema de Arquivamento Aberto de Informação, assim como, também, elas apresentam um conjunto de critérios para o planejamento de repositórios digitais. No centro dessas propostas, a maioria proveniente da área arquivística, já estava contido o conceito de arquivos ou repositórios digitais confiáveis. A confiabilidade vem diretamente do princípio arquivístico de custódia responsável.

Na figura 3, Innarelli (2007) cita os 10 mandamentos da preservação digital:

Figura 3 – Os 10 mandamentos da preservação digital



Fonte: Innarelli (2007).

Segundo Innarelli (2007), levando em consideração a preservação digital e manutenção de “nossa memória”, foram idealizados “Os dez mandamentos da preservação digital”, os quais levam em consideração os princípios abordados pelo tema. Porém, a política deve ser estudada e definida pela empresa conforme suas necessidades. Vale a pena lembrar que documentos contam a vida de uma instituição e nem sempre devem ser eliminados somente por não possuírem valor legal.

METODOLOGIA DA PESQUISA

Esta pesquisa é documental e descritiva, com base em revisão de literatura, e relaciona as estratégias, critérios e políticas de preservação de documentos arquivísticos.

A pesquisa descritiva visa descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis. Envolve o uso de técnicas padronizadas de coleta de dados: questionário e observação sistemática. Assume, em geral, a forma de levantamento. Para realização deste trabalho foram feitos levantamentos bibliográficos buscando principalmente textos recentes. Primeiramente foi feita uma seleção de materiais encontrados em bases de dados *on-line* de periódicos na área de ciências da informação, artigos, teses, textos e páginas *web*, para um maior esclarecimento do tema a ser abordado pela pesquisa, como os ligados a arquivologia e tecnologia da informação.

RESULTADOS

As estratégias para a preservação digital representam as ações concretas a serem executadas no âmbito das organizações com vistas à preservação de documentos digitais por longo prazo. As principais ações de preservação de documentos digitais são: Migração, Emulação/Encapsulamento, Conservação de Hardware e Software, Reprografia, Refrescamento, Atualização de versões, Conversão

para formatos concorrentes, Normalização, A Pedra de Rosetta Digital, Arqueologia digital.

MIGRAÇÃO

A migração se preocupa com o conteúdo intelectual do documento, com a informação contida nele e não somente com o suporte, o objeto digital em seu formato digital. A importância da migração é transferir a informação para novos formatos enquanto for possível, preservando a integridade da informação.

Migração é a transferência de materiais digitais de uma plataforma computacional, hardware e software, em vias de descontinuidade para outra mais moderna, preservando assim a integridade dos objetos digitais. É a transferência periódica do recurso digital de uma mídia que está se tornando obsoleta ou fisicamente deteriorada, ou ainda menos estável para um suporte mais novo, de um formato ultrapassado para um formato mais atual ou padronizado. As estratégias baseadas na migração centram-se na preservação de seu conteúdo intelectual. (FERREIRA, 2006, p. 36).

A migração de documentos digitais deve se preocupar com a compatibilidade existente em relação as novas tecnologias, assim como, a estrutura interna e o conteúdo do material devem ser preservados e transferidos igualmente, para que o “novo” objeto seja uma representação fiel do original.

[...] Migração envolve transferir a informação Digital de uma mídia que está se tornando obsoleta ou fisicamente deteriorada, ou instável, para um suporte mais novo ou tecnologicamente mais atualizado, por exemplo: de um CD-ROM para outro CD-ROM mais novo, de disquete para um DVD; de um formato ultrapassado para um formato mais atual ou para um formato padronizado, por exemplo: de MSWord para Xml ou ASCII; de uma plataforma computacional em vias de descontinuidade para outra mais moderna [...]. (SAYÃO, 2005, p. 135).

Arellano (2008, p. 63) acrescenta que

o propósito da migração é preservar a integridade dos objetos digitais e assegurar a habilidade dos

clientes para recuperá-los, expô-los e usá-los de outra maneira diante da constante mudança da tecnologia. A importância da migração é transferir para novos formatos enquanto for possível, preservando a integridade da informação.

As instituições que comportam grande número de acervos utilizam a migração da informação digital, de um *hardware* ou *software* para outro, assegurando a sua recuperação, transferindo para novos formatos, preservando a integridade da informação.

Na escolha de uma estratégia de preservação, o tipo e a quantidade de recursos deve ser levado em consideração. A migração depende de fatores como experiência técnica, expectativa de usuários, orçamento institucional, equipamento existente e tempo disponível. A migração tem como objetivo manter os objetos digitais compatíveis com tecnologias atuais de modo a que um utilizador comum seja capaz de os interpretar sem necessidade de recorrer a artefatos menos convencionais, como por exemplo, emuladores. (FERREIRA, 2006, p. 36).

Segundo Ferreira (2006, p. 37), “existem diversas variantes de migração que poderão ser consideradas: migração para suportes analógicos, atualização de versões, conversão para formatos concorrentes, normalização, migração a pedido e migração distribuída”. Uma das vantagens da migração é que ela permite o acesso rápido ao recurso, já que o documento estará sempre em um formato compatível com os padrões tecnológicos de *hardware* e *software* da época, apesar da perda de algum atributo visual.

EMULAÇÃO/ENCAPSULAMENTO

A emulação consiste no desenvolvimento de um sistema que funcione da mesma forma que outro *software* já obsoleto, objetivando permitir a leitura de documentos que perderam suporte por *software* e *hardware*. A emulação necessita do desenvolvimento de técnicas de encapsulamento de documentos, seus metadados, *software* e especificações de emulador

a fim de assegurar sua coesão e prevenir sua corrupção. Os dados podem ser encapsulados junto com a aplicação de *software* utilizado na sua criação.

É também importante salientar que, com o tempo, o próprio emulador irá sofrer de obsolescência, havendo então necessidade de o converter para uma nova plataforma ou desenvolver um novo emulador capaz de emular o primeiro. (FERREIRA, 2006, p. 35).

Arellano (2008, p. 68) descreve a técnica de emulação como

as técnicas de emulação sugerem a preservação do dado no seu formato original, por meio de programas emuladores que poderiam imitar o comportamento de uma plataforma de hardware obsoleta e emular o sistema operacional relevante. O processo consiste na preparação de um sistema que funcione da mesma forma que outro do tipo diferente, para conseguir processar programas.

Segundo Ferreira (2006, p. 33), as estratégias de emulação baseiam-se essencialmente na utilização de um *software*, designado emulador, capaz de reproduzir o comportamento de uma plataforma de *hardware* e/ou *software*, numa outra que à partida seria incompatível.

A emulação é a única que pode preservar os objetos digitais originais e a capacidade de serem executados como eles foram inicialmente. Deve ser usada quando o recurso digital não pode ser convertido em formatos de *softwares* independentes, e migrados no futuro.

Para Cunha e Lima (2007), a estratégia de encapsulamento consiste em “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”.

O encapsulamento permite manter o formato original do recurso digital, contudo este deve fazer-se acompanhar por um conjunto de instruções que permitam interpretar os formatos do documento e o

conteúdo da informação. O encapsulamento é uma estratégia de preservação que consiste em preservar todos os detalhes de como interpretar o objeto digital. Preserva-se juntamente com o objeto digital, toda a informação (descrição formal e detalhada do ambiente de *software* e *hardware* requerido para seu funcionamento) necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores e ou emuladores.

CONSERVAÇÃO DE HARDWARE E SOFTWARE

Trata-se de uma estratégia dispendiosa e complexa em nível tecnológico, e encontra-se em declínio apesar de ser utilizada ainda por algumas empresas; a estratégia de conservação de *software* e *hardware* sob a perspectiva de preservação digital é cara e pouco prática, sugere que os dados digitais sejam preservados em mídia estável e copiados para nova mídia, caso seja necessário, e associados às cópias preservadas de uma aplicação de *software* original.

O *software* e o *hardware* tornam-se obsoletos em questão de anos, e, embora as versões sucessivas de programas possam ser compatíveis, os fabricantes de *software* normalmente não garantem a compatibilidade por longo período. “Criar *softwares* que considerem avançada compatibilidade e *hardwares* que facilitem compatibilidade entre sistemas é visto como vital para o contínuo acesso ao documento digital”. (CHILVERS, 2000).

A conservação do *software* trata-se de uma estratégia recente, a combinação de duas estratégias anteriormente referidas: o encapsulamento e a emulação. A preservação ou conservação da tecnologia implica desvantagens, como custo de operação, espaço físico, suporte técnico (manutenção), além do acesso à informação que fica restrito apenas ao local físico onde estão os *hardwares* preservados. É necessário que *software* e *hardware* se mantenham em condições que permitam consultar a informação neles armazenada.

REPROGRAFIA

Os arquivistas baseados no princípio arquivístico da “custódia responsável”, foram a primeira comunidade a apoiar o desenvolvimento dos repositórios digitais como aqueles que reproduzem os espaços físicos dos arquivos tradicionais.

A intersecção entre a preservação de documentos em suportes tradicionais e a preservação de objetos digitais ocorre em função do processo de digitalização. Em si, esse processo tem sido utilizado como vetor da preservação, pois os objetos digitais gerados atualmente podem conter uma alta fidelidade aos originais, o que permite poupar o acesso direto e o manuseio dos originais. (BODÉ, 2008, p. 20).

A tecnologia da digitalização é a mais flexível ferramenta de arquivamento, preservação e acesso a documentos através do armazenamento de suas imagens em formato digital. Para administrar tais serviços, criou-se o GED (Gerenciamento Eletrônico de Documentos), que converte informações, que podem ser voz, texto e imagens, para forma digital. Funciona com *softwares* e *hardwares* específicos e usa as mídias ópticas, em geral, para armazenamento. Um sistema de GED usa a tecnologia de informática para captar, armazenar, localizar e gerenciar versões digitais das informações.

O processo de digitalização engloba a conversão de documentos em papel, foto, microfilme, microficha, jaqueta ou cartão janela para uma imagem digital através de um *scanner*. Essa tecnologia possibilita o armazenamento de grandes volumes de documentos em algum meio digital.

A digitalização é um processo de conversão dos documentos para o formato digital, que consiste em unidades de dados binários, denominadas de *bits* - que são 0 (zero) e 1 (um), agrupadas em conjuntos de 8 *bits* (*binary digit*) formando um *byte*, e com os quais os computadores criam, recebem, processam, transmitem e armazenam dados. (CONARQ, 2010, p. 5).

O Conselho Nacional de Arquivos (Conarq) está disponibilizando em seu *site* a Resolução nº 31, que dispõe sobre a adoção das Recomendações para Digitalização de Documentos Arquivísticos Permanentes (figura 6). Tais recomendações visam auxiliar as instituições detentoras de acervos arquivísticos de valor permanente, na concepção e execução de projetos e programas de digitalização.

As recomendações buscam auxiliar as instituições detentoras de acervos arquivísticos de valor permanente, na concepção e execução de projetos e programas de digitalização. A digitalização de acervos é uma das ferramentas essenciais ao acesso e à difusão dos acervos arquivísticos, além de contribuir para a sua preservação, uma vez que restringe o manuseio aos originais, constituindo-se como instrumento capaz de dar acesso simultâneo local ou remoto aos seus representantes digitais como os documentos textuais, cartográficos e iconográficos em suportes convencionais, objeto desta recomendação. (CONARQ, 2010).

Essas recomendações tratam do seguinte:

1. captura digital em imagem de documentos planos e encadernados: impressos, manuscritos, mapas, plantas, desenhos, gravuras, cartazes, microformas, diapositivos, negativos, cópias e ampliações fotográficas;
2. padrões e boas práticas mínimas para a captura digital de imagens; produção de matrizes e derivadas, identificação do representante digital e controle de qualidade;
3. formatos digitais para representantes digitais matrizes e derivados;
4. metadados técnicos;
5. boas práticas gerais para armazenamento, segurança e preservação dos representantes digitais;
6. utilização de serviços terceirizados para a captura digital, armazenamento e acesso aos representantes digitais.

A reprodução em microfilme garante, se este for gerado com qualidade e devidamente preservado, acesso às informações contidas por cerca de meio milênio. “Sabe-se, contudo, que a reprodução a partir de um microfilme é de qualidade inferior à reprodução a partir do original, e o problema do documento multimídia permanece”. (SANT’ANNA, 2002, p. 40).

REFRESCAMENTO

A grande vantagem do refrescamento é tratar-se de um processo que não implica grande investimento em equipamentos, nem elevado grau de conhecimentos técnicos, a menos que a informação contida no suporte se encontre, de algum modo, protegida.

No contexto da preservação digital surge também o refrescamento (refreshing), como um método que consiste [...] na transferência de informação de um suporte físico de armazenamento para outro mais atual antes que o primeiro se deteriore ou se torne irremediavelmente obsoleto. (FERREIRA 2006, p. 33).

Por sua vez, a desvantagem mais apontada reside no fato de o refrescamento apenas resolver os problemas de degradação e de obsolescência dos suportes.

ATUALIZAÇÃO DE VERSÕES

A atualização de versões é, possivelmente, a estratégia de preservação mais vulgarmente utilizada pela generalidade dos utilizadores. Essencialmente, consiste em atualizar os materiais digitais produzidos por determinado *software* recorrendo a uma versão mais atual do mesmo.

Segundo Ferreira (2006, p. 37), é bastante comum encontrar aplicações de *software* capazes de abrir ou importar objetos digitais produzidos por versões anteriores dessa mesma aplicação. Essas aplicações permitem geralmente gravar os objetos importados no formato mais atual produzido pela mesma. Esta operação designa-se por atualização da versão do formato.

Ferreira (2006) cita que a atualização de versões é a atualização de materiais digitais produzidos por determinado *software* através de gravação em uma versão mais atual do mesmo.

CONVERSÃO PARA FORMATOS CONCORRENTES

O processo e a atualização de versões são geralmente controlados pela organização que desenvolveu a respectiva aplicação de *software*. A qualidade da migração depende, assim, da capacidade dos importadores fornecidos pelo fabricante e do grau de retrocompatibilidade oferecido pelo novo formato.

Idealmente, o fabricante deveria assegurar que todos os atributos presentes numa dada versão de um formato se encontram disponíveis na nova versão que o vem substituir. No entanto, independentemente do sucesso econômico de um fabricante ou produto de *software*, os formatos estão constantemente sujeitos à descontinuidade. Uma forma de garantir que os objetos digitais sobrevivem a este tipo de rupturas tecnológicas consiste em convertê-los para formatos de uma linha de produtos concorrente. (FERREIRA, 2006, p. 38).

Segundo Ferreira (2006);

a conversão de 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.

Existem, no entanto, formatos que não são dependentes de qualquer aplicação de *software*. Tal acontece com grande parte dos formatos de imagem (JPEG, TIFF, PNG). Isto possibilita que os objetos sejam convertidos entre formatos análogos, independentemente da aplicação utilizada na sua criação.

NORMALIZAÇÃO

A normalização de formatos pode ser implementada de diversas maneiras. Determinados repositórios procedem à conversão automática dos objetos recebidos para um formato único de preservação. Outros definem políticas de arquivo que limitam os formatos em que aceitam informação, significando isto, que cabe aos produtores da informação converter os seus objetos digitais para os formatos estipulados. (FERREIRA, 2006, p. 39).

Havendo um número controlado de formatos, a mesma estratégia de preservação poderá ser aplicada transversalmente a maior número de objetos digitais, o que poderá conduzir a uma redução generalizada dos custos de preservação.

Existe um leque variado de opções no que diz respeito a formatos para representação de imagens bidimensionais (BMP, GIF, JPEG, PNG, TARGA). Se durante o processo de ingestão, todas as imagens digitais forem convertidas para um único formato, futuras intervenções ao nível da sua preservação poderão ser realizadas de forma mais simples e, consequentemente, mais econômica. (FERREIRA, 2006, p. 39).

O argumento que suporta esta abordagem assenta no pressuposto de que os produtores de informação serão as entidades mais indicadas para avaliar a qualidade da conversão efetuada.

PEDRA DE ROSETTA DIGITAL

Nesta estratégia, em vez de preservar as regras que permitem descodificar o objeto digital, são reunidas amostras de objetos que sejam representativas do formato que se pretende recuperar. As amostras deverão existir num formato que possa ser diretamente interpretado pelo ser humano. Trata-se do conjunto de referência, por ex., a versão grega do decreto inscrito na Pedra de Rosetta (figura 7). Com essa informação seria possível inferir as regras necessárias para traduzir/converter o objeto original para qualquer formato contemporâneo.

Um exemplo de aplicação desta estratégia consiste em imprimir em papel um conjunto representativo de documentos de texto juntamente com a sua representação binária. No futuro, as regras necessárias para interpretar e migrar os objetos para um novo formato poderiam ser inferidas, comparando os documentos impressos com a sua representação binária. Esta estratégia deverá ser considerada apenas em situações em que todos os esforços de preservação falharam. Trata-se, sobretudo de uma ferramenta de arqueologia digital e não propriamente de uma estratégia de base para preservação de objetos digitais. (FERREIRA, 2005, p. 45).

Ferreira (2006) cita que,

a Pedra de Roseta Digital baseia-se em três momentos diferentes os quais sejam: processo de preservação do conhecimento; registro da codificação do formato de arquivo e do conteúdo em binários; recuperação dos dados e 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 *softwares* os arquivos digitais advindos de tecnologias já obsoletas, com parâmetros que permitam uma tradução.

ARQUEOLOGIA DIGITAL

Arqueologia digital é todo o processo de recuperação de informação armazenada em suportes danificados, degradados ou obsoletos e/ou formatos obsoletos.

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. (FERREIRA, 2006).

A arqueologia digital é uma solução parcial de preservação. Não há garantia de que seja possível recuperar a totalidade dos dados, nem a sua legibilidade. Mas existe ainda outra desvantagem: não existindo documentação suficiente, a interpretação dos dados só poderá ser feita por “adivinhação”, tornando-se difícil estabelecer a identidade,

integridade e contexto do documento digital recuperado.

Apesar de haver cada vez mais especialistas (“arqueólogos digitais”) a oferecerem este tipo de serviços, a arqueologia digital requer mão de obra altamente qualificada e, normalmente, é levada a cabo por organizações privadas de recuperação de dados, que possuem grande variedade de tecnologias e instalações especiais (por exemplo, divisões esterilizadas, onde se desmonta o hardware e suportes danificados). Trata-se, naturalmente, de serviços pagos, através dos quais é possível recuperar a cadeia de bits dos suportes danificados, sobretudo dos suportes magnéticos. (KENNEY *et al.*, 2003 *apud* FERREIRA, 2006, p. 64).

Por ser demasiado dispendiosa, a arqueologia digital é recomendada apenas para a recuperação e restauro de dados contidos em suportes danificados ou formatos obsoletos cujo valor ou importância dos dados justifique esse investimento, pois, por vezes, o valor da informação não justifica o custo da operação (WEBB, 2003, p. 145 *apud* FERREIRA, 2006, p. 64).

CONSIDERAÇÕES FINAIS

Como se pode observar na literatura utilizada, várias estratégias estão sendo estudadas e avaliadas a fim de que se consiga uma solução em longo prazo para preservação de documentos digitais. O esforço é justificado pela importância que a informação cada vez mais exerce na sociedade. Concluiu-se que muitas são as ações neste sentido, mas que ainda não existem consensos no que diz respeito às técnicas e metodologias utilizadas, apesar do número de estratégias de preservação ser crescente, nenhuma destas foi até ao momento universalmente aceita. A preservação dos documentos em formato digital desafia os arquivistas. O desafio está na criação de ambientes organizacionais apropriados para a ação.

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. (CUNHA; LIMA, 2007, p.19).

A respeito das estratégias e técnicas apontadas na literatura como as viáveis para preservação digital, o que se percebe é que a manutenção da disponibilidade e a reprodução sem perda de informações ainda não são viáveis nas técnicas disponíveis. Não há uma estratégia única de preservação até o presente momento, muitas estratégias têm sido propostas, mas nenhuma delas é apropriada para todos os tipos de objetos, situações e instituições, o que se tem são soluções específicas para casos específicos.

É importante que os documentos digitais sejam preservados, pois muitos dos documentos já se perderam ou estão em más condições. Garantir que as informações digitais, que são geradas continuamente, estejam acessíveis não apenas hoje, mas também para as gerações futuras, é um dos desafios mais urgentes da chamada era da informação. E, para lidar com esse desafio, são necessárias soluções não apenas técnicas, mas principalmente econômicas e sociais.

Esta pesquisa bibliográfica pretende servir como base para estudos mais aprofundados na área e contribuir para as investigações das transformações que os sistemas de informação vêm enfrentando. A utilização de padrões abertos e metadados tem se mostrado um fator fundamental, proporcionando facilidades para a gestão, recuperação e utilização da informação digital, bem como para a adoção de estratégias como migração e emulação. Esse é um caminho que deve ser percorrido pelas organizações responsáveis pela preservação e disponibilização de documentos digitais.

Tanto nas áreas da arquivística, quanto da ciência da informação, existe ainda o desafio de estabelecer padrões, políticas e estratégias para as melhores práticas de arquivamento digital. Há a necessidade de trabalhar com padrões técnicos mínimos (homogêneos), como padrões de resolução para

captura digital que permitam adequada visualização. O Conselho Nacional de Arquivos (Conarq), nesse sentido, está lançando uma Recomendação para Digitalização de Documentos Arquivísticos Permanentes.

A preservação de *hardware* se mostra inviável, sobretudo no que diz respeito ao acesso restrito à informação que pode ser consultada somente no local físico onde estão os *hardwares* preservados.

Acredita-se que estratégias como o encapsulamento são ainda muito caras e difíceis para se manter, já que exigem maior aporte tecnológico, além de especialistas comprometidos com o processo da representação estrutural e detalhes da interpretação do documento pelos *softwares* e *hardwares*.

Entre as atuais estratégias utilizadas, apresentadas nesta pesquisa, as que parecem mais viáveis e confiáveis para serem implantadas seriam a migração e a emulação, por serem técnicas mais utilizadas por organizações com grandes acervos, além de o foco ser o conteúdo e não apenas o suporte.

Preservar é uma medida de segurança que permite a transferência da informação para gerações futuras. A preservação em longo prazo de documentos digitais, no entanto, é um tema complexo, com questões ainda não definidas. Apesar da existência de várias estratégias, e que algumas possam ser mais viáveis que outras, ainda não há provas conclusivas que garantam a eficácia de nenhuma delas para longo prazo.

A partir da literatura utilizada, percebe-se que a prática da preservação digital constitui-se ainda parte de um ideal que se pretende alcançar; a preservação digital em arquivos implica um compromisso de investimento ao longo de todo o ciclo de vida do documento que, se não respeitado por curto período que seja, pode resultar em perdas irreversíveis e tornar todo o esforço despendido.

Os profissionais das áreas documentais podem e devem procurar trabalhar em conjunto com

especialistas em tecnologia para levar adiante as ações necessárias, assumindo uma postura pró-ativa diante da necessidade de providências para a preservação de documentos arquivísticos digitais, encontrar e propor as soluções cabíveis a altura do desafio que a sociedade do conhecimento moderno impõe. Estudar a preservação de documentos digitais é analisar como ele sofre alterações, ou virá a sofrer.

É fundamental realçar que existem muitos outros aspectos da preservação digital em arquivos que não foram descritos, tão pouco mencionados ao longo desta pesquisa. As informações presentes neste trabalho são suficientes para que o leitor seja capaz de identificar e compreender os principais conceitos associados a essa temática, bem como divisar estratégias e critérios adequados à preservação de objetos digitais em arquivos.

REFERÊNCIAS

ARELLANO, M.A. *Crítérios para a preservação digital da informação científica*. 2008. 356f. Tese (Doutorado em Ciência da Informação). Universidade de Brasília, Brasília, 2008. Disponível em: http://eprints.rclis.org/15412/1/Tese_Miguel_%C3%81ngel_M%C3%A1rdero_Arellano.pdf. Acesso em 10 jan. 2011.

BODÊ, E.C. *Preservação de documentos digitais: o papel dos formatos de arquivo*. 2008. 153 f. Dissertação (Mestrado em Ciência da Informação)-Universidade de Brasília, Brasília, 2008.

CHILVERS, A.H. *Managing long-term access to digital data approach: a metadata approach*. 2000. Thesis (Ph.D.) - Loughborough University, Inglaterra, 2000.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). Publicações digitais. *Carta para a Preservação do Patrimônio Arquivístico Digital: preservar para garantir o acesso*. Rio de Janeiro: UNESCO, 2004, 5 p. Disponível em: <http://www.documentoseletronicos.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm>. Acesso em: 09 abr. 2011.

- _____. *Diretrizes do produtor: a elaboração e a manutenção de materiais digitais: diretrizes para indivíduos*. Disponível em: <http://www.interpares.org/display_file.cfm?doc=ip2_creator_guidelines_booklet--portuguese.pdf>. Acesso em: 10 out. 2011.
- _____. *Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos: e-ARQ*, Rio de Janeiro: UNESCO, 2009, 178 p. Disponível em: <<http://www.conarq.arquivonacional.gov.br>>. Acesso em: 05 abr. 2011.
- _____. *Recomendações para digitalização de documentos arquivísticos permanentes*. Abril 2010. 28 p. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/recomenda/recomendaes_para_digitalizacao.pdf>. Acesso em: 24 mar. 2011.
- CUNHA, J. de A.; LIMA, M.G. *Preservação digital: o estado da arte*. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8, 28-31 out. 2007, Salvador. Anais do VIII ENANCIB. Salvador: UFBA/PPGCI; Ancib, 2007. Disponível em: <<http://www.enancib.ppgci.ufba.br/artigos/GT2--043.pdf>>. Acesso em 15 abr. 2011.
- FERREIRA, M. *Introdução à preservação digital: conceitos, estratégias e actuais consensos*. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. 85 p. Disponível em: <<https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>>. Acesso em: 10 abr. 2011.
- HISTÓRIA viva. *Pedra de Roseta digital*. 2009. Disponível em: <http://www2.uol.com.br/historiaviva/noticias/pedra_de_roseta_digital.html>. Acesso em: 21 jun. 2011.
- INNARELLI, H.C. *Preservação de documentos digitais: confiabilidade de mídias CD-ROM e CD-R*. 2006. 147 f. Dissertação (Mestrado Engenharia Mecânica) – Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, 2006. Disponível em: <<http://libdigi.unicamp.br/document/?code=vtls000384479>>. Acesso em: 07 abr. 2011.
- INNARELLI, H.C. *Os dez mandamentos da preservação digital*. In: SANTOS, V. B.; INNARELLI, H.C.; SOUSA, T.R.B. *Arquivística: temas contemporâneos*. Brasília: Senac, 2007.
- SANT'ANNA, M.L. *A digitalização de documentos de arquivo: o caso das plantas de parcelamento do solo de Belo Horizonte*. Dissertação (Mestrado). Escola de Governo da Fundação João Pinheiro. 2002. Disponível em: <<http://www.powerbrasil.com.br/pdf/leone2002.pdf>>. Acesso em: 30 abr. 2011.
- SAYÃO, L. F. *Preservação digital no contexto das bibliotecas digitais: uma breve introdução*. In: MARCONDES, C.H. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. Salvador: EDUFBA; Brasília: IBICT, 2005. p. 115-146.
- SAYÃO, L.F. *Preservação digital: uma breve introdução*. Disponível em: <<http://www.documentoseletronicos.arquivonacional.gov.br/Media/publicacoes/ctdemetadadospreservacaodigitalsayao.pdf>>. Acesso em: 10 ago 2010.
- SAYÃO, L. F. *1 Preservação digital: uma brevíssima introdução*. 2 OAIS: Open Archival Information System. Disponível em: <http://libdigi.unicamp.br/document/?view=8398> Acesso em: 30 abril. 2012.