



Universidade de Brasília – UnB

Faculdade de Ciência da Informação – FCI

**Metodologias aplicadas à preservação de documentos digitais
na Biblioteca Central da Universidade de Brasília**

Pedro Paulo Mizael Junior Cavalcante Reis

Brasília

Janeiro/2011



Universidade de Brasília – UnB

Faculdade de Ciência da Informação – FCI

**Metodologias aplicadas à preservação de documentos digitais
na Biblioteca Central da Universidade de Brasília**

Pedro Paulo Mizael Junior Cavalcante Reis

Monografia apresentada à Faculdade de
Ciência da Informação da Universidade de
Brasília como requisito parcial da obtenção do
grau de Bacharel em Biblioteconomia.

Orientadora: Profa. Dra. Ivette Kafure Muñoz

Brasília
Janeiro/2011

Título: Metodologias aplicadas à preservação de documentos digitais na Biblioteca Central da Universidade de Brasília.

Aluno: Pedro Paulo Mizael Júnior Cavalcante Reis.

Monografia apresentada à Faculdade de Ciência da Informação da Universidade de Brasília, como parte dos requisitos para obtenção do grau de Bacharel em Biblioteconomia.

Brasília, 19 de janeiro de 2011



Ivette Kafure Muñoz – Orientadora

Professora da Faculdade de Ciência da Informação (UnB)
Doutora em Ciência da Informação (UnB)



Murilo Bastos da Cunha – Membro

Professor da Faculdade de Ciência da Informação e Documentação (UnB)
Doutor em Ciência da Informação.



Dulce Maria Baptista – Membro

Professora da Faculdade de Ciência da Informação (UnB)
Doutora em Ciência da Informação (UnB)

*À minha carinhosa e dedicada mãe,
por todo apoio que tem me dado em
importantes momentos da minha
vida.*

AGRADECIMENTOS

Agradeço à colaboração de todos os colegas do curso, aos professores, a minha orientadora, aos funcionários da Faculdade de Ciências da Informação (FCI) e da Biblioteca Central (BCE) da Universidade de Brasília (UnB) que foram importantes para a realização deste trabalho.

RESUMO

O presente trabalho visa identificar a aplicação das metodologias de preservação de documentos digitais utilizadas na Biblioteca Central da Universidade de Brasília. Trabalha com a pesquisa bibliográfica sobre os conceitos de documento digital e de preservação digital, a fim de fundamentar a importância de sua utilização. Tem como metodologia o estudo de caso e a observação das formas de preservação digital aplicadas no acervo da Biblioteca Central da Universidade de Brasília, a fim de verificar a utilização das mesmas. Obteve-se como resultado da pesquisa que a Biblioteca Central da Universidade de Brasília utiliza-se de duas metodologias aplicadas a preservação digital, a migração de formatos de arquivo e a normalização .

Palavras-chave: preservação digital, documento digital, Biblioteca Central da Universidade de Brasília.

ABSTRACT

This study aims to identify the methodologies of preservation of digital documents used in the University of Brasilia Central Library. Works with the literature on the concepts of digital documents and digital preservation in order to substantiate the importance of its use. Its methodology case study and observation forms of digital preservation applied in the Library Center at the University of Brasilia, to verify their use. Was obtained as a result of research that the University of Brasilia Central Library of the makes use of two methods applied to digital preservation, migration of file formats and standards.

Keywords: digital preservation, digital document, University of Brasilia Central Library.

LISTA DE FIGURAS

Figura 1 – Diferentes níveis de abstração de um objeto digital.....	17
Figura 2 – Refrescamento de mídia.....	21
Figura 3 – Migração de formatos de arquivo.....	24
Figura 4 – Página inicial da RIUnB.....	27
Figura 5 – Página inicial da BDM.....	27
Figura 6 – Página inicial da BDTD.....	28
Figura 7 – Página inicial da BDS.....	29
Figura 8 – Exemplo de migração de formato de arquivo.....	33

LISTA DE TABELAS

Tabela 1 - Classificação de formatos de arquivo pelo conteúdo.....	16
Tabela 2 - Quantidade de Documentos Digitais na BCE/UnB.....	30
Tabela 3 - Formatos de Arquivos.....	30
Tabela 4 - Distribuição dos formatos de arquivo nos repositórios.....	31
Tabela 5 – Metodologias de preservação digital aplicadas nos repositórios....	31

SUMÁRIO

1. INTRODUÇÃO.....	10
2. JUSTIFICATIVA.....	11
3. PROBLEMA.....	12
4. OBJETIVOS.....	13
4.1. Objetivo Geral.....	13
4.2. Objetivos Específicos.....	13
5. METODOLOGIA.....	14
6. REVISÃO DE LITERATURA	15
6.1. Documento digital.....	16
6.1.1. O que é documento digital.....	17
6.1.2. Formato de arquivo.....	17
6.2. Preservação digital.....	18
6.2.1. Preservação de Tecnologia.....	20
6.2.2. Refrescamento.....	20
6.2.3. Migração.....	21
6.2.4. Emulação.....	22
6.2.5. Normalização.....	23
7. ESTUDO DE CASO DA BIBLIOTECA CENTRAL DA UNIVERSIDADE DE BRASÍLIA.....	25
7.1. Objeto da pesquisa.....	26
7.2. Coleta de Dados.....	29
7.3. Análise dos dados coletados.....	31
8. CONCLUSÃO.....	34
9. REFERÊNCIAS.....	37

1. INTRODUÇÃO

A latente necessidade da humanidade de expandir o conhecimento e o surgimento de novas tecnologias como a Internet, viabilizou uma maneira mais rápida para a troca de documentos via rede. Para tanto se torna necessário à aplicação de técnicas. A fim de manter a integridade da informação é preciso aplicar medidas para preservar os documentos digitais.

Devido ao crescente aumento na utilização de documentos digitais por parte de Centros de Informação, se torna necessário o conhecimento de tais técnicas para melhor armazenar e preservar estes documentos.

Neste trabalho pretende-se abordar os conceitos sobre documento digital e um estudo de caso sobre a utilização as metodologias de preservação digital pela Biblioteca Central da Universidade de Brasília, a fim de demonstrar sua aplicação e importância.

Este trabalho teve como metodologia um estudo de caso sobre a Biblioteca Central da Universidade de Brasília para melhor conhecer os documentos digitais existentes nesta biblioteca. Conseguiu-se analisar os 4 repositórios existentes na Biblioteca Central da Universidade de Brasília, sendo eles: RIUnB, BDM, BDTD, BDS.

2. JUSTIFICATIVA

Diante do constante aumento da utilização dos documentos em formato digital para a disseminação da informação e a necessidade de manter estes documentos acessíveis ao longo do tempo, torna-se importante saber como a Biblioteca Central da Universidade de Brasília utiliza-se das metodologias aplicadas a preservação dos documentos digitais para a preservação dos documentos para a posteridade.

Acredita-se ser de importância a pesquisa sobre preservação digital, pela escassez de literatura nesta área do conhecimento. E também pela pouca existência de pesquisas nesta área tendo como foco a Biblioteca Central da Universidade de Brasília.

3. PROBLEMA

A problemática deste trabalho parte do interesse em saber sobre a utilização das metodologias de preservação digital aplicadas aos documentos digitais existentes na Biblioteca Central da Universidade de Brasília.

4. OBJETIVOS

4.1. Objetivo Geral

O objetivo geral deste trabalho é identificar as metodologias aplicadas para a preservação dos documentos digitais na Biblioteca Central da Universidade de Brasília. Pretende-se identificar quais as metodologias de preservação digital e como se aplicam nos documentos da Biblioteca Central da Universidade de Brasília.

4.2. Objetivos Específicos

Para a identificação das metodologias aplicadas a preservação dos documentos digitais na Biblioteca Central da Universidade de Brasília será necessário responder aos seguintes objetivos específicos:

- Caracterizar os conceitos de documento digital e preservação digital através de revisão de literatura;
- Descrever as metodologias aplicáveis à preservação de documentos digitais;
- Analisar quais as metodologias de preservação digital utilizadas pela Biblioteca Central da Universidade de Brasília.

5. METODOLOGIA

O presente trabalho assumirá como metodologia o caráter de pesquisa qualitativa exploratória com um estudo de caso que visa uma aproximação maior para melhor entendimento sobre o objeto de estudo, a Biblioteca Central da Universidade de Brasília. Esta pesquisa se dividirá em dois momentos:

- Será realizada uma pesquisa bibliográfica de literatura que aborde os temas, tais como: documentos digitais, metodologias de preservação digital e aplicação destas metodologias.
- No segundo momento ocorrerá a realização de um estudo de caso buscando uma aproximação com o objeto de estudo e melhor entendimento da utilização das metodologias de preservação digital na Biblioteca Central da Universidade de Brasília para observar a aplicação das metodologias de preservação digital.

6. REVISÃO DE LITERATURA

6.1. Documento digital

Desde seu início a humanidade tem a necessidade de se comunicar e registrar seus conhecimentos e informações. Para tanto têm sido criados formas de registro da informação em diferentes suportes há milhares de anos. Conway (2001, p. 12) relata que “a mais antiga e conhecida evidencia da escrita – os signos pictóricos sobre placas de argila secas ao sol – remontam aproximadamente a seis mil anos atrás. [...] exemplares de escrita sumeriana e babilônica existem hoje nos principais centros de pesquisa do mundo”.

Com o passar do tempo os suportes evoluíram. Depois das placas de argila surgiu o papiro que, por ser muito frágil, tinha como maior deficiência o fato de apenas poder ser utilizado de um lado. Algum tempo depois se passou a utilizar também o pergaminho, que era um suporte fabricado de couro de animais (porcos, carneiros, etc.), com maior resistência em relação ao papiro. Posteriormente criou-se o papel, feito de celulose. Nos dias de hoje o papel, ainda é o principal suporte para gravar/transmitir informações (BAPTISTA, 2009).

Gravações de som em discos também foram muito utilizadas no final do século XIX e no início do século XX. Posteriormente estas gravações foram acrescentadas aos filmes e gravações televisivas produzidos na época. Bodê (2008, p. 41) cita como exemplo de evolução posterior: “No final do século XX surgiram os *Compact Discs* (CDs), inicialmente para gravações de áudio, surgindo depois os modelos específicos para vídeo (DVD’s)”.

Em meados da década de 1940, após a Segunda Grande Guerra Mundial, os computadores passaram a ser mais utilizados para armazenar informações. Mas o grande ápice deste invento vem a culminar com o que Bodê (2008, p. 41) escreve: “O uso cada vez maior de computadores inicialmente pelas grandes corporações, mas a partir da década de 1980 do século XX, também pelo cidadão comum representou um grande salto para o **registro**, o **armazenamento** e a **recuperação** de documentos” (grifo do autor).

6.1.1. O que é documento digital?

Documentos digitais são definidos de acordo com Thibodeau (2002 *apud* FERREIRA, 2006, p. 21) “como todo e qualquer objeto de informação que possa ser representado através de uma sequência de dígitos binários” (zero e um). Esta representação pode ser originada de duas formas: documentos já criados em meio digital ou documentos nascidos em formato analógico e digitalizados posteriormente.

O documento digital pode estar inscrito em diferentes suportes físicos (*CD-ROM*, *DVD*, *Hard Drive*, *pen drive*, etc.). Dentre os suportes existentes há disparidades do modo de armazenamento, isso se deve pela diferença entre os materiais e lógicas utilizados, por exemplo, um *CD-ROM* armazena um arquivo em orifícios refletores dispostos em espiral formada em uma base de policarbonato, já em um *Hard Drive* (HD) essas informações são armazenadas de forma magnética em um disco metálico localizado em seu interior.

Para a interpretação das informações contidas nos suportes, existem inúmeros softwares responsáveis pela transformação dos dados em

informação. Que pode ser ou não interpretada pelo ser humano, tornando-se assim o que Ferreira (2006, p. 23) chama de “objeto experimentado”, como podemos ver na figura 1.

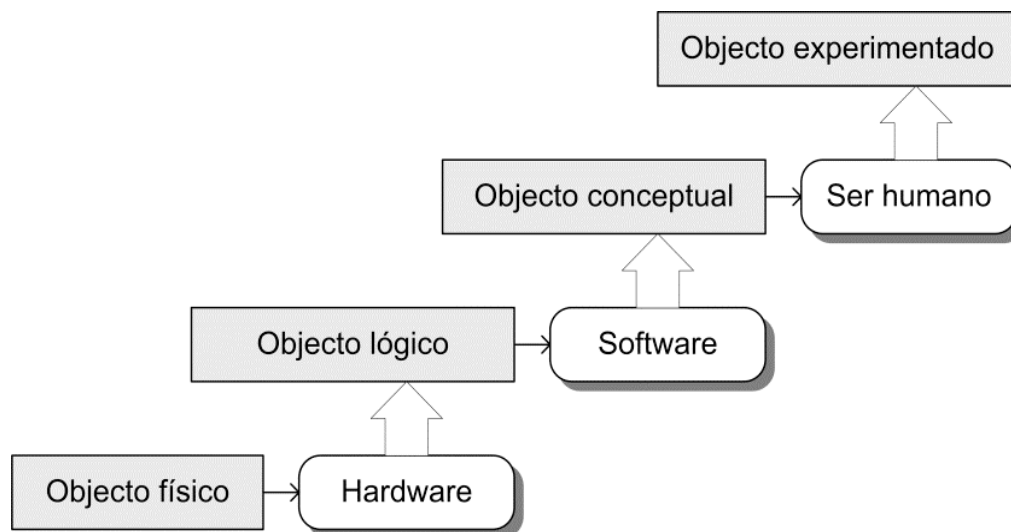


Figura 1 – Diferentes níveis de abstração de um objeto digital (FERREIRA, 2006 p. 23)

6.1.2. Formato de Arquivo

Para definir o conceito de formato de arquivo Bodê, (2008, p. 53) cita um relatório chamado *The Representation and Rendering Project*, da Universidade de Leeds, Reino Unido que o define como:

Em seu nível mais baixo, objetos digitais são sequências de zeros e uns que representam dados codificados. Diferentes Formatos de Arquivo especificam como esses códigos representam conteúdo intelectual criado por um autor de um objeto digital. (UNIVERSITY OF LEEDS, [s.d.], p. 4).

Esta definição “chama a atenção para o fato de que um formato de arquivo qualquer especifica como um determinado conteúdo está estruturado” (BODÊ 2008, p. 54). Mais adiante o mesmo autor afirma que “Uma especificação para um formato de arquivo X nada mais é senão a determinação de quais informações (conteúdo, metadados e outros) e qual a

ordem sequencial (ou não) de gravação no arquivo físico composto por **códigos binários** [...]” (grifo do autor p. 55). O que pode se entender analisando estas definições é que diferentes formatos de arquivo possuem diferentes estruturas de organização de suas sequências de zeros e uns (*bitstream*) e características peculiares a cada um deles. Como por exemplo poderemos ver na Tabela 1.

Tipo predominante de conteúdo	Exemplos de Formatos de Arquivo
Texto	RTF, Openoffice, ODF, DOC, AmiPro e outros
Imagens Fixas	BMP, EXIF, GIF, JPG, TIFF e outros
Imagens 3D	CAD, BIFF, X4D e outros
Sonoro	MEU, KAR, MP3, MP4 e outros
Imagens em movimento	AVI, MOV, MPEG, SWF e outros

Tabela 1 - Classificação de formatos de arquivo pelo conteúdo (BODÊ, 2008 p. 58)

Em um formato de arquivo, seja ele qual for, não estão apenas inseridos os dados referentes ao conteúdo presente no documento mas também existem metadados que são definidos por Ferreira (2006) como sendo: “informação utilizada para descrever um determinado objeto ou recurso”. Para Lopez (2004, p. 71) “[...] o metadados garante que o conteúdo informativo não seja desprovido dos dados contextuais da origem arquivística do ato administrativo que o produziu, além de garantir a permanência de seu valor probatório.” Não obstante, Bodê (2008, p. 55) diz que “[...], um arquivo real necessita também de **metadados** mínimos, como a data de criação do arquivo, o tamanho desse arquivo em *bytes*, o *software* utilizado para a criação, etc.” (grifo do autor).

6.2. Preservação digital

Com o crescimento da criação de documentos digitais, cresce também a necessidade de torná-los acessíveis ao longo do tempo. Não podemos esquecer que a confiabilidade/durabilidade dos hardwares de armazenamento destes documentos não é tão grande. Os softwares também podem acarretar problemas quanto à incompatibilidade com as versões dos documentos, impossibilitando assim o acesso. Para tentar solucionar estes e outros problemas é necessário recorrer às técnicas de preservação digital.

Para definir o conceito de preservação digital Boeres (2004, p. 26) cita a definição retirada do *Modern Information Retrieval Glossary* que diz:

Garantir que um objeto digital continue a estar acessível e útil ao longo do tempo, o que geralmente requer uma conversão tanto da mídia (copiando de um antigo formato para um novo antes que o velho não seja mais lido) quanto à conversão do formato (mudando de alguma estrutura de arquivo ou codificando para um mais novo, que continuará a ser usado e entendido).

Não obstante, Cunha e Cavalcanti (2008, p. 290) definem preservação digital como sendo um “processo de armazenamento, em condições adequadas para o uso, de documentos ou objetos produzidos em formato digital.”

Pode-se inferir através destas duas definições que o grande propósito da preservação digital é conseguir manter os documentos digitais acessíveis quando solicitados em qualquer época. Para isso devem-se utilizar estratégias de preservação digital, a fim de manter o documento sempre à disposição do usuário.

As técnicas ou estratégias de preservação digital podem estar associadas tanto a parte física (*hardware*), como a parte lógica (formato do documento, *software*, entre outros). Devem ser aplicadas de acordo com a

necessidade e o tamanho do acervo em questão. Nos próximos subitens serão ilustradas algumas destas estratégias para preservar os documentos digitais.

6.2.1. Preservação de Tecnologia

Esta técnica trata da preservação do *hardware* e do *software*. Consiste na guarda dos equipamentos e dos aplicativos para serem utilizados de maneira “original” de acordo com a geração do documento. Trata-se sobretudo da criação de museus de tecnologia. Aqui, o foco da preservação não se concentra no objeto conceitual, mas sim na preservação do objeto digital na sua forma original (FERREIRA, 2006, p. 32).

Esta técnica tem como grande problema o alto custo para manutenção dos equipamentos, além de ocupar grande espaço físico.

6.2.2. Refrescamento

Um objeto torna-se persistente no momento em que é inscrito no suporte físico de armazenamento (disco rígido, CD-ROM). Garantir a integridade do suporte é fundamental para que a informação nele armazenada possa ser corretamente interpretada (FERREIRA 2006, p. 33).

O refrescamento é uma técnica que consiste na troca da mídia em que o documento está inscrito, por uma em melhores condições (preferencialmente nova) ou para uma mais avançada tecnologicamente. Nesta transferência o documento não sofre nenhuma alteração.

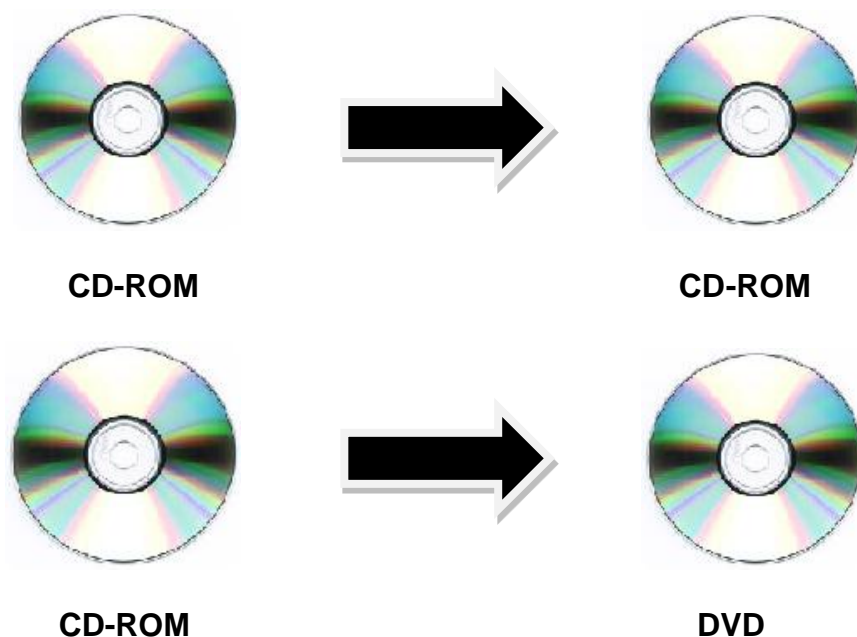


Figura 2 – Refrescamento de mídia

Em muitas situações o refrescamento irá ocorrer entre suportes de tecnologias diferentes, geralmente, tendo o novo suporte maior capacidade. Assim, o novo suporte poderá aglutinar o conteúdo de várias unidades do suporte obsoleto [...] (AMORIM *et al.*, 2005, p. 38).

6.2.3. Migração

A migração periódica da informação digital a partir de um ambiente *hardware* ou de um *software* para outro é a estratégia operacional para a preservação digital mais frequentemente usada pelas instituições detentoras de grandes acervos (MÁRDERO ARELLANO, 2008, p. 62).

Márdero Arellano (2008, p. 63) acrescenta que:

O propósito da migração é preservar a integridade dos objetos digitais e assegurar a habilidade dos clientes para recupera-los, expô-los e usa-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.

Esta técnica pode ser representada em várias operações como: transferência de um arquivo de texto, conversão de uma imagem, ou atualização de versão etc.

Amorim (*et. al.* 2005, p. 39), elencam algumas análises que precedem o processo de migração:

- Avaliar criteriosamente os formatos digitais com o objetivo de utilizar aquele que mais se adapte aos objetivos da migração.
- Realizar testes preliminares antes da completa execução da migração, para verificar se o procedimento está de acordo com as expectativas.
- Estabelecer um controle de qualidade dos resultados da migração.
- Gerar metadados com informações sobre o processo visando à documentação das alterações pelas quais o registro venha a passar durante sua existência.

6.2.4. Emulação

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

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.

Dando outro enfoque, Ferreira (2006, p. 34) destaca que “tal como acontece em estratégias baseadas na preservação de tecnologia, as técnicas de emulação centram-se na preservação do objeto lógico no seu formato original.” Com base na afirmação pode-se inferir que diferente da migração a técnica de emulação mantém os dados do formato do arquivo “originais” e a tecnologia aplicada se aproxima da “original”.

“A emulação tem algo em comum com a preservação da tecnologia e envolve os seguintes critérios:

- a) Uma mídia de informação estável deve ser armazenada em um sistema;
- b) A mídia digital será preservada enquanto o documento está sendo preservado como uma “*machine language*”;
- c) Dados serão representados como um formato de mídia novo através da conversão e reformatação;
- d) A integridade da informação digital será incrementada através de um processo de cópias;
- e) A aplicação original deve ser preservada e usada para criar ou acessar o recurso digital” (MÁRDERO ARELLANO, 2008, p. 70).

Alguns aspectos podem ser levantados em desfavor da migração. Dentre eles que: o usuário tenha que se inteirar do funcionamento de plataformas anteriores ao seu tempo para acessar o documento.

6.2.5. Normalização

A normalização é uma técnica segundo Thibodeau (2002 *apud* FERREIRA, 2006 p. 38) que “tem como objetivo simplificar o processo de preservação através de redução do número de formatos distintos que se encontram no repositório de objetos digitais”. A utilização desta técnica poderá implicar no futuro, uma redução nos custos de preservação dos documentos digitais.

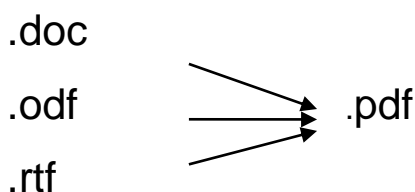


Figura 3 – Normalização de formatos de arquivo

Esta técnica é bastante associada à migração. A princípio parece ser a mesma, porém a normalização é feita com o intuito de reduzir a quantidade de formatos de arquivos distintos.

Para a garantia de sucesso desta técnica torna-se necessário a escolha correta do formato. Sempre que possível, deverão ser escolhidos formatos conhecidos pela comunidade de interesse baseados em normas internacionais abertas (HESLOP; DANIS; WILSON, 2002 *apud* FERREIRA, 2006).

Esta normalização pode ser definida de acordo com a política do repositório sendo que alguns só receberão o documento se estiver no formato estipulado pelo mesmo, ou poderão fazer a conversão no momento da entrega.

7. ESTUDO DE CASO DA BIBLIOTECA CENTRAL DA UNIVERSIDADE DE BRASÍLIA

Este estudo de caso tem por interesse responder ao seguinte propósito: identificar as metodologias aplicadas para a preservação dos documentos digitais na Biblioteca Central da Universidade de Brasília. Para tanto traremos primeiramente informações sobre a Universidade e a Biblioteca Central.

A Universidade de Brasília foi inaugurada em 21 de abril de 1962. Em janeiro de 2011, possui mais de 1900 professores e aproximadamente 2525 servidores. A UnB oferece 103 cursos de graduação, 64 de mestrado, 45 de doutorado e 73 especializações. Os órgãos complementares auxiliam os alunos no desenvolvimento de trabalhos de ensino, pesquisa e extensão. São eles: o Hospital Universitário de Brasília, a Biblioteca Central, a Fazenda Água Limpa, o Centro de Informática, a UnB TV, e a Editora UnB (UNIVERSIDADE DE BRASÍLIA).

A Biblioteca Central da Universidade de Brasília foi criada, em 1962, opondo-se à tradição de múltiplas bibliotecas dispersas nas várias unidades de ensino das universidades – um sistema oneroso que gerava duplicações desnecessárias de acervo e de processos técnicos e administrativos. Desde então, percorreu uma trajetória ímpar de mudanças, recuos e avanços (UNIVERSIDADE DE BRASÍLIA).

Detentora de um acervo formado por livros, teses, dissertações, periódicos, CD-ROM, mapas entre outros materiais de diferentes suportes e tipologias a Biblioteca Central da Universidade de Brasília possui um acervo com aproximadamente 350.000 títulos e mais de 1.450.000 exemplares. Possui

mais de 10.000 documentos digitais (teses, dissertações, monografias, narração de livros para deficientes visuais) disponibilizados em repositórios digitais com acesso livre via internet. Ainda fornece através de convênio com a CAPES acesso ao seu portal de periódicos contendo inúmeros títulos de periódicos de diversas áreas do conhecimento.

Tem como missão: “promover e garantir para a comunidade universitária o acesso à informação e o compartilhamento no âmbito do Sistemas de Bibliotecas da UnB contemplando o ensino, a pesquisa e a extensão” (UNIVERSIDADE DE BRASÍLIA)

7.1. Universo da pesquisa

O universo da pesquisa deste trabalho consiste nos documentos digitais existentes na BCE/UnB. Para se ter uma amostra consistente com o objetivo da pesquisa limitou-se o universo aos repositórios institucionais mantidos pela Biblioteca Central da Universidade de Brasília. Durante levantamento junto ao setor de Gerenciamento da Informação Digital (GID), responsável pelo gerenciamento de documentos digitais na BCE/UnB, foi relatada a existência de quatro repositórios. São eles:

- Repositório Institucional (RIUnB);

O Repositório Institucional da UnB, criado em setembro de 2008, é um conjunto de serviços oferecidos pela Biblioteca Central para a gestão e disseminação da produção científica e acadêmica da Universidade de Brasília. Todos os seus conteúdos estão disponíveis publicamente, e por estarem amplamente acessíveis proporcionam maior visibilidade e impacto da

produção científica da universidade. Este repositório utiliza-se do software DSpace para sua montagem e execução. Tendo sua página inicial representada pela figura 4.



Figura 4 – Página inicial do RIUnB
Fonte: <http://repositorio.bce.unb.br>

- Biblioteca Digital de Monografias (BDM);

A Biblioteca Digital de Monografias criada em setembro de 2009, é um serviço oferecido pela BCE/UnB, que têm como principal foco disseminar as monografias criadas por alunos de Especialização e Graduação da UnB. Este repositório faz uso do software DSpace para gerenciamento de seus arquivos, ao qual auxilia bastante no processo de preservação do objeto digital. Sua página inicial na *web* está representada na figura 5.



Figura 5 – Página Inicial da BDM

Fonte: <http://bdm.bce.unb.br>

- Biblioteca Digital de Teses e Dissertações (BDTD);

A Biblioteca Digital de Teses e Dissertações é um projeto instituído pelo IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia) que funciona na UnB desde 2006, têm por objetivo integrar os sistemas de informação de teses e dissertações existentes no país. Para isso o IBICT tornou cada Universidade responsável pela disponibilização de suas teses e dissertações defendidas. Podemos ver sua interface representada na figura 6.

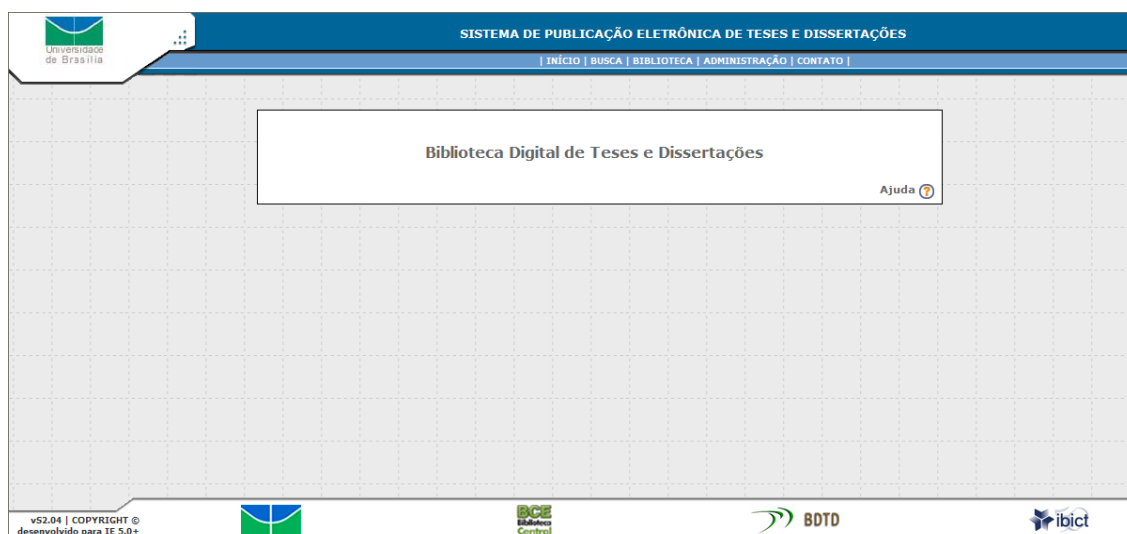


Figura 6 – Página Inicial da BDTD

Fonte: <http://bdtb.bce.unb.br>

- Biblioteca Digital e Sonora (BDS).

Desenvolvida pelo Programa de Apoio às Pessoas com Necessidades Especiais da UnB (PPNE) juntamente com a Biblioteca Central em 2008, a Biblioteca Digital e Sonora visa atender às demandas de informação dos deficientes visuais de toda a comunidade. Seu acervo físico é composto por obras em Braille, CDs gravados e o acervo virtual é composto por obras gravadas e digitalizadas disponíveis no site. O objetivo é proporcionar acessibilidade e promover aprendizagem de pessoas com deficiência visual. As equipes do PPNE e da BCE trabalham no desenvolvimento de alternativas metodológicas que visam à acessibilidade do conteúdo digital da BDS, e conta com a contribuição dos usuários para a inserção de documentos no repositório. Sua interface está representada na figura 7.



Figura 7 – Página Inicial da BDS

Fonte: <http://bds.bce.unb.br>

7.2. Coleta de dados

Os dados referentes à pesquisa foram coletados no dia 01/12/2010 por meio de pesquisa realizada com a responsável pelo setor de Gerenciamento da Informação Digital (GID) da BCE/UnB e dados estatísticos conseguidos junto à mesma.

Desde a criação do repositório mais antigo (BDTD) no ano de 2006, contabilizam-se um total de 12.038 documentos digitais distribuídos nos 4 repositórios e representados de acordo com a tabela 2.

Repositórios					
Ano	RIUnB	BDM	BDTD	BDS	Total
2006	0	0	0	0	0
2007	0	0	0	0	0
2008	240	0	0	0	240
2009	1689	624	0	0	2313
2010	3025	504	5179	777	9485
Total	4954	1128	5179	777	12038

Tabela 2 – Quantidade de Documentos Digitais na BCE/UnB

É importante salientar que não foi possível obter durante a coleta dos dados o número de arquivos inseridos anualmente dos repositórios BDTD e BDS. Foi possível obter apenas os dados gerais sobre os documentos existentes nestes repositórios.

Nestes 12.038 documentos são utilizados três formatos de arquivo distintos, 2 deles utilizados para apresentação de textos, PDF (*Portable Document Format*), HTML (*Hipertext Markup Language*), e 1 utilizado para arquivos de áudio, MP3(*MPEG 1/2 Audio Layer 3*), apresentados de acordo com a tabela 3.

Formatos		
PDF	HTML	MP3
11261	761	16

Tabela 3 – Formatos de arquivo

Os formatos de arquivos utilizados nos repositórios da BCE/UnB se enquadram ao que diz respeito ao conteúdo dos documentos. Por exemplo por ser um formato de arquivo criado para arquivos de conteúdo de áudio, o formato MP3 é utilizado na BDS para a gravação de áudio e a narração de livros, contos etc. Assim como o formato de arquivo PDF é utilizado para documentos textuais existentes nos repositórios, RIUnB, BDM, BDTD. A seguir, na tabela 4 a distribuição dos formatos de arquivo.

Repositórios					
	RIUnB	BDM	BDTD	BDS	TOTAL
PDF	4954	1128	5179	0	11261
HTML	0	0	0	761	761
MP3	0	0	0	16	16

Tabela 4 – Distribuição dos formatos de arquivo nos repositórios

No que se refere à utilização de metodologias para à preservação dos documentos digitais utilizadas pela BCE/UnB, pode-se constatar a utilização de metodologias em todos os repositórios. Como será demonstrado na tabela 5.

	Tecnologia	Refrescamento	Migração	Emulação	Normalização
RIUnB	Não	Não	Sim	Não	Sim
BDM	Não	Não	Sim	Não	Sim
BDTD	Não	Não	Sim	Não	Sim
BDS	Não	Não	Não	Não	Sim

Tabela 5 – Metodologias de preservação digital aplicadas nos repositórios

7.3. Análise dos dados coletados

A análise dos dados coletados junto a Biblioteca Central da Universidade de Brasília teve o intuito de responder à problemática do trabalho. Saber sobre a utilização das metodologias de preservação digital aplicadas aos documentos digitais existentes na Biblioteca Central da Universidade de Brasília.

Quanto aos documentos digitais existentes na BCE/UnB, constatou-se a presença de 12.038 documentos digitais e a utilização de três formatos de arquivo distintos. São eles: PDF, MP3, HTML.

O formato de arquivo PDF que foi encontrado em maior número, 11.261 documentos, é utilizado em sua totalidade em três repositórios: RIUnB, BDM, BDTD, como se pode constatar na tabela 4 apresentada no subitem anterior. Quanto aos outros formatos de arquivo MP3 e HTML são utilizados em sua totalidade nos documentos localizados na BDS, devido a características de tipologia do arquivo, por exemplo o formato de arquivo MP3 utilizado para arquivos de áudio apenas se enquadram para utilização nos documentos da BDS.

Isto nos leva a constatar, que a BCE/UnB utiliza-se adequadamente dos formatos de arquivo de que dispõe para melhor atender as necessidades informacionais de seus repositórios.

Quanto ao que se refere à preservação dos documentos digitais encontrados nos repositórios da BCE/UnB, observa-se nos dados coletados que a BCE/UnB utiliza-se de técnicas de preservação de documentos digitais. Esta afirmação torna-se verdadeira no momento em que podemos identificar, de acordo com a pesquisa bibliográfica realizada no presente trabalho, as características que nos levam a constatar a utilização das técnicas de **normalização e migração**. A **normalização** dos formatos de arquivo é realizada pela BCE/UnB com o intuito de reduzir o número de formatos de arquivo existentes em seus repositórios, representado anteriormente pela figura 3, facilitando assim a preservação dos mesmos para serem observados por usuários futuros.

A outra técnica de preservação digital utilizada pela BCE/UnB em seus documentos digitais é a **migração** de formato. Esta migração é realizada nos documentos digitais localizados nos repositórios RIUnB, BDM e BDTD. Dá-se, quando o arquivo enviado pelo usuário é recebido pela instituição e não se encontra no formato de arquivo utilizado pela BCE/UnB para disseminação do documento.

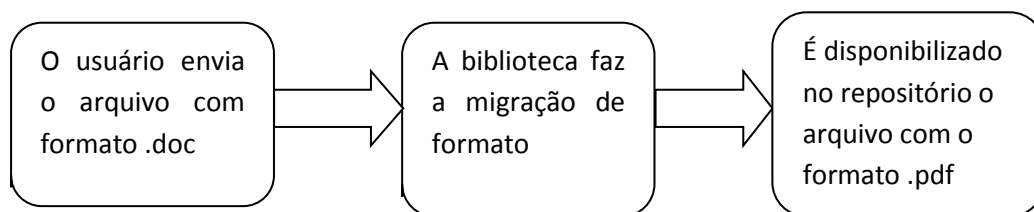


Figura 8 – Exemplo de migração de formato de arquivo

Como já alertado anteriormente no trabalho à migração e a normalização de documentos digitais são duas técnicas que estão associadas. A normalização destes documentos é feita através do processo de migração dos formatos de arquivo.

8. CONCLUSÃO

Serão feitas aqui algumas considerações acerca do que foi realizado neste trabalho e algumas reflexões acerca do trabalho que vem sendo realizado na BCE/UnB.

A respeito dos resultados constatados na pesquisa realizada junto à biblioteca para responder ao objetivo geral que norteia este trabalho, pode-se afirmar que a BCE/UnB utiliza-se de parte das metodologias para a preservação de seus documentos digitais. Esta afirmação baseia-se nos dados coletados na pesquisa e na revisão de literatura realizada anteriormente.

Algumas das técnicas relatadas neste trabalho como a preservação de tecnologia, a emulação e o refrescamento, não teriam aplicação no acervo de documentos digitais da biblioteca. Levando-se em conta que as técnicas de preservação de tecnologia, emulação e refrescamento devem se aplicar a acervos já com alguma troca de tecnologia ou danos, sejam eles de *hardware* ou *software*. Apesar de cada dia as trocas de tecnologia estejam cada vez mais rápidas, ainda não se tornam necessárias à aplicação de tais técnicas ao caso dos documentos digitais que se encontram na BCE/UnB. Leva-se em consideração neste caso que o repositório com mais tempo de atividade tem menos de 6 anos, e os *softwares* e *hardwares* aplicados para visualização destes documentos ainda estão sendo utilizados.

É muito importante abordar o fato de que a Biblioteca Central da Universidade de Brasília busque atender todos os seus usuários, sejam eles: graduandos, mestrandos, doutorandos, professores, pesquisadores e usuários com deficiência física. Os diversos repositórios digitais disponibilizados nesta

biblioteca vêm a contribuir com o melhor acesso a informação por parte de todos os seus usuários.

Observa-se também de acordo com a observação da tabela 2 (p. 28), o aumento da inserção de documentos nos repositórios disponibilizados pela Biblioteca Central por parte dos membros da UnB. O que vem a concordar com o que já se foi falado anteriormente no trabalho, o aumento da produção de documentos em formato digital e a maior utilização desses documentos para a disseminação da informação.

No que tange os objetivos deste trabalho, analisaremos aqui o que foi feito para responder a cada um deles.

O primeiro objetivo específico que se propunha a caracterizar os conceitos de documento digital e preservação digital, foi realizada pesquisa bibliográfica acerca dos temas em questão para caracterizar cada um deles. Obteve-se como resultado desta caracterização um conceito de documento digital “como todo e qualquer objeto de informação que possa ser representado através de uma sequência de dígitos binários” Thibodeau (2002 *apud* FERREIRA, 2006). Para o conceito de preservação digital, baseou-se em definições de Cunha e Cavalcanti (2008, p. 290) e um relatório (*Modern Information Retrieval Glossary*) citado por Boeres (2004, p. 26). Pode-se inferir através da análise destas duas definições que o grande propósito da preservação digital é conseguir manter os documentos digitais acessíveis quando solicitados em qualquer época.

Em relação ao segundo objetivo específico se propõe descrever as metodologias aplicáveis à preservação de documentos digitais, buscou-se revisão de literatura acerca do tema. Foram encontradas e caracterizadas cinco

metodologias aplicadas à preservação de documentos digitais. São elas: preservação de tecnologia, refrescamento, migração, emulação e normalização.

O objetivo específico terceiro teve o intuito de analisar quais as metodologias de preservação digital utilizadas pela Biblioteca Central da Universidade de Brasília. Este objetivo só pode ser respondido com o estudo de caso da BCE/UnB e com o apoio da revisão de literatura realizada sobre o tema. Dentre as metodologias aplicadas à preservação digital vistas anteriormente, foram encontradas duas delas sendo aplicadas na BCE/UnB sendo elas a migração e a normalização.

Com o auxílio de estudos feitos acerca dos conceitos de documento digital, preservação digital, assim como as metodologias aplicadas a preservação digital e o estudo de caso da BCE/UnB, foi possível responder ao objetivo principal que norteia este trabalho que é: identificar as metodologias aplicadas para a preservação dos documentos digitais na Biblioteca Central da Universidade de Brasília. Foi relatada a utilização de duas metodologias aplicadas à preservação de documentos digitais nos repositórios da BCE/UnB. Estas técnicas são a migração e a normalização de formatos de arquivos.

Recomendam-se novos estudos na área da preservação de documentos digitais em centros de informação. Podendo assim aumentar a literatura sobre o tema e auxiliar também novos trabalhos nesta área da Ciência da Informação.

9. Referências

AMORIM, Eliane Dutra; LOPES, Carlos Eduardo Rodrigues; DO VALLE JUNIOR, Eduardo Alves. **Introdução à preservação de acervos digitais**. Belo Horizonte: Secretaria de Estado de Cultura/ Arquivo Público Mineiro, 2005, p. 42.

BAPTISTA, Dulce Maria. Notas de aula da disciplina: Historia do Livro e das Bibliotecas. FCI/UnB, 2009.

BODÊ, Ernesto Carlos. **Preservação de Documentos Digitais: o Papel dos Formatos de Arquivo**. 2008. 153 f. (Dissertação de Mestrado em Ciência da Informação) – Departamento de Ciência da Informação, Universidade de Brasília, Brasília.

BOERES, Sonia Araújo de Assis. **Política de preservação da informação digital em bibliotecas universitárias brasileiras**. 2004. 180 f. Dissertação (Mestrado em Ciência da Informação e Documentação) - Universidade de Brasília, Brasília, 2004.

CONWAY, Paul. **Preservação no Universo Digital**. Rio de Janeiro: Arquivo Nacional, 2001. 32 p.

CUNHA, Murilo Bastos da. CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de biblioteconomia e arquivologia**. Brasília: Briquet de Lemos, 2008.

FERREIRA, Miguel. **Introdução à Preservação Digital: conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. 85 p.

LOPEZ, Andre Porto Ancona. **Principios arquivísticos e documentos digitais**. Arquivo Rio Claro, Rio Claro, n 2, p. 70-85, 2004

MÁRDERO ARELLANO, Miguel Ángel. **Critérios para a preservação digital da informação científica**. 2008. 354 f. (Doutorado em Ciência da Informação) – Departamento de Ciência da Informação, Universidade de Brasília, Brasília.

THIBODEAU, K.. **Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years**. The State of Digital Preservation: An International Perspective Washington D.C., 2002.

UNIVERSIDADE DE BRASÍLIA. Site da Universidade de Brasília .Disponível em <<http://www.unb.br/>> Acesso em: 10/12/2010.

UNIVERSIDADE DE BRASÍLIA. Site da Biblioteca Central. Disponível em: <<http://bce.unb.br/>> Acesso em: 10/12/2010.

UNIVERSITY OF LEEDS. Survey and assessment of sources of information on file formats and software documentation. The representation and rendering project. Reino Unido, [s.d.] 48 p.