

Clarissa Costa e Lima

PRESERVAÇÃO DIGITAL:
A EXPERIÊNCIA DA PESQUISA GUIGNARD

Dissertação apresentada ao Programa de Pós-Graduação em Artes da Escola de Belas Artes da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Artes.

Área de concentração: Arte e Tecnologia da Imagem.

Orientador: Prof. Arnaldo de Albuquerque Araújo
DCC- ICEx - UFMG

Belo Horizonte

Escola de Belas Artes

2007



Universidade Federal de Minas Gerais
Escola de Belas Artes
Mestrado em Artes Visuais: Arte e Tecnologia da Imagem

Dissertação intitulada “Preservação Digital: A experiência da Pesquisa Guignard”, de autoria da mestrandona Clarissa Costa e Lima, aprovada pela banca examinadora constituída pelos seguintes:

Prof. Dr. Arnaldo de Albuquerque Araújo – DCC/ICEEx/UFMG – Orientador

Dra. Claudina Maria Dutra Moresi - CECOR/EBA/UFMG

Prof. Dr. Marcello P. Bax - ECI/UFMG

Profa. Dra. Marília Andrés Ribeiro - FAFICH/UFMG

Belo Horizonte, 04 de julho de 2007

Ao meu Castelo Forte e
à Pesquisa Guignard

AGRADECIMENTOS

A Deus.

Ao Professor Dr. Arnaldo de Albuquerque Araújo,
Pelas orientações e exortações, não apenas durante o período do mestrado, mas em todo este tempo de Pesquisa Guignard.

À Claudina,
Pelo estímulo, importantes sugestões e oportunidade de aplicar a informática a serviço de um universo encantador de arte e cultura.

Ao Cris,
Pela colaboração.

Ao SERPRO pelo incentivo.

À minha amada mãe, Norma,
Pelo esforço e dedicação incondicionais, sempre.

Aos também amados Geraldo, Karine e Ludmila.

À querida Suzi, pelo companheirismo.

Aos amigos, Quequel, Raquel, Elba e Camillo,
Pelo apoio e atenção, muitas vezes pontos de equilíbrio.

Finalmente, agradeço a todos os colegas da Pesquisa Guignard,
Pelo respeito, carinho e cordialidade, sempre acolhedores. Minha admiração por todos.

RESUMO

Preservação digital significa manter uma coleção digital de forma usável e interpretável por um longo período. Isso implica na manutenção da integridade, do contexto tecnológico, do histórico de procedência, referências e usabilidade de cada objeto digital. Diante da obsolescência tecnológica, a maior ameaça às coleções digitais, a preservação digital se faz desafiadora. Através deste estudo de caso, verificam-se os principais aspectos a serem considerados na elaboração de uma política de preservação digital e identificam-se os requisitos e necessidades de sistemas de informação voltados para este domínio de aplicação. Dessa forma, é apresentada a aplicação da preservação digital na Pesquisa Guignard, responsável pelo estudo da vida e obra de Alberto da Veiga Guignard, artista Modernista, e que possui boa parte de seu patrimônio no formato digital, inclusive se originando assim. No entanto, não é possível se determinar um plano ideal para manutenção de todo e qualquer patrimônio digital. Todavia, a preservação digital é uma política de controle de riscos, que objetiva a permanência de uma herança cultural, histórica e científica.

Palavras-chave: Preservação digital, longevidade, metadados, sistemas de informação.

ABSTRACT

Keeping a digital collection for a long period in a usable and interpretable way is called digital preservation. It implicates the maintenance of the integrity, of the technological context, and of the origin description, references and usability of every digital object. Faced to technological obsolescence, the greatest threaten to digital collections, digital preservation is a challenge. The essential aspects considered in a digital preservation policy elaboration are examined in this case study. Besides, it identifies the requirements and needs for information systems designed to this application domain. This work presents the digital preservation applied to the Guignard Research, a study about the life and the work of Alberto da Veiga Guignard, Modernist artist. A great part of the Guignard Research's collection is held in digital format. Nevertheless, it is impossible to set an ideal, definite and general purpose maintenance plan for any digital collection. However, digital preservation is a risk control policy which aims to a lasting cultural, historical and scientific legacy.

Keywords: Digital Preservation, longevity, metadata, information systems.

LISTA DE ILUSTRAÇÕES

FIGURA 1 - Foto das camadas de um CD-ROM.....	27
FIGURA 2 - Modelo ambiental de um OAIS.....	41
FIGURA 3 - OAIS – Modelo com entidades funcionais.....	43
FIGURA 4 - Processo de desenvolvimento de sistemas.	67
FIGURA 5 - Ciclo de vida para desenvolvimento de interfaces em estrela.....	68
FIGURA 6 - Atividade de fluxo de requisitos.....	69
FIGURA 7 - Tela do File Maker Pro, sistema de banco de dados utilizado inicialmente pela Pesquisa Guignard para documentação bibliográfica.....	77
FIGURA 8 - Tela do Sistema de Obtenção de Dados – Bibliografia.....	78
FIGURA 9 - Diagrama esquemático da representação da informação sobre Obras.	79
FIGURA 10 - Diagrama de caso de uso para o sistema de acesso ao acervo da Pesquisa.....	82
FIGURA 11 - Leiaute esquemático.	83
FIGURA 12 - Tela Entrevistas e Depoimentos.	84
FIGURA 13 - Tela Busca por Resultados de Estudos de Obras.	84
FIGURA 14 - Diagrama de estados para a Tela Busca por Resultados de Estudos de Obras. 84	
FIGURA 15 - Diagrama de funcionamento do sistema de recuperação de imagens com base em imagem exemplo.....	86
FIGURA 16 - Tela Busca por Imagem Exemplo aos Trabalhos Artísticos Estudados pela Pesquisa.	87
FIGURA 17 - Diagrama de estado para a Tela Busca por Imagem Exemplo aos Trabalhos Artísticos Estudados pela Pesquisa.	87
GRÁFICO 1 - Condições de temperatura e umidade e risco de hidrólise.....	23
QUADRO 1 - Características-chave de armazenamento de acesso e armazenamento arquivístico	24
QUADRO 2 - Parâmetros recomendáveis para armazenamento em diferentes fontes	28
QUADRO 3 - Paralelo entre os formatos JPEG, GIF, PNG e PDF	51
QUADRO 4 - Recomendações da Universidade de Columbia	53
QUADRO 5 - Comparação de formatos de arquivos digitais de áudio	56
QUADRO 6 - Tamanhos de arquivos para uma hora de áudio	57
QUADRO 7 - Funções do produto	81
QUADRO 8 - Descrição de atores	82

LISTA DE SIGLAS

ASCII	Padrão de Código Americano para troca de informações, do inglês <i>American Standard Code for Information Interchange</i> .
Bit	Dígito Binário, do inglês <i>Binary digit</i> .
CAMiLEON	<i>Creative Archiving at Michigan & Leeds: Emulating the Old on the New</i> .
CBIR	Recuperação de Imagens baseada no Conteúdo, do inglês <i>Content-Based Image Retrieval</i> .
CCSDS	<i>Consultative Committee for Space Data System</i> .
CD	Disco Compacto, do inglês <i>Compact Disc</i> .
CDA	Disco Compacto de Áudio, do inglês <i>Compact Disc Audio</i> .
CD-R	Disco Compacto Gravável, do inglês <i>Compact Disc - Recordable ou Writable</i> .
CD-RAM	Dico Compacto de Memória Acessada Randomicamente, do inglês <i>Compact Disc Random Access Memory</i> .
CD-ROM	Disco Compacto de Memória apenas para leitura, do inglês <i>Compact Disc Read Only Memory</i> .
CD-RW	Disco Compacto Regravável, do inglês <i>Compact Disk – Rewritable</i> .
CDWA	Categorias para Descrição de Trabalhos Artísticos, do inglês <i>Categories for the Description of Works of Art</i> .
CECOR	Centro de Conservação e Restauração de Bens Culturais Móveis.
CEDARS	CURL <i>Exemplars in Digital Archives</i> .
CONARQ	Conselho Nacional de Arquivos.
CT&IT	Coordenadoria de Transferência e Inovação Tecnológica.
CTDE	Câmera Técnica de Documentos Eletrônicos
DAT	Fita de Áudio Digital, do inglês <i>Digital Audio Tape</i> .
DDS	Armazenamento de Dados Digitais, do inglês <i>Digital Data Storage tape</i> .

DPI	Pontos por Polegada, do inglês <i>Dot per inch.</i>
DSEP	Sistema de arquivamento virtual para Publicações Eletrônicas, do inglês <i>Deposit Systems for Electronic Publications.</i>
DVD	Disco Digital Versátil, do inglês <i>Digital Versatile Disc.</i>
FEDORA	<i>Flexible Extensible Digital Object and Repository Architecture.</i>
FLAC	Codec de Áudio Livre Sem Perdas, do inglês <i>Free Lossless Audio Codec.</i>
GIF	Formato para Intercâmbio de Gráficos, do inglês <i>Graphics Interchange Format.</i>
HTML	Linguagem para marcação de texto, do inglês <i>Hyper Text Markup Language.</i>
ICR	Reconhecimento inteligente de caracteres, do inglês <i>Intelligent Character Recognition.</i>
INPE	Instituto Nacional de Pesquisas Espaciais.
ISO	Organização Internacional para Padronização, do inglês <i>International Organization for Standardization.</i>
JPEG	<i>Joint Picture Expert Group.</i>
KHz	Kilohertz.
MARC	Catalogação legível por máquina, do inglês <i>Machine Readable Cataloging.</i>
MP3	<i>MPEG- ½ Audio Layer 3.</i>
NASA	<i>National Aeronautics and Space Administration.</i>
NEDLIB	Rede Européia de Arquivamento Virtual para Bibliotecas, do inglês <i>Network European Deposit Library.</i>
NISO	Organização Nacional de Padrões Informacionais, do inglês <i>National information Standards Organization.</i>
OAIS	Sistema de Informação Aberto para Arquivos, do inglês <i>Open Archival Information System.</i>
OCLC/RLG	<i>Online Computer Library Center/ Research Libraries Group.</i>

OCR	Reconhecimento óptico de caracteres, do inglês <i>Optical Character Recognition</i> .
PANDORA	<i>Preserving and Accessing Networked Documentary Resources of Australia.</i>
PDF	Formato para Documentos Portáveis, do inglês <i>Portable Document Format</i> .
pH	Potencial Hidrogênio iônico.
PHP	<i>Hipertext PreProcessor.</i>
PNG	Gráficos Portáveis de Rede, do inglês <i>Portable Network Graphics</i> .
PRAXIS	PRocesso para Aplicativos eXtensíveis InterativoS.
RDF	Sistema de Descrição de Recursos, do inglês <i>Resource Description Framework</i> .
RPM	Rotações por Minuto.
S-VHS	<i>Super Video Home System.</i>
TIFF	<i>Tagged Image/Interchange File Format.</i>
UML	Linguagem de Modelagem Unificada, do inglês <i>Unified Modeling Language</i> .
VHS	<i>Video Home System.</i>
WAV	<i>WAVeform audio file format.</i>
Web	Teia do tamanho do mundo, do inglês <i>Word Wide Web ou WWW</i> .
WMA	<i>Windows Medi Audio.</i>
XHTML	Linguagem de Marcação Extensível para Hipertexto, do inglês <i>eXtensible Hypertext Markup Language..</i>
XML	Linguagem de Marcação Extensível, do inglês <i>eXtensible Markup Language</i> .

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Patrimônio digital.....	15
1.2	Objetivos.....	16
1.3	Estrutura da dissertação.....	17
2	ASPECTOS FUNDAMENTAIS DA PRESERVAÇÃO DIGITAL.....	19
2.1	Longevidade digital	19
2.1.1	Mídias magnéticas	20
2.1.2	Discos Ópticos.....	26
2.2	Seleção de objetos digitais a serem mantidos.....	30
2.2.1	Direitos Autorais	33
2.3	Autenticidade no acervo digital.....	34
2.4	Estratégias para preservação digital	36
2.4.1	Refrescamento	36
2.4.2	Migração.....	37
2.4.3	Emulação	38
2.4.4	Arqueologia digital	39
2.5	Padrão para Preservação ISO 14721 – <i>Open Archival Information System (OAIS)</i>.....	40
2.6	Iniciativas institucionais.....	43
2.6.1	<i>Networked European Deposit Library - NEDLIB</i>	44
2.6.2	e-ARQ.....	45
2.6.3	Projeto Portinari.....	46
3	FORMAÇÃO DE ACERVO DIGITAL MULTIMÍDIA.....	48
3.1	Conversão de imagens.....	49
3.2	Conversão de áudio	54
4	METADADOS	58
4.1	Metadados para imagens	60
4.2	Metadados para arquivos sonoros	62
5	DESENVOLVIMENTO DE SISTEMAS ORIENTADOS À PRESERVAÇÃO	64
5.1	Gestão documental	65
5.2	Processo de desenvolvimento de sistemas.....	66
5.3	Engenharia de requisitos	68
5.4	Engenharia da Web	69
6	PRESERVAÇÃO DIGITAL NA PESQUISA GUIGNARD	72
6.1	Seleção para disponibilização de acesso	72
6.2	Formação do acervo digital	73
6.3	Cuidados com os meios digitais	75
6.4	Sistemas	76
6.4.1	Sistema de Obtenção de Dados	76
6.4.2	Sistema de busca.....	80
6.4.3	Ferramenta de busca pelo conteúdo.....	85
6.4.4	Considerações finais	87
6.5	Propostas futuras	88
7	CONCLUSÃO.....	90
	REFERÊNCIAS	93

1 INTRODUÇÃO

A Pesquisa Guignard teve sua origem no CECOR – Centro de Conservação e Restauração de Bens Culturais Móveis - no início do ano 2000. Ela propõe o estudo sistemático do artista Alberto da Veiga Guignard, através de uma metodologia própria.

O desenhista e pintor Alberto da Veiga Guignard nasceu em Nova Friburgo, Rio de Janeiro, em 1896 e faleceu em Belo Horizonte, em 1962. Guignard é reconhecido hoje como um dos maiores expoentes da arte Moderna brasileira, formador de escola em Belo Horizonte e Rio de Janeiro, bem como de várias gerações de artistas. Estudou na Europa (1907-1929), ingressou na Real Academia de Belas Artes de Munique em 1917, onde estudou com o pintor Hermann Groeber e com o artista gráfico e ilustrador Adolpho Hengeler, pertencentes ao grupo “*Sezession*”. Participou de vários salões nacionais e internacionais; conheceu Picasso, Matisse, dentre outros pintores famosos, recebendo importantes prêmios ao longo de sua carreira. Ao retornar ao Brasil, residiu no Rio de Janeiro e em Itatiaia. Em 1944, transferiu-se para Belo Horizonte, a convite do Prefeito Juscelino Kubitschek de Oliveira, para dirigir um curso livre de desenho e pintura, atual Escola Guignard da Universidade Estadual de Minas Gerais (UEMG). Ele iniciou o Modernismo em Belo Horizonte, permanecendo em Minas até a sua morte (VIEIRA, 2000).

A produção do artista compreende desenhos, pinturas sobre madeira (portas de armário, violão, compensados, etc.), pinturas sobre tela, pinturas murais e forros. Os temas pintados foram paisagens, retratos, auto-retratos, flores, naturezas-mortas, temas religiosos e outros (ZANINI, 1982). Guignard, considerado um dos maiores pintores brasileiros, ainda não foi devidamente estudado, sobretudo no que diz respeito à caracterização da sua técnica construtiva.

O estudo sistemático usando metodologia científica de análise, com identificação de materiais, técnicas e pineladas é um procedimento já difundido em vários países, mas pouco explorado no Brasil. Estes estudos possibilitam o resgate dos métodos empregados na produção de obras e permitem caracterizar a maneira de pintar e desenhar do artista.

Dante desse quadro e da importância da atuação do mestre Guignard em Minas, a Pesquisa Guignard, constitui amplo programa de pesquisa e documentação da vida e da obra desse artista, relacionando aspectos históricos, estéticos e caracterização dos materiais e técnicas, através de metodologias técnico-científicas aplicadas às artes plásticas.

Os procedimentos metodológicos envolvem ações multidisciplinares; desta forma, pela sua abrangência, a Pesquisa Guignard é um projeto onde as seguintes unidades da Universidade Federal de Minas Gerais (UFMG) estão envolvidas: Centro de Conservação e Restauração de Bens Culturais Móveis (Cecor) e Departamento de Fotografia, Teatro e Cinema (FTC), Escola de Belas Artes (EBA); Núcleo de Processamento Digital de Imagens (NPDI), Departamento de Ciência da Computação (DCC), Instituto de Ciências Exatas (ICEEx); Programa de História Oral (PHO), Centro de Estudos Mineiros (CEM), Faculdade de Filosofia e Ciências Humanas (FAFICH); Departamento de Ciências Aplicadas à Educação (DECAE), Faculdade de Educação (FaE).

Essa parceria institucional possibilita a estruturação da Pesquisa em seis vertentes distintas e complementares, correspondendo às fases:

- Bibliografia: Fase que envolve o levantamento e análise de material bibliográfico por assunto e ilustrações das obras. Serve de fonte de consulta, confirmação e comparação de dados, que são coletados pelo método de História Oral, os quais têm sido comparados com os resultados obtidos pelas análises químicas dos procedimentos técnicos e construtivos das obras, possibilitando ampliar e complementar o conhecimento sobre o artista e sua obra.
- História Oral: Metodologia de pesquisa que procura resgatar e registrar memórias, opiniões, posturas, sentimentos, visões dos sujeitos no seu percurso de vida, acerca de eventos, conjunturas, instituições, modos de vida ou outros aspectos da história contemporânea. Na Pesquisa Guignard, é utilizada a história oral temática, cujo tema estabelecido é a Vida e Obra de Guignard em Minas Gerais.
- Estudo da Obra: Sua metodologia envolve exames e técnicos estruturais, experimento de cores e identificação da técnica e dos materiais por métodos físico-químicos de análise. Este estudo visa caracterizar a obra e ampliar o conhecimento sobre o processo criativo do artista.

- Documentação Fotográfica e Digitalização: A documentação fotográfica constitui importante recurso de apoio à pesquisa, facilitando o estudo das características materiais e estéticas, presta-se também à produção de material informativo para divulgação.
- Preservação Digital: Fase na qual este trabalho se encaixa. Corresponde ao programa para preservação digital dos resultados de estudos da Pesquisa através de instruções de procedimentos para acondicionamento e conservação do material digital e desenvolvimento de sistemas de informação para armazenamento e divulgação de dados. Essa fase dá suporte às demais vertentes da Pesquisa, possibilitando a organização e distribuição da informação e orientação sobre a infra-estrutura do acervo digital.
- Conservação: Acondicionamento e conservação de todo material textual e audiovisual produzido e adquirido na Pesquisa no formato analógico.

Reunindo especialistas de diferentes áreas do conhecimento, o projeto funciona como um núcleo cooperativo, capaz de promover o intercâmbio de conhecimento entre os parceiros e a formação de competência científico-tecnológica nos campos da preservação de bens móveis em meio digital, conservação-restauração, história da arte, artes plásticas, história oral, fotografia, educação, informação e informática.

A metodologia de análise da obra de Guignard proposta pela Pesquisa pretende servir de parâmetro para o estudo de outros artistas brasileiros, criando uma referência técnica para o conhecimento e controle de acervos e coleções de arte no Brasil. Considerando os atuais desafios enfrentados pela conservação, especificamente as perdas decorrentes do desconhecimento específico de materiais e técnicas empregadas em obras de arte modernas e contemporâneas (HUMMELEN; SILLÉ, 1999), a Pesquisa tem gerado novos conhecimentos sobre os processos de degradação de obras de arte, informações estas cruciais para a formulação de propostas e procedimentos de preservação. Por conseguinte, o CECOR tem aprimorado sua metodologia de caracterização dos objetos artístico-culturais, melhorando a qualidade das suas atividades de ensino e pesquisa, com ênfase na área de informatização de dados científicos, ampliando uma nova área – a computação aplicada à conservação.

1.1 Patrimônio digital

A produção artística de Guignard está dispersa principalmente em coleções particulares. O estudo técnico científico é dificultado por essa razão. Em contra partida, a tecnologia digital vem auxiliando o estudo da obra e possibilitando a divulgação das obras do artista e dos estudos da Pesquisa. A maior parte do patrimônio da Pesquisa Guignard se encontra em forma digital, inclusive com uma parte se originando assim.

Do mesmo modo, alguns recursos existem cada vez mais em formas digitais, e progressivamente apenas no formato digital, e neste caso, quando um recurso nasce na forma digital não há outro formato além do original em digital. Alguns destes recursos têm valor passado e significativo, abrangendo qualquer área do conhecimento e expressão humana, constituindo um patrimônio que deve ser protegido e preservado para as gerações correntes e futuras. O patrimônio digital é constituído, dentre outras possibilidades, por arquivos de texto, bases de dados, programas de computador, páginas *Web*¹, imagens estáticas e em movimento, gráficos, áudio, abrangendo uma vasta e crescente variedade de formatos (NATIONAL LIBRARY OF AUTRALIA, 2003).

As tecnologias utilizadas na criação e no uso dos recursos digitais têm algumas vantagens que explicam o seu extraordinário avanço em várias partes do mundo. Dentre elas, a possibilidade de edição de imagens constitui um recurso precioso para o estudo da obra de arte, oferecendo flexibilidade de experimentos e análises documentais, impensáveis de serem aplicados em originais não digitais. A agilidade e a facilidade para obtenção de acesso a documentos é outro ponto favorável, além da produção, transmissão e armazenagem a baixos custos, com ganho no processo de criação e troca de informação arquivística.

Porém, existem preocupações na guarda de um patrimônio na forma digital. A mídia utilizada para armazenamento é instável e a tecnologia necessária para o acesso é rapidamente superada por novas tecnologias. As tecnologias perdem seu suporte e, desta forma, o acesso ao patrimônio digital pode ser perdido. Exemplos destes problemas são a perda de dados da exploração de Marte pela *Viking* em 1976, o caso das observações via satélite do Brasil nos

¹ *Web - World Wide Web*, em português, teia do tamanho do mundo.

anos 1970, importantes para o estabelecimento da linha de mudanças climáticas da bacia amazônica, perdidas em fitas obsoletas; e o censo americano em 1960, onde os computadores e programas para acesso aos dados se tornaram obsoletos e a tecnologia atual não é capaz de ler os registros informacionais (WATERS; GARRETT, 1996).

O propósito de se preservar um patrimônio digital consiste em assegurar a acessibilidade, autenticidade e integridade do acervo, mantendo a memória para a sociedade. Instituições como a *United Nations Educational, Scientific and Cultural Organization* (UNESCO)², que tem como objetivo contribuir para a paz e segurança no mundo, mediante a educação, ciência, cultura e comunicação já levantou sua preocupação a cerca da perda da memória digital e tomou medidas para impedi-la com a resolução sobre preservação digital (CONSELHO NACIONAL DE ARQUIVOS, 2005) e o guia para preservação do patrimônio digital (NLA, 2003). Essa inquietação é também salientada por Howell (2000) ao lembrar que a humanidade tem dois mil anos de experiência na preservação de papel manuscrito e duzentos anos na preservação de papel feito por máquinas, mas não tem experiência na preservação de documentos digitais. O conhecimento sobre a preservação digital é recente, muitas pesquisas estão sendo realizadas e tendem a evoluir, porém, já existem alguns resultados divulgados, como os ilustrados neste trabalho.

Diante desse quadro, a Pesquisa Guignard, que tem quase toda parte de seu valor ou seu produto, agregado de forma digital, ou seja, em informação, se vê com a necessidade de lidar com a preservação digital. Dessa forma, é fundamental a discussão deste tema.

1.2 Objetivos

Dada a importância do mestre Guignard para a arte brasileira, o desconhecimento da totalidade do seu acervo, a falta de um maior aprofundamento dos materiais e técnicas empregados pelo artista, tem-se a relevância da formação de um banco de dados, com acesso via *Internet* para democratização de resultados de estudos científicos e a aplicação da preservação digital.

² Organização das Nações Unidas para a Educação, Ciência e Cultura.

Neste trabalho, há um estudo de caso do emprego da preservação digital das informações textuais e audiovisuais geradas na Pesquisa Guignard, que poderá servir a outras instituições que mantêm a custódia de acervos. O programa de preservação trata sobre:

- A formação do acervo digital, sua guarda, organização, gestão e divulgação;
- A conservação do conteúdo digital;
- O respeito aos direitos autorais;
- O uso de sistemas de informação para gestão documental e para divulgação do trabalho da Pesquisa.

Dessa forma, este trabalho visa contribuir para difusão da obra de Guignard e resultados de estudos de sua obra ao colaborar para a compreensão sobre como preservar o patrimônio artístico, por meio de recursos tecnológicos.

1.3 Estrutura da dissertação

Este capítulo apresentou a Pesquisa Guignard, ressaltando a importância de seus estudos nos cenários das artes e da cultura. O patrimônio digital foi caracterizado, mostrou-se sua abrangência, fragilidade e fontes de origem. Foram apresentados os problemas ligados à guarda de um acervo no formato digital. A partir disso, deu-se a definição do problema motivador deste trabalho: A aplicação da preservação digital na Pesquisa Guignard. Como consequência, os objetivos do presente trabalho foram apresentados, dando ênfase à explanação de um estudo de caso. Desta forma, foram estruturados os seguintes capítulos:

- O Capítulo 2 expõe os primeiros aspectos a serem considerados dentro de uma política de preservação digital, englobando: os cuidados de guarda e manuseio a serem tomados com os suportes de informação digital, a fim de aumentar sua longevidade; os critérios para seleção do acervo a ser mantido, diante de uma grande massa de arquivos digitais; a importância da autenticidade numa coleção; a apresentação das principais e mais difundidas estratégias de preservação digital; a figuração do padrão

*Open Archival Information System*³ (OAIS); o relato de algumas iniciativas institucionais.

- O Capítulo 3 discorre sobre a formação do acervo digital, em especial, com imagens e áudio.
- O Capítulo 4 traz a definição de metadados e a importância de seu emprego na preservação do patrimônio digital.
- O Capítulo 5 explana sobre o desenvolvimento de sistemas de informação, ferramenta fundamental na promoção da gestão e disponibilização do acervo.
- O Capítulo 6 traz o estudo de caso da aplicação da preservação digital na Pesquisa Guignard.
- O Capítulo 7 expõe as conclusões do trabalho.

³ *Open Archival Information System*, em português, Sistema de Informação Aberto para Arquivos.

2 ASPECTOS FUNDAMENTAIS DA PRESERVAÇÃO DIGITAL

A preservação por longos períodos dos registros intelectuais e culturais da sociedade tem ocupado bibliotecários, arquivistas e curadores de museus por séculos. Mas este assunto tem se revigorado com a proliferação das tecnologias para criação e armazenamento do formato digital e se tornado um problema tecnológico, envolvendo maior diversidade de profissionais.

A preservação digital vem então se tornando um imperativo pela agilidade de expansão da informação, e um desafio, uma vez que o processo de preservar a informação por longo período e mantê-la acessível exige um grande e contínuo esforço. Esforço esse, justificado pela fragilidade da mídia, pela volatilidade tecnológica, pela necessidade de provas documentais para se interpretar o passado e por perdas já sofridas pela humanidade.

As instituições, portanto, devem se preparar adotando uma política de preservação adequada ao patrimônio digital, que possui suas particularidades. O item a ser mantido na preservação digital é a informação contida num suporte, não o objeto em si, como ocorre na preservação de patrimônio analógico (SITTS, 2000).

Através deste capítulo, pode-se observar que um projeto de preservação de conteúdo informacional deve atentar principalmente ao manuseio ideal para conservação das mídias que armazenam os conteúdos digitais. Deve expor critérios para seleção de materiais a serem mantidos como parte do patrimônio digital, além de explicitar qual a melhor estratégia para se preservar a informação, apesar da volatilidade tecnológica. Estas instituições podem também se nortear ou seguir padrões já estabelecidos para preservação do bem digital.

2.1 Longevidade digital

Embora o advento dos meios de armazenamento eletrônico seja bastante novo, uma quantidade substancial de informações armazenadas eletronicamente tem se deteriorado e

desaparecido. Existe um dilema: a capacidade de registrar informações aumentou exponencialmente ao longo do tempo, enquanto que a longevidade dos meios utilizados para armazená-la decresceu de modo equivalente (CONWAY, 97), (SITTS, 2000).

Dentre as mídias mais utilizadas para guarda da informação com fins de preservação estão a fita magnética, os discos compactos (CDs)⁴ e os discos digitais versáteis (DVDs)⁵. A fita magnética possui uma grande capacidade de armazenamento a baixo custo. Já os discos ópticos são mais estáveis do que a fita magnética. O seu surgimento promoveu a popularização da gravação e armazenamento de todo tipo de conteúdo digital, por gerarem um acesso mais ágil do que as fitas magnéticas. Entretanto, nenhum dos dois, nem os discos ópticos, nem as fitas magnéticas, são estáveis como o microfilme ou o papel (BOGART, 2001).

Não obstante, medidas podem ser tomadas para que as mídias sejam conservadas por um período maior. Devem-se observar as condições de temperatura e umidade, pois essas podem afetar acentuadamente a vida útil desses meios. Ambientes especiais de condicionamento podem vir a ser necessários. É, também importante, que se faça a transcrição periódica dos meios antigos para novos, não somente porque o suporte é instável, mas também porque a tecnologia de gravação se tornará obsoleta (BOGART, 2001).

2.1.1 Mídias magnéticas

Os meios magnéticos são suportes para o registro e armazenamento de informação numérica e textual, som, imagem estática e em movimento. Eles possuem vários formatos, para vídeo, por exemplo, existem o U-matic⁶, VHS⁷, S-VHS⁸, 8mm e BetaCam⁹, para armazenamento de

⁴ Do inglês, *Compact Disc*.

⁵ Do inglês, *Digital Versatile Disc*.

⁶ U-matic: fitas de ¾ polegadas para videocassete.

⁷ VHS - *Video Home System*, fita de videocassete.

⁸ S-VHS – *Super Video Home System*, fita de melhor qualidade que a VHS, que separa o sinal de luminescência do de crominância.

⁹ BetaCam: fita de meia polegada.

dados computacionais há os disquetes, discos rígidos, fitas DAT¹⁰ e DDS¹¹, por exemplo. De modo geral, os meios magnéticos podem estar disponíveis em rolos, cassetes ou cartuchos. Sua durabilidade se estende às décadas, cerca de 10 a 30 anos, com algumas variações, alcançando uma longevidade maior quando arquivados (acima de 60 anos) (BOSTON, 2000), (BOGART, 2001).

A gravação das fitas magnéticas de modo digital (ou seja, o *bit* 0 é representado pela magnetização positiva, o *bit* 1 será a negativa e vice-versa) possui características que chamam a atenção em relação a sua preservação. Ao contrário da gravação analógica, na gravação digital a percepção da deterioração não ocorre de forma paulatina e perceptível. Uma fita digitalmente gravada exibirá, se houver, pouca deterioração de qualidade até o instante de falha catastrófica, onde grandes seções de informações gravadas já estarão completamente perdidas. Isso se deve ao fato do gravador poder distinguir entre o sinal registrado e o ruído da fita, graças ao sistema de correção de erros, que utiliza *bits* redundantes para reconstruir áreas em que o sinal foi perdido. Essa recuperação de dados diminui perdas de informação no instante da reprodução da fita, quando o desgaste não é muito agressivo, porém, retarda a percepção do verdadeiro estado de preservação da fita (BOGART, 2001).

Portanto, os cuidados com a preservação dos meios magnéticos digitais requerem maior atenção com a conservação. Para se conservar por um período maior, as fitas magnéticas devem ser limpas e usadas de forma adequada. De modo geral, deve-se (BOGART, 2001), (ST. LAURENT, 1997), (BOGART; MERZ, 1995):

- Utilizar e armazenar os rolos e cassetes de fita magnética em um ambiente limpo;
- Não armazenar fitas de rolo em sacos plásticos no interior da caixa da fita. O saco plástico reterá umidade;
- Evitar a contaminação das fitas por sujidades, poeira, impressões digitais, comida, fumaça, cinzas de cigarro e poluentes do ar;
- Não deixar as fitas, ou os cassetes discos rígidos caírem;
- Rebobine (exercite) as fitas de rolo a cada três anos e meio;
- Manter as fitas protegidas da luz solar intensa e evitar o contato com água;

¹⁰ DAT - Digital Audio Tape.

¹¹ DDS – Digital Data Storage tape

- Armazenar as fitas longe de aquecedores, peitoris de janelas, televisores, equipamentos eletrônicos e máquinas em geral;
- Manter uma ventilação apropriada e uma circulação de ar permanente nas estantes para evitar qualquer micro-clima;
- Devolver as fitas à prateleira de armazenamento e guardá-las em pé, quando não estiverem sendo utilizadas. Não se deve permitir que as fitas permaneçam deitadas (flanges da bobina paralelas à tampa da mesa) por longos períodos de tempo;
- Evitar expor a fita ou disco rígido a campos magnéticos fortes, a fim de impedir a perda da informação (como os detectores de metais). Os discos rígidos podem apresentar também problemas com flutuações de energia na gravação de arquivos.

Além dos fatores citados acima, as fitas devem ser inseridas e ejetadas tendo expostas áreas que não contenham informação gravada, pois o mecanismo que puxa a fita do cassete pode danificar a fita se os pinos de direcionamento não estiverem apropriadamente alinhados. Sujidades depositadas sobre o mecanismo de inserção/carregamento podem arranhar a superfície da fita. Rolos de fitas apropriadamente bobinados podem sobreviver a variações maiores de temperatura e umidade sem danos permanentes, comparado aos que não são (BOGART, 2001).

Cuidados com transporte também devem ser tomados, evitando-se o deslocamento quando a temperatura for superior a 43° C, dando preferência às estações de temperaturas mais amenas. Assim como para o armazenamento, as fitas cassetes e disquetes são melhores transportados na posição vertical, de pé, com o peso da fita sendo sustentado pelo eixo da bobina. Naturalmente, é aconselhável o acondicionamento das fitas em embalagens, principalmente as que absorvem choque (embalagens especiais, plástico bolha) podendo ter a vantagem adicional de proporcionar um isolamento que contribua para a proteção dos meios contra as grandes variações de temperatura e umidade (BOSTON, 2000), (BOGART, 2001).

No entanto, para se estender o tempo de vida dos meios, a preocupação mais importante que se pode ter é com o acondicionamento da fita magnética em um ambiente limpo e controlado. Uma pequena variação de temperatura e umidade é recomendada (QUADRO 1) para se evitar o principal fator de degradação das fitas magnéticas, a hidrólise. A hidrólise é um processo químico que age no polímero que constitui a fita, fazendo com que as ligações químicas do

tipo éster do polímero de poliéster sejam rompidas através de uma reação com a água, gerando ácidos orgânicos responsáveis pela decomposição hidrolítica (GRAF. 1). Na fita em que ocorre a hidrólise, pode haver a síndrome da fita grudenta, caracterizada pelo amolecimento da superfície que se torna pegajosa, podendo aderir à cabeça de leitura/gravação, por vezes impedindo a recuperação dos dados. Fitas mal conservadas tornam-se ácidas exalando um odor semelhante ao vinagre. As condições controladas do ambiente também evitam que as montagens da fita fiquem mais “apertadas” – o que resulta na distorção do suporte da fita e em um aumento de *dropouts*¹² permanentes, uma vez que as sujidades incorporadas durante o bobinamento são forçadas para o interior da camada magnética. A adesão entre camadas de fitas, conhecida por *tape blocking*, também pode ocorrer após longos períodos de armazenamento a temperaturas elevadas (BOSTON, 2000), (BOGART, 2001), (ROSS, 99).

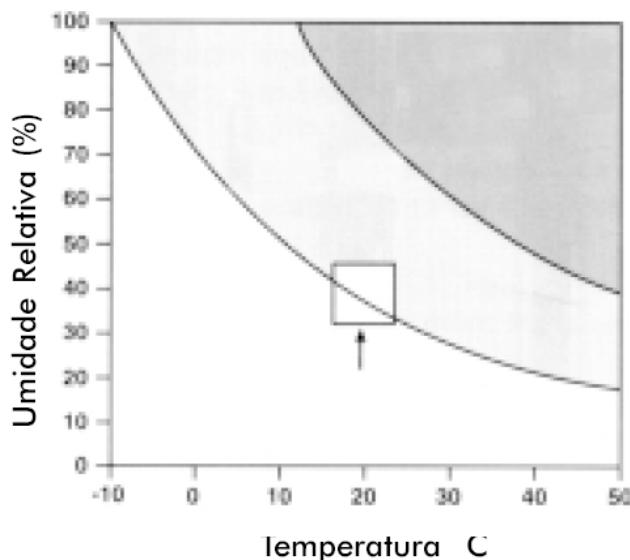


GRÁFICO 1 - Condições de temperatura e umidade e risco de hidrólise

Fonte: BOGART, 2001.

A umidade relativa (UR) também é muito prejudicial, principalmente se associada à alta temperatura, quando pode vir a causar o crescimento de fungos. A umidade relativa elevada causa um aumento na degradação do aglutinante, devido ao elevado conteúdo de umidade no interior da montagem da fita. Também pode provocar aumento das trações (estresses) na

¹² Perdas breves de sinal na leitura da fita devido às sujidades ou deterioração.

montagem da fita – ela absorve umidade e se expande, causando distorção do suporte e aumento de *dropouts* permanentes (BOSTON, 2000), (BOGART, 2001).

É importante observar que meios conservados em baixas temperaturas precisam de um período de lenta aclimatação antes de serem disponibilizados para acesso. Por isso, recomenda-se que haja um armazenamento de acesso e um arquivístico, ou seja, para preservação. Além disso, um armazenamento arquivístico é uma estratégia eficaz para se evitar a perda de informação, já que as fitas estão sujeitas a desaparecer (perda, roubo, destruição por enchente ou incêndio, etc.) (BOSTON, 2000), (BOGART, 2001).

QUADRO 1 - Características-chave de armazenamento de acesso e armazenamento arquivístico

Característica-Chave	Armazenamento de Acesso	Armazenamento Arquivístico
Função	Propiciar um armazenamento para os meios que permita acesso e reprodução imediatos	Propiciar um armazenamento que preserve os meios pelo maior tempo possível
Aclimatação Necessária antes da reprodução	Não	Sim
Expectativa de vida dos meios	Pelo menos dez anos quando armazenado sob as condições de temperatura e umidade indicadas	O máximo permitido ao tipo de meio particular
Ponto de ajuste de temperatura	Próximo ou igual à própria temperatura ambiente interna. Na faixa de 15 a 23°C	Significativamente inferior à temperatura ambiente interna. Tão baixo quanto 5°C
Ponto de ajuste de umidade	Próximo ou igual à própria umidade ambiental interna. Na faixa de 25 a 55% UR	Significativamente inferior à umidade ambiental interna. Tão baixo quanto 20% UR
Variações de temperatura	A diferença entre o valor máximo e o valor mínimo não deve exceder 4°C	A diferença entre o valor máximo e o valor mínimo não deve exceder 4°C
Variações de umidade	A diferença entre o valor máximo e o valor mínimo não deve exceder 20% UR	A diferença entre o valor máximo e o valor mínimo não deve exceder 10% UR

Fonte: BOGART, 2001, p. 30.

Quanto à limpeza, ela deve ser feita para evitar que sujidades diminutas causem perda do sinal de reprodução através da perturbação do contato íntimo necessário entre a superfície da fita e a cabeça de reprodução. As sujidades podem ser responsáveis pela quebra da cabeça de leitura/gravação dos discos rígidos e outros tipos de cabeças giratórias. A poeira ainda, captura a umidade, favorecendo a ocorrência de hidrólise (ST. LAURENT, 1997), (BOSTON,

2000), (BOGART, 2001). Para retirada das sujidades a montagem da fita de rolo pode ser aspirada com um aspirador que tenha uma mangueira comprida, evitando-se a proximidade do motor do aspirador à fita, o que pode vir a causar uma magnetização e consequentemente, perda de dados. As fitas também podem ser limpas com produtos como o *Tape Cleaning Fabric*¹³ da 3M (610-1-150). Este produto têxtil macio coletará sujidades desprendidas, normalmente encontradas nas superfícies das fitas após serem desalojadas pelas fibras do tecido (ST. LAURENT, 1997). Os discos rígidos, especificamente, não costumam serem acometidas com problemas como sujidades, pois são bem acondicionados, mesmo os externos (exceto em casos de sério comprometimento, por exemplo, se submerso em água salgada). Porém, de modo geral, a limpeza mais profunda deve ser realizada por pessoal especializado (BOGART, 1995).

Uma boa prática para controle desse tipo de acervo é a medição sistemática do fator pH¹⁴ das fitas. Para isso, pode ser utilizado um papel indicador, onde uma reação química é provocada tornando o papel amarelado, caso haja contaminações nas fitas (MANNIS, 2005). Outras práticas para conservação das fitas também são recomendadas, como o retencionamento, a regravação e a transcrição. Com o uso a fita vai perdendo sua tensão original devido ao enrolamento da fita. Isso pode fazer com que a fita fique desalinizada no rolo e gere problemas de leitura/gravação. O retencionamento devolve a tensão à fita aplicando força e velocidade, de acordo com o tipo de fita. Já a regravação possibilita o refrescamento do sinal magnético, gravando-se a mesma informação sobre a mesma fita ou sobre outra. Para se alcançar um tempo de vida arquivística, sistemas de gravação, partes sobressalentes suficientes e manuais técnicos também devem ser arquivados juntamente com a fita magnética. Mas em certos casos, como para arquivos de áudio e vídeo, a transcrição será inevitável, em períodos que poderão ser inferiores a dez anos. A transcrição consiste na cópia da fita para outra de mesmo formato ou de formato diferente (BOGART, 2001).

¹³ Tecido para Limpeza de Fita.

¹⁴ PH – Potencial hidrogênio iônico, é um índice indicador de acidez, neutralidade ou alcalinidade de um meio.

2.1.2 Discos Ópticos

Ambos, CDs e DVDs são mídias ópticas, o que quer dizer que são mídias que utilizam a tecnologia da luz (mais especificamente, o laser) para armazenamento e disponibilização de dados. Esses discos são diferencialmente identificados por suas características específicas tais como graváveis, regraváveis e somente leitura. Por exemplo, CD-R, DVD-R, e DVD+R são discos que permitem uma única gravação. CD-RW, DVD-RW, e DVD+RW são discos que permitem apagar algumas informações para que outras possam ser gravadas. DVD-RAM são discos que podem ser graváveis e regraváveis, eles são formatados para acesso randômico¹⁵, como os discos rígidos dos computadores. CD-ROM e DVD-ROM são prensados e modelados, são discos não graváveis, servindo apenas para leitura. No caso específico dos DVDs há ainda outra diferenciação, os graváveis (DVD-R, DVD-RW, DVD-RAM) podem ser lidos ou gravados em um dos lados ou nos dois lados, ou seja, podem ser fabricados com uma camada para gravação em cada lado, e o DVD ROM, que é pré-gravados, pode possuir dois lados para leitura (BYERS, 2003).

De acordo com fabricantes, com o armazenamento apropriado, os discos ópticos podem ultrapassar o tempo de duração da tecnologia. Os CD-R, DVD-R, e DVD+R têm uma expectativa de vida de 100 a 200 anos ou mais. Os discos CD-RW, DVD-RW, DVD+RW, e DVD-RAM têm uma expectativa de vida de 25 anos ou mais. Os CD-ROM e DVD-ROM (incluindo áudio e vídeo) possuem uma expectativa que varia de 20 a 100 anos. A expectativa de vida dos discos é definida pelo tempo que um disco permanece utilizável, portanto, dentro deste período estimado pelos fabricantes o disco certamente já terá iniciado seu processo de degradação, porém este não será de pronto percebido devido ao detector de erros e à capacidade de correção desenvolvida em alguns sistemas (BYERS, 2003). No entanto, os discos puramente ópticos parecem ser mais resistentes a danos do que as fitas magnéticas, por serem imunes a campos magnéticos e por obterem resultados mais eficientes após um processo de recuperação.

De modo geral, os CDs e DVDs são formados por camadas. A mais superficial delas é constituída de um substrato de policarbonato, que abriga a área lida pelo laser (o plástico

¹⁵ Acesso randômico é aquele que ocorre aleatoriamente, podendo começar de qualquer ponto do disco, ao contrário do acesso seqüencial, onde os dados são lidos, como o nome implica, de forma linear, um a um.

transparente que constitui a maioria dos discos), provendo a profundidade necessária para manter o foco do laser nas outras camadas mais internas (FIG. 1). Há também a camada de dados, que como o nome implica, guarda os dados. O dado aparece como marcas ou pequenas depressões que tanto absorvem a luz do raio laser, como transmitem a luz de volta ao fotosensor de laser através da camada refletiva de metal. Os tipos de metais refletivos tipicamente utilizados são alumínio, prata ou liga de prata. Nos DVDs com camada dupla de gravação e/ou leitura, o silicone pode ser utilizado como uma das camadas semi-refletivas (BYERS, 2003).



FIGURA 1 - Foto das camadas de um CD-ROM.

Fonte: INNARELLI, 2006.

Os discos ROMs são discos replicados – os dados são fisicamente comprimidos no disco quando ele está sendo fabricado. ROMs são geralmente produções em massa e contêm músicas, vídeos, jogos interativos, gráficos, aplicações de computador. Os discos *Recordabel* (Rs) possuem o mesmo padrão de leitura dos ROMs e são mais apropriados para poucas cópias de uma mesma informação (SIMÕES, 2002). Os discos ROMs e Rs são mais recomendados e utilizados para uso arquivístico, dada a característica de serem graváveis por uma única vez. Essa qualidade atribui a eles uma garantia de fidedignidade e autenticidade dos dados gravados (INNARELLI, 2006). Portanto, neste capítulo, esses tipos de discos merecem destaque.

A camada de dados costuma ser a primeira a sofrer degradação por ser mais suscetível a fatores ambientais, porém, o manuseio incorreto costuma agredir primeiro a camada de policarbonato, pois este pode se flexionar ou curvar-se se armazenado por um longo período numa posição não vertical (BYERS, 2003). Os fatores que mais atingem os discos são a

temperatura, umidade, tempo de uso e a qualidade da mídia (INNARELLI, 2006). Os principais cuidados a serem tomados são (BYERS, 2003), (ST. LAURENT, 1997):

- Controle de temperatura e umidade relativa. Discos mantidos em ambientes refrigerados, com baixa umidade e não sujeitos à mudanças ambientais bruscas devem durar mais. Temperaturas de acondicionamento e umidade relativa recomendados por várias fontes técnicas estão apresentadas no quadro (QUADRO 2).
- Uso de acondicionamento e armazenagem adequados. Os discos devem ser guardados em caixas apropriadas, de forma isolada e mantidos na posição vertical, para evitar que fiquem arqueados;
- Manter os discos protegidos da luz solar, evitando, desta forma, o calor causado por ele e evitar contato com a água. Para os discos Rs a ação da luz é agressiva, pois pode danificar a camada de gravação que é constituída por uma tinta. Isso já não ocorre com os discos ROMs, entretanto, ambos os discos são mais acometidos pelo calor;
- Evitar a contaminação com sujidades, principalmente no lado de leitura do laser;
- Evitar arranhões, principalmente no lado de leitura do laser e os que seguem a direção da leitura (tangentes ao centro do disco);
- Evitar o contato com solventes orgânicos fortes, como acetona ou benzeno, que podem causar danos à camada de policarbonato;
- Evitar flexionar ou entortar os discos;
- Quando possível, evitar o uso de etiquetas para discos que devem ser armazenados por períodos superiores a cinco anos. A etiqueta possui um material diferente e está sujeita a deterioração típica de seus componentes químicos.
- Procurar remover os discos de suas caixas pressionando o polegar e o terceiro dedo sobre as bordas do disco próximas à parte superior e à parte inferior da caixa e pressionando o fecho de plástico no centro da caixa com a outra mão.

QUADRO 2 - Parâmetros recomendáveis para armazenamento em diferentes fontes

Fontes	Mídias	Temperatura	Gradiente de temperatura máxima	Umidade relativa (UR)	Gradiente de (UR) máxima
ISSO TC 171/SC Jan. 2002	CD-R CD-ROM	-5°C a 20°C	4 °C/hora	30% a 50%	10%/hora
IT9.25 e ISSO 18925	CDs DVDr	-10°C a 23°C	...	20% a 50%	Ciclo entre: +/-

Fev. 2002					10%
NARA, FAQ sobre mídias ópticas, Abr. 2001	CDs DVDs	20°C	+/- 0.6°C/dia	40%	5%/dia
National Archives of Australia, Abr. 1999	CDs	18°C a 20°C	...	45% a 50%	10%/dia
Library Technical Report Nov.-Dez. 1997	CDs	-10°C a 50°C	...	10% a 90%	...
DVD Demystified, 2 ed., Jim Taylor, 2001	DVD-R DVD-ROM	-20°C a 50°C	15°C/hora	5% a 90%	10%/hora
	DVD-RAM	-10°C a 50°C	10°C/hora	3% a 85%	10%/hora
	DVD+RW	10°C a 55°C	15°C/hora	3% a 90%	10%/hora
National Library of Canada, 1996	CDs	15°C a 20°C	2°C/dia	25% a 45%	5%/dia
Media Sciences, Inc. Jerome L. Hartke Jul. 2001	CD-R	10°C a 15°C	...	20% a 50%	...

Fonte: BYERS, 2003.

Outros cuidados devem ser tomados com os discos, se forem armazenados a temperatura relativa muito inferior à do ambiente, eles deverão ser gradualmente aclimatados a fim de reduzir a condensação da umidade e as tensões. Manter o disco dentro de sua embalagem irá ajudar numa aclimatação gradual. As tensões advindas de freqüentes mudanças de temperatura também podem ser minimizadas optando-se por manter os discos usados freqüentemente a temperaturas similares a do ambiente no qual eles serão utilizados (BYERS, 2003), (BOSTON, 1999).

A limpeza dos discos deve ser adequada ao tipo de sujidade e realizada por profissionais¹⁶. Luvas adequadas (de algodão, nylon ou látex) devem ser usadas para evitar a transferência de óleos corporais, impressões digitais e ranhuras na superfície do disco (INNARELLI, 2006). A

¹⁶ Para maior aprofundamento, ver referências (BOGART, 1995), (BYERS, 2003).

limpeza deve seguir a direção radial do centro do disco às bordas, evitando-se a fricção. Um pano macio embebido numa solução deve ser usado, depois deve-se enxaguar, passar o pano novamente, desta vez, embebido apenas em água destilada e enxugar bem o local, usando um pano de algodão macio (BYERS, 2003). Uma pistola de ar também deve ser utilizada para eliminar resquícios deixados sobre o disco (ST. LAURENT, 1997). Caso seja necessário remover impressões digitais ou outras manchas, pode ser utilizada a solução de Tergitol em água destilada. Entretanto, a água isoladamente não consegue remover alguns tipos de sujeiras, podendo ser necessário o uso de solventes suaves, como o álcool iso-propílico ou metanol, que evaporam rapidamente e não dissolvem o policarbonato. Eles talvez, contudo, dissolvam ou causem danos à etiqueta ou a cobertura opcional no lado da etiqueta do disco (ST. LAURENT, 1997), (BYERS, 2003).

Controles de qualidade periódicos com uso de programas e máquinas especiais podem permitir avaliar a deterioração progressiva da mídia possibilitando a seleção de registros informacionais em mídias que tenham alcançado excelente qualidade nos teste ou mesmo permitindo a transcrição da informação em tempo (VALLE, 2003), (INNARELLI, 2006).

2.2 Seleção de objetos digitais a serem mantidos

As decisões tomadas para a seleção do patrimônio digital são cruciais para o sucesso do projeto de preservação. Essa seleção abrange a escolha fundamentada do patrimônio digital, que precisa ser convertido para outro formato digital mais atual ou migrar de suporte, com vistas à longevidade. Deve-se decidir o que deve ser descartado, o que deve ser preservado e por quanto tempo (NLA, 2003), (KENNEY; RIEGER, 2000).

Embora a digitalização por si não constitua preservação, as razões que influenciam a conversão de materiais para o formato digital são também partes do processo de seleção para preservação digital. Dentre elas está a consideração fundamental de que a seleção deve se basear em fatores estabelecidos pela política de desenvolvimento da coleção, nas necessidades das partes interessadas, nos objetivos da instituição e no conhecimento dos materiais e de seu contexto (KENNEY; RIEGER, 2000).

Muitas instituições têm desenvolvido seus guias para seleção de materiais. Estes guias freqüentemente servem tanto para enunciar políticas e esboçar as intenções da instituição, quanto para apontar estratégias para o desenvolvimento de coleções digitais. Para criação destes guias e definição de critérios de seleção, deve-se ter uma preocupação com a autenticidade e com as estratégias empregadas, considerando-se o valor do material de acordo com a missão da organização responsável pela preservação, os custos, dificuldades prováveis para a manutenção e as necessidades da comunidade para a qual o material é conservado. Deve-se ainda ter um cuidado com a documentação registrada sobre o bem digital. Por exemplo, é possível prever que alguns programas terão que permitir aos usuários o uso de versões antigas de simuladores, jogos eletrônicos ou arte digital para reproduzir a experiência original ao invés da experiência acelerada que podem oferecer as tecnologias posteriores. Assim, quando os materiais digitais só podem ser compreendidos fazendo-se referência a conjuntos de regras, com informações contextuais, o processo de seleção também deve determinar qual documentação deve ser preservada (KENNEY; RIEGER, 2000).

Outros aspectos que podem auxiliar no processo de seleção foram levantados pela UNESCO, em forma de perguntas que devem ser feitas pela instituição responsável pela preservação (NLA, 2003, p. 70, *tradução da autora*):

- [...] Porque é válido conservar os materiais? O que os dá o valor que justifica o esforço para sua preservação? Este valor está vinculado com:
 - Evidências, provas?
 - Informação?
 - Fatores artísticos ou estéticos?
 - Uma inovação importante?
 - Uma associação histórica ou cultural?
 - O que um usuário pode executar alguma atividade utilizando o material? O que o usuário pode fazer com que o material faça?
 - Características culturalmente significativas?
- O valor do material depende de sua aparência? (Ele deve ter seu valor perdido ou significativamente decrescido se apresentar diferença estética?)
- O valor do material depende do modo como ele funciona? O valor se perderia se certas funcionalidades fossem perdidas?
- O valor depende do contexto do material? O valor se perderia se as ligações (links) relacionadas ao material não funcionassem? Ou se o usuário não pudesse ver a prova de que o material está conectado com seu contexto original?
- Necessita de alguns elementos de navegação e funções de visualização, que não podem ser substituídas?

- Caso seja difícil de se definir as necessidades que devem ser mantidas, pode ser mais fácil considerar o impacto de um elemento que não se mantenha e determinar as funções ou os elementos que definitivamente não são necessários.

Mais fatores determinantes no processo de seleção estão citados a seguir (KENNEY; RIEGER, 2000), (GERTZ, 1998):

- O conteúdo intelectual, o valor literário dos materiais, ou valor duradouro;
- Importância histórica;
- Consonância com a missão da instituição;
- Prover acesso a uma coleção, evitando o manuseio e uso de materiais originais frágeis ou muito utilizados, que estão danificados ou correm risco de extermínio;
- O custo da preservação está compatível com o valor do item;
- Direitos autorais.

A manutenção de materiais digitais para conservação de originais analógicos é uma justificativa suficiente para a escolha de objetos para a preservação. Porém, a seleção baseada apenas no valor intelectual provavelmente não será motivo satisfatório para justificar a preservação de uma coleção, pois devem ser considerados fatores como a mensuração da quantidade de uso e o valor histórico do material. Mas, de modo geral, o conteúdo do material pode ser avaliado considerando-se o seu valor histórico e físico, se ele é um dos melhores dentro do assunto que trata, se o conteúdo é amplo, se é preciso e tratado com profundidade (KENNEY; RIEGER, 2000). O direito autoral merece especial atenção e será tratado a seguir.

Além de todas essas considerações, diante de recursos escassos para sustentação de uma grande coleção, será necessária uma seleção mais criteriosa, limitando o acervo a materiais de altíssimo valor. Os custos da preservação também estão vinculados a diversidade de problemas e formatos que devem ser administrados, de modo que, também pode ser razoável se limitar os tipos de materiais selecionados a poucos formatos (NLA, 2003).

2.2.1 Direitos Autorais

Os direitos autorais influem nas decisões para a seleção de materiais para preservação e devem receber primazia. Materiais com direitos muito restritivos a ponto de impedir a negociação de acordos sobre o acesso no futuro podem ser escolhidos para descarte na seleção para preservação, pois se tornaram inúteis no patrimônio (KENNEY; RIEGER, 2000).

As regras do direito autoral parecem bater de frente ao interesse público e ao direito da nação à cultura, uma vez que restringem o uso das obras, protegendo a propriedade do autor. No entanto, estas regras buscam incentivar a divulgação da obra e a expressão de seu criador. O autor, estando seguro da proteção de sua obra de usos ilícitos, a divulgará à comunidade (GARZON, 2006). É possível vislumbrar, portanto, um equilíbrio entre os benefícios da aplicação do direito autoral à sociedade e ao proprietário intelectual, pois ocorrem limitações e condições que beneficiam ambas as partes.

No Brasil, os direitos autorais correspondem ao direito exclusivo do autor sobre seu trabalho intelectual, literário, artístico ou científico, expresso por qualquer meio ou fixado em qualquer suporte, tangível ou não. Abrangendo portanto, entre outros, a música, a literatura, as artes plásticas e os programas de computador. Ele inclui o direito do autor de controlar a reprodução, cópia, performance, exibição e outros usos de um trabalho. E é mais abrangente do que o *copyright* (direito de cópia) por corresponder além do direito de exploração econômica aos direitos morais. Ele possui efeito a partir do momento da criação do objeto de direito, ou seja diante do ato de produção, quando a criação é expressa por qualquer meio (BRASIL, 1998).

Toda obra intelectual digital ou não, independente de seu meio (papel, *Internet*) tem a sua proteção válida dentro de um período. O direito patrimonial tem sua validade até setenta anos após a morte do autor, contados a partir de primeiro de janeiro do ano subsequente ao do seu falecimento. Para obras audiovisuais e fotográficas a validade é menor, sendo de setenta anos, contado a partir de primeiro de janeiro do ano seguinte à divulgação do trabalho (MARTINS FILHO, 1998).

É necessária a autorização do autor para emprego de sua obra para, dentre outros usos, reprodução, edição, adaptação e tradução. No entanto, a utilização da obra não ferirá o direito de autor quando tiver seu uso permitido e em alguns casos, quando o nome do autor for mencionado, por exemplo (BRASIL, 1998).

A lei restringe sua validade ao seu território. Para proteger o direito do autor pelo mundo existem convênios internacionais (MARTINS FILHO, 1998). A importância do estabelecimento de convênios internacionais incorporados à legislação interna dos países teve sua importância agravada com a disseminação e o acelerado compartilhamento de conteúdo provocado pela *Internet*. A partir desse agravante, surgiram também estruturas colaborativas como os modelos de software livre e o projeto *Creative Commons* (bens comuns criativos), que vêm adequar os direitos autorais a uma nova realidade, permitindo ao criador a liberação de sua obra de acordo com a sua conveniência mediante licenças (GARZON, 2006). Muitos países já aderiram ao uso do *Creative Commons*. No Brasil, as licenças já se encontram totalmente adaptadas à legislação brasileira.

Os direitos do autor devem ser observados como um importante critério no momento da seleção para preservação digital, seja no momento inicial da formação do acervo digital, com a escolha de itens para digitalização, seja no instante de tornar o acervo acessível à comunidade. A instituição responsável pela preservação ao divulgar o seu acervo em meios como a *Internet* deve resguardar os direitos autorais, estando atenta aos modos de se proteger contra o mal uso de seu patrimônio e à melhor maneira de estabelecer um diálogo com a sociedade.

2.3 Autenticidade no acervo digital

Para se preservar um objeto digital, deve-se assegurar que ele é autêntico a fim de que recursos não sejam consumidos com a manutenção e conservação de objetos sem valor para a sociedade e para sua instituição.

Outros cuidados também devem ser observados, uma vez que, “converter um objeto digital de um programa para outro, ou o migrar de versão a versão, pode apresentar problemas de autenticidade que podem ou não ser resolvidos com uma cuidadosa atenção com a procedência” (CONCIL ON LIBRARY AND INFORMATION RESOURCES, 2000a, *tradução da autora*). É como ocorre com os CD-ROM e CD-R, que prometem garantia de autenticidade e fidelidade ao seu conteúdo, uma vez que a gravação nestas mídias é definitiva, não podendo ocorrer alterações (INNARELLI, 2006).

Um registro é dito autêntico quando ele é o documento que declara ser. Prover autenticidade a um documento garante que ele não resulta de uma manipulação, substituição, ou falsificação ocorrida no seu processo de criação, e que ele é o que ele se propõe a ser. Dessa forma, a autenticação da reprodução de um documento, realizada por um órgão oficial autorizado para execução dessa função, confere à cópia a mesma força do documento transscrito ou reproduzido, seja ele um rascunho, um original ou outra cópia, mas também com seu mesmo grau de fidelidade (DURANTE, 1995).

A autenticidade não implica na fidelidade de um registro. Se o documento reproduzido não poder ser tratado como um fato quando criado e no seu status original, a autenticação da cópia não confere essa capacidade, ou seja, não dá ao documento fidelidade, não o torna genuíno (DURANTE, 1995).

Como ocorre em documentos escritos em papel, os documentos digitais devem ser autenticados, garantindo que não houve manipulação ou substituição do conteúdo digital. “Porque objetos digitais possuem menos evidência da autoria, originalidade, procedência e outros atributos aceitáveis para objetos analógicos, o autor está sujeito a uma adicional suspeita” (CLIR, 2000a, *tradução da autora*). Dessa forma, testes precisam ser realizados para autenticação.

Existem recursos para se atestar autenticidade em documentos digitais, tais como, criptografia, marca d’água e a assinatura digital (usada para autenticar o remetente de uma mensagem).

2.4 Estratégias para preservação digital

É vital que haja um lugar para uma estratégia de preservação e proteção dos mecanismos de armazenamento e visualização, onde as coleções poderão ser sustentáveis por um longo período, apesar da obsolescência da tecnologia.

A escolha da estratégia de preservação deve ser fundamentada em fatores como qualidade e tipo do acervo (de imagens, texto, multimídia), custo/benefício, ambiente computacional e programas. É fortemente necessário para tanto, estar atento ao cenário tecnológico e suas mudanças a fim de se tomar medidas que tentem impedir perdas de informações a tempo, alcançando um controle de riscos. As técnicas para isso são o refrescamento (transferência do conteúdo digital para um suporte mais atual), a migração de dados (conversão de formato de arquivos, mudança de ambiente computacional), a emulação (criação de simuladores de ambientes operacionais, na tentativa de preservar a experiência homem-máquina). Essas estratégias, que podem ser utilizadas em conjunto, requerem um esforço de trabalho, pois não são automáticas e nem automatizadas, sendo necessário planejamento e análise ao se vislumbrar a necessidade de emprego de alguma delas. Também é indicada a observação das restrições legais que podem afetar a prática destas estratégias (KENNEY; RIEGER, 2000).

2.4.1 Refrescamento

Para o padrão OAIS de preservação¹⁷, refrescamento é simplesmente a cópia de todos os *bits* de uma mídia para outra idêntica, porém nova¹⁸ (KENNEY; RIEGER, 2000). No entanto, ele também se refere à mudança de um mesmo formato para atender compatibilidade de versões do programa, por exemplo (HOWELL, 2001). Ele pode ser empregado, como um cuidado preventivo para impedir que a mídia chegue ao seu limite de longevidade. Como visto no Capítulo 2.1, sobre fitas magnéticas, o refrescamento também pode ser usado para se replicar

¹⁷ OAIS, abordado no Capítulo 2.5.

¹⁸ Apesar do refrescamento se referir à cópia exata da informação, *bit a bit*, durante seu processo podem ocorrer perdas não intencionais (KENNEY; RIEGER, 2000).

a mesma informação numa mesma fita, a fim de reforçar a magnetização dos dados nela gravados.

Portanto, é uma estratégia paliativa, sendo válida dentro de um período restrito (enquanto a indústria der suporte à mídia). Ela pode ser entendida como parte da rotina de manutenção de uma coleção digital (KENNEY; RIEGER, 2000), (HOWELL, 2001), quando observados os direitos de cópia, afinal, o uso da mídia pode ter um controle de acesso limitado (KENNEY; RIEGER, 2000).

2.4.2 Migração

“Migração é o processo de transferência da informação digital de um ambiente computacional (máquina, sistema e programa) para outro ou de uma geração de computadores para outra” (KENNEY; RIEGER, 2000, *tradução da autora*).

A migração pode envolver conversões, reformatação, ou reescrita do código de um programa, a fim de manter sua compatibilidade com o ambiente computacional atual (DOLLAR, 1999), sem contudo, comprometer a integridade dos dados (KENNEY; RIEGER, 2000). Ela se faz necessária para o alcance dos benefícios da nova tecnologia (como maior agilidade no processamento dos arquivos), para evitar a perda de dados em suportes que estão à margem da obsolescência, para atualizar sistemas com defeitos conhecidos (como erro em programas) e para criação de um ambiente computacional mais estável (HOWELL, 2001), (KENNEY; RIEGER, 2000).

Os riscos da migração irão depender do processo adotado. Dentre os possíveis problemas decorrentes da migração, naturalmente, estão os relacionados à conversão, desta forma, sucessivas conversões podem corromper o documento digital ou afetar a experiência homem-máquina desejada para o arquivo (em decorrência da mudança de ambientação) (SITTS, 2000), (KENNEY; RIEGER, 2000). Existe também a possibilidade de escolha de formatos que podem diminuir substancialmente o tamanho do arquivo, comprimindo-o. Essa compressão pode ser com ou sem perdas. Na escolha de formatos com perdas, a informação

desprezada no objeto digital¹⁹ (exemplo de compressão de imagens e som), poderá vir a ser útil no futuro para sistemas que façam busca pelo conteúdo do arquivo (SITTS, 2000). Outro tipo de problema pode estar relacionado à incompatibilidade do arquivo digital migrado com o novo ambiente ou com a perda não intencional de dados. Deve-se também observar as restrições legais que podem impedir que um arquivo tenha seu formato alterado (KENNEY; RIEGER, 2000).

2.4.3 Emulação

A emulação tenta salvaguardar o registro informacional de forma fiel, preservando o seu comportamento e modo de interação junto ao usuário (SITTS, 2000), pois “a emulação envolve a recriação do ambiente técnico para visualização e uso da coleção digital” (KENNEY; RIEGER, 2000, *tradução da autora*). Isso pode ser possível quando a documentação sobre os requisitos do ambiente computacional (máquinas, sistemas e programas) é mantida, além do conteúdo digital original, ou o que tenha sofrido refrescamento ou migração (KENNEY; RIEGER, 2000), (ROTHENBERG, 1998). Para Rothenberg (1998), a emulação é a única forma de se assegurar o comportamento apropriado (original) de um documento digital. Em contra partida, Bearman (1999), além de acreditar na inviabilidade da emulação, afirma que ela não funciona como uma estratégia de preservação, dando preferência à salvaguarda da funcionalidade em detrimento da integridade e autenticidade dos dados.

A emulação envolve um alto custo sem garantias de obtenção de um bom resultado (KENNEY; RIEGER, 2000). Alguns ambientes computacionais são muito difíceis de serem emulados por envolverem a utilização de equipamentos ultrapassados como teclado e *joystick* específicos (HEDSTROM; LAMPE 2001). Antes de se empregar a emulação é interessante verificar se o conhecimento do uso do ambiente antigo foi mantido pelos usuários (SITTS, 2000), e observar que a preservação que excede o que o usuário necessita pode ser muito dispendiosa inutilmente. Em alguns casos, os usuários podem preferir a não adoção de um

¹⁹ A informação retirada do arquivo não é relevante para o homem. Ele não consegue perceber a perda devido às limitações de seus sistemas auditivo ou visual.

ambiente emulado por questões como difícil usabilidade e performance (HEDSTROM; LAMPE 2001).

Um estudo baseado na interação homem-computador do projeto CAMiLEON²⁰ submeteu um grupo de usuários a um teste empírico com um jogo de computador antigo e conhecido, em um ambiente emulado e numa versão migrada, tendo como base o jogo no seu ambiente original preservado. Segundo a análise dos autores deste estudo, os resultados obtidos não apontam diferenças significativas estatisticamente do nível de satisfação dos usuários, performance ou facilidade de uso. Desta forma, através da migração também é possível alcançar resultados de interação homem-máquina similares aos obtidos com a emulação (HEDSTROM; LAMPE 2001).

Em oposição, Rothenberg (1998) lembra, que uma vez criado um ambiente emulado para um tipo de dados, esse, certamente poderá ser aproveitado por outros tipos de dados. Rothenberg (1998) sugere assim, que as instituições se associem para alcance dessa meta em comum. Porém, não se tem conhecimento de instituições que tenham adotado essa medida (BEARMAN, 1999).

A emulação também levanta questões relacionadas à propriedade intelectual para a aquisição e manutenção de informações proprietárias necessárias (especificações ou documentações do programa, sistema ou máquina) (KENNEY; RIEGER, 2000).

2.4.4 Arqueologia digital

A arqueologia digital envolve métodos e procedimentos para salvar o conteúdo digital da obsolescência ou danos sofridos por uma mídia ou ambiente computacional. Essa estratégia pode ser, portanto, uma alternativa para casos, onde a migração e o refrescamento não foram realizados a tempo (KENNEY; RIEGER, 2001).

²⁰ CAMiLEON – *Creative Archiving at Michigan & Leeds: Emulating the Old on the New.*

Nessa estratégia, os suportes do conteúdo digital podem passar por procedimentos de restauração para recuperação dos dados. Tais procedimentos podem ser simples como os de limpeza, abordados no Capítulo 2.1 ou podem ser mais avançados, sendo necessário o uso de equipamentos mais sofisticados, como os de microscopia magnética. Dependendo da solução de recuperação, essa estratégia, como uma medida corretiva, pode vir a ter um custo impeditivo.

2.5 Padrão para Preservação ISO 14721 – *Open Archival Information System (OAIS)*

A aplicação do modelo OAIS resulta num conjunto de boas práticas que são indispensáveis para a preservação de longo prazo. Suas diretrizes têm alcançado ampla aceitação, sendo adotado por projetos como o *Network European Deposit Library*. - NEDLIB²¹, abordado no Capítulo 2.6.

O modelo de referência OAIS foi desenvolvido pelo *Consultative Committee for Space Data System* - CCSDS²², ligado à NASA²³, para a ISO²⁴, gerando a norma ISO 14721:2002. Este modelo provê um sistema conceitual para preservação e acesso de informações digitais de longo período, incluindo terminologia e conceitos para descrição e comparação de arquiteturas de arquivos, estratégias e técnicas de preservação. Ele tem o objetivo de facilitar a melhor compreensão dos requisitos para preservação e acesso, provendo um padrão que possa ser utilizado por várias intuições (KENNEY; RIEGER, 2000). Ele possui a participação de diversos países, incluindo o Brasil, através do INPE²⁵ (CCSDS, 2002).

De modo geral, o Modelo de Referência OAIS procede de um modelo ambiental conceitual de alto nível, a partir do qual, derivações em níveis mais detalhados vão ocorrendo. A ilustração (FIG. 2) representa o modelo simples, descrevendo o ambiente que envolve um OAIS.

²¹ *Networked European Deposit Library*. Disponível em: <<http://nedlib.kb.nl/>>.

²² Disponível em: <<http://www.ccsds.org>>.

²³ NASA - *National Aeronautics and Space Administration*. Disponível em: <<http://www.nasa.gov/>>.

²⁴ ISO - *International Organization for Standardization*.

²⁵ INPE – Instituto Nacional de Pesquisas Espaciais. Disponível em: <<http://www.inpe.br/>>.



FIGURA 2 - Modelo ambiental de um OAIS.

Fonte: CCSDS, 2002, *tradução da autora.*

Neste modelo, atuam quatro intervenientes, que podem ser entendidos como segue (HOWELL, 2001): Produtores se referem a pessoas ou sistemas clientes que fornecem a informação a ser preservada. O Gerenciamento é representado pela entidade responsável pelo estabelecimento de políticas no arquivo. Já os Usuários são as pessoas ou sistemas clientes que integram os serviços do OAIS para recuperar e adquirir informações preservadas de interesse.

As informações submetidas a um OAIS por um Produtor e toda difusão estabelecida a partir de um OAIS a um Usuário ocorrem através de pacotes. Os pacotes de informação se referem à idéia conceitual de um conjunto de informações e metadados.

No próximo nível de detalhamento do OAIS, seis entidades funcionais são identificadas (CCSDS, 2002):

- Recolhimento: Responsável pela aceitação de pacotes de Informação Submetidos pelos Produtores (ISP) (ou elementos internos, acima do controle administrativo) e pela preparação dos conteúdos para armazenamento e gerenciamento de arquivos, se responsabilizando pela garantia da qualidade e geração de informações descritivas e Pacotes de Arquivamento de Informação (PAI).
- Armazenamento: Armazena, mantêm e recupera pacotes de arquivamento de informação (PAI), recebendo-os e guardando-os na área de armazenamento permanente. Sendo responsável pela gerencia hierárquica de armazenamento, refrescamento de mídias nas quais os arquivos estão armazenados, verificação de erros, recuperação de falhas e fornecimento de dados (PAIs) para acesso.

- Gerenciamento de dados: Envolve serviços e funções para armazenar, manter e acessar a informação descritiva, que identifica e documenta os acervos do arquivo, e os dados administrativos utilizados para gerenciá-los. Responsável pela captura e gerência de metadados necessários para operar o sistema, atualizações e execução de buscas e emissão de relatórios. Essa entidade funcional administra as funções do banco de dados, gerando definições de esquemas e visões.
- Administração: Parte responsável por serviços e funções para a gerência operacional do arquivo como um todo, realizando solicitações e negociações de acordos de submissão com os produtores, gerência de configuração do sistema, atualização/migração de informações do arquivo, controle físico de acesso, auditoria de submissões e atendimento ao cliente. Sendo também responsável pelo estabelecimento e manutenção de padrões e políticas para arquivos, provendo suporte aos consumidores.
- Planejamento de preservação: Define e gerencia estratégias que permitam que objetos digitais armazenados, pacotes de arquivamento de informação (PAI), atravessem o tempo sem sofrer perdas inaceitáveis e mudanças de conteúdo ou funcionalidade. Monitora o ambiente OAIS e fornece recomendações que garantem que a informação armazenada permanece acessível para a comunidade interessada ao longo do tempo, mesmo que o ambiente computacional original se torne obsoleto. Para tanto, observa a transitoriedade tecnológica, desenvolve estratégias para padrões de preservação e planos de migração.
- Acesso: Coordena atividades de acesso, gerando serviços e funções para determinação da existência, descrição, localização e disponibilização da informação armazenada no OAIS, gerando o Pacote de Disseminação de Informação (PDI), resultados e relatórios.

A ilustração (FIG. 3) descreve a relação conceitual entre as seis entidades funcionais e as três variações de pacotes de informação.

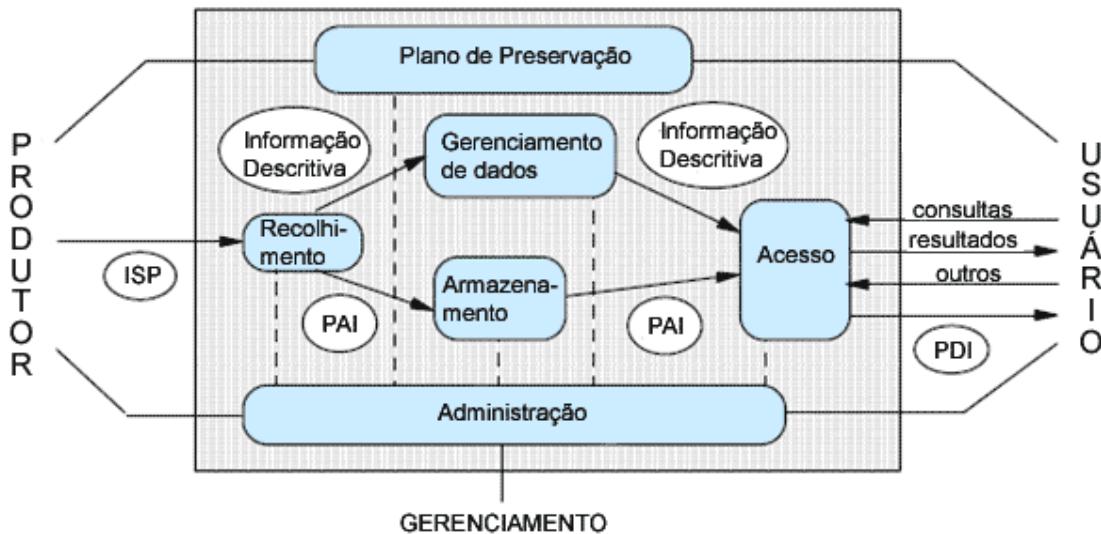


FIGURA 3 - OAIS – Modelo com entidades funcionais.

Fonte: CCSDS, 2002, tradução da autora.

O modelo OAIS é um modelo, como dito anteriormente, de referência, portanto, ele é muito geral para ser o suficiente numa implementação de projeto. Ele precisa ser refinado e estendido para ser útil em um domínio específico. Devido a essa característica, Gladney (2002) o considera como sendo apenas uma ontologia²⁶. Contudo, projetos de preservação podem adotar o modelo adicionando seus requisitos específicos e conhecimentos para arquivamento.

2.6 Iniciativas institucionais

A preocupação com a longevidade digital tem estimulado diversas iniciativas, a exemplo de projetos como CAMILEON²⁷, CEDARS²⁸, PANDORA²⁹, FEDORA³⁰ e OCLC/RLG³¹ *Working Group on Preservation Metadata*.

²⁶ Ontologia: Modelo de dados que representa um conjunto de conceitos dentro de um domínio e os relacionamentos entre estes. Uma ontologia é utilizada para realizar inferência sobre os objetos do domínio.

²⁷ CAMILEON – *Creative Archiving at Michigan & Leeds: Emulating the Old on the New*. Disponível em: <<http://www.si.umich.edu/CAMILEON/about/aboutcam.html>>.

²⁸ CEDARS – *CURL Exemplars in Digital Archives*. Disponível em: <<http://www.leeds.ac.uk/cedars/>>.

²⁹ PANDORA – *Preserving and Accessing Networked Documentary Resources of Australia*. Disponível em: <<http://pandora.nla.gov.au/index.html>>.

Neste capítulo, é possível se ter um breve panorama de como a preservação digital vem sendo tratada por instituições definidoras de diretrizes, no exterior e no Brasil, através dos projetos NEDLIB, e-ARQ. Por fim, apresenta-se um exemplo de instituição detentora de patrimônio digital no Brasil, o Projeto Portinari. Existem, porém, outras instituições, tais como, a Superintendência de Museus do Estado de Minas Gerais e as fundações Iberé Camargo³² e Inimá de Paula³³.

2.6.1 *Networked European Deposit Library - NEDLIB*

NEDLIB é um projeto colaborativo da Biblioteca Nacional Européia. Finalizado em 2001, seus objetivos visam a disponibilização de publicações no presente e no futuro através do desenvolvimento de uma infra-estrutura, com ferramentas básicas para a criação de um sistema de arquivamento virtual para publicações eletrônicas - DSEP³⁴ (WERF-DAVELAAR, 1999), (NEDLIB, 2000). Apesar de ser voltado para bibliotecas, o NEDLIB também pode ser adequado para uso em instituições arquivísticas (NEDLIB, 2000).

Seus principais objetivos foram alcançados. Eram eles (WERF-DAVELAAR, 1999):

- Identificar requisitos funcionais comuns aos arquivos de bibliotecas para obtenção de um desenho genérico de alto nível para o DSEP que pudesse servir de base para implementações de outros sistemas para bibliotecas.
- Definir as características de arquivos digitais e a preservação adequada a cada tipo, mantendo a autenticidade.
- Desenvolver um sistema demonstrativo (DSEP) com ferramentas e programas prontos para uso.

³⁰ FEDORA – *Flexible Extensible Digital Object and Repository Architecture*. Disponível em: <<http://www.fedora.info>>.

³¹ OCLC/RLG - *Online Computer Library Center/ Research Libraries Group*. Disponível em: <<http://www.oclc.org/research/pmwg/>>.

³² <http://www.iberecamargo.org.br/>

³³ <http://www.inima.org.br/>

³⁴ DSEP - *Deposit Systems for Electronic Publications*.

Em 1998, o projeto adotou o OAIS como base para modelagem do DSEP. Com isso foi possível o detalhamento de um modelo para o processo e um modelo para os dados (metadados baseados no contexto OAIS). Desta forma, o projeto se tornou aplicável a qualquer biblioteca, pois estava especificado o suficiente para possibilitar desenhos implementais e o desenvolvimento de novos trabalhos (WERF-DAVELAAR, 1999). Afinal, com a adoção do OAIS se teve a utilização de um modelo referencial, terminologias comuns e um sistema conceitual comum, o que veio facilitar a troca de idéias e o compartilhamento de experiências.

O NEDLIB também considerou aspectos como custo/benefício, restrições legais, acordos com editoras, e atendimento de necessidades de usuários para acesso às informações (NEDLIB, 2000). Houve um esforço para se estabelecer estratégias de preservação, com a definição de testes de técnicas e mecanismos para preservação, como o experimento de emulação realizado por Jeff Rothenberg (2000).

Embora o projeto já tenha sido finalizado, atualizações são desenvolvidas para o DSEP e o projeto continua disponível na *Internet*.

2.6.2 e-ARQ

A Câmera Técnica de Documentos Eletrônicos do Conselho Nacional de Arquivos (CTDE/Conarq) define normas, diretrizes, procedimentos técnicos e instrumentos legais a respeito da gestão arquivística e preservação de documentos digitais (CONARQ, 2006).

Dentre outros trabalhos de orientações sobre a gestão e manutenção de acervos, a Conarq definiu um modelo de requisitos mínimos para Sistemas Informatizados de Gestão Arquivística de Documentos digitais (SIGAD), independente de plataforma, chamado e-ARQ. Este modelo tem como base o modelo OAIS, entre outros padrões.

Esse sistema se preocupa com questões ligadas à preservação. Nele são estabelecidas tabelas de temporalidade e uma tabela de tempo estimado de duração das mídias. A partir disso, o sistema é capaz de realizar um controle da vida útil dos suportes, informando quais suportes estão próximos do seu fim. O sistema preocupa-se também em garantir a confiabilidade, autenticidade e segurança dos dados armazenados, estabelecendo rotinas para verificação de erros e realizações de cópias de segurança. Outra preocupação é com o acesso às informações, para tanto o e-ARQ estabelece critérios para geração de metadados.

O e-ARQ serve como modelo de requisitos para desenvolvimento de novos sistemas, bem como para avaliar sistemas já implantados. Ele é aplicável a qualquer organização, não se restringindo a um ramo de atividade em específico. Seu acervo pode abranger documentos digitais de todos os tipos, imagens estáticas e em movimento, áudio, mensagens de correio, páginas *Web*, bases de dados, etc., além de referências a documentos não digitais. Ele também pode ser adotado como padrão ou norma pela administração pública com o objetivo de uniformizar o desenvolvimento e aquisição de sistemas que visam produzir e manter documentos arquivísticos em formato digital (CONARQ, 2006).

No entanto, para adoção do e-ARQ é necessário que uma política arquivística já esteja madura. Adequações dos requisitos às necessidades específicas de organizações também são cabíveis.

2.6.3 Projeto Portinari

O Projeto Portinari teve início em 1979, liderado pelo filho do artista Cândido Portinari, João Cândido Portinari. Dada à dispersão do acervo de Portinari, antes da formação do Projeto, suas obras eram um tanto desconhecida no Brasil. Este Projeto visa o “levantamento, a pesquisa, a organização e o acesso às informações sobre a obra, a vida e a época de Portinari [...]” (PORTINARI, 2000). Ele possui grande preocupação com a autenticidade das obras, pois, tinha como principal objetivo a formação do catálogo *raisonné*, que foi lançado em 2004.

Inicialmente, o acervo do Projeto Portinari era constituído de transparências coloridas, papel e outros elementos materiais, que começaram a sofrer degradação em decorrência do manuseio, da presença de sujidades e fatores ambientais. Foi então, na década de 80, que o Projeto se engajou na formação de seu acervo digital (PORTINARI, 2000). Cuidados com indexação, vocabulário controlado e tesauros foram tomados na formação de seu banco de dados, a partir daí deu-se o desenvolvimento da aplicação hipermídia para que o acervo fosse disponibilizado (LANZELOTTE et al., 1993).

Este projeto pode ser entendido como ligado à preservação digital por de certa forma resguardar o original através da digitalização, com a formação de banco de dados e indexações que permitem a facilidade e possibilidade de acesso ao acervo (CONWAY, 1997). Neste caso, o acervo digital vem contribuir também para que a memória não seja perdida, uma vez que algumas obras de Portinari foram destruídas, modificadas ou desaparecidas (PORTINARI, 2000). Certamente, este projeto adota uma política de preservação digital.

3 FORMAÇÃO DE ACERVO DIGITAL MULTIMÍDIA

Segundo Kenney e Rieger (2000, *tradução da autora*), “os componentes-chaves de uma política de preservação digital se dividem em quatro categorias: Infra-estrutura organizacional; política de seleção, conversão, e reseleção; estratégias de preservação; e prognóstico tecnológico”.

Desta forma, como parte da preservação digital ocorre a digitalização de arquivos analógicos, e a conversão de arquivos nato-digitais, através da seleção para adoção de estratégias que busquem a sua manutenção, como é o caso da migração.

De modo geral, a digitalização de um acervo pode vir a ocorrer para (IMAGE QUALITY WORKING GROUP OF ARCHIVESCOM, 1997), (SITTS, 2000):

- Disponibilização para consulta, inclusive remotamente, promovendo o alcance da pesquisa e da educação;
- Proteger o item original da deterioração advinda do manuseio;
- Obtenção de reproduções fiéis. Em meio digital é possível se obter reproduções idênticas de cópias para cópias de geração a geração;
- Obtenção de alta qualidade visual e sonora, possibilitando percepções até mais fáceis do que as obtidas no objeto original. Os arquivos podem ser manipulados, altas resoluções podem ser atribuídas às imagens possibilitando a visão de características imperceptíveis a olho nu, os arquivos sonoros podem ser trabalhados para retirada ou diminuição de ruídos;
- Aperfeiçoamento do acesso, através de ferramentas de buscas. Pode ocorrer economia de espaço para armazenamento e melhora no desempenho dos sistemas de informação na rede.

No entanto, alguns princípios gerais devem ser observados (KENNEY; RIEGER, 2000), (SITTS, 2000):

- Um projeto de digitalização deve ser elaborado para formação do acervo digital. Além de considerar os interesses da instituição detentora do patrimônio e o orçamento disponível, devem ser observados os propósitos para os quais o produto digital irá servir, a natureza e características do objeto original, as possibilidades de alterações tecnológicas que podem vir a ocorrer, o tempo que o produto digital deve perdurar e o impacto que o objeto original pode sofrer durante o processo de digitalização;
- Por vez, no momento da formação do acervo, alguns metadados são criados, como os provenientes do OCR³⁵, estas informações descritivas do arquivo digital também são objetos de preservação, bem como os programas e dispositivos que possibilitem sua utilização (visualização, impressão, escaneamento, etc.), principalmente em casos onde a estratégia de preservação vai além da retenção do conteúdo, como a emulação;
- Arquivos mestres devem ser salvaguardados. Estes arquivos são aqueles digitalizados em altíssima qualidade, a maior possível, em formatos em conformidade com padrões sem compressão ou compressão sem perdas. A partir destes arquivos se é possível realizar cópias sem que novas intervenções no original analógico sejam necessárias;
- No caso de migração de formato de arquivo, algumas perdas poderão ocorrer, o novo formato escolhido deverá ser o que mais minimize essas perdas;
- Tanto o processo de conversão quanto o de digitalização devem ser devidamente documentados;
- Por prudência, os originais analógicos devem ser mantidos para que possam ser redigitalizados no caso de perda catastrófica dos arquivos mestres e cópias, ou no caso da migração de formatos incorrerem em perdas de qualidade inaceitáveis.

3.1 Conversão de imagens

A digitalização de imagens é motivo de estudo há algum tempo, vários órgãos já estabeleceram guias norteadores do processo de digitalização de imagens. Uma vez que poucas mudanças têm afetado este segmento, consideraram-se principalmente o conjunto de guias do Conselho em Bibliotecas e Recursos de Informação (CLIR, 2000b), o guia

³⁵ OCR - Reconhecimento Óptico de Caracteres, do inglês *Optical Character Recognition*. Abordado no Capítulo 4.

Handbook For Digital Projects: A Management Tool for Preservation and Access (SITTS, 2000) e as recomendações técnicas para projetos de digitalização de imagens da Universidade de Columbia (IQWGAC, 1997). Assim sendo, para maior aprofundamento do tema, é aconselhável que se recorra a esses guias, que apesar da longa data da publicação continuam atuais. Naturalmente, além de recorrer a esses guias, as instituições deverão realizar testes a fim de alcançar os parâmetros que melhor atendam aos seus interesses.

Para o fim de preservação do acervo, é aconselhável a digitalização de imagens utilizando o formato *Tagged Image/Interchange File Format* - TIFF, estas imagens devem ser bem armazenadas e utilizadas como matrizes, sendo recorridas para formação de imagens derivadas em formatos comprimidos, que possibilitem o acesso rápido. O formato TIFF é o mais utilizado em projetos de digitalização na formação de imagens matrizes, pois é muito versátil, independente de plataforma, e é um formato de uso livre (KENNEY; RIEGER, 2000), (NLA, 2003). A resolução da imagem matriz dependerá do nível de detalhe que se quer obter através dela, devendo-se realizar testes para verificação da qualidade desejada.

A escolha do formato para imagens derivadas deve considerar principalmente se o formato é ou não livre e se há restrição do número de cores. De modo geral, as imagens derivadas servirão para impressão e para visualização em tela. Para tanto, elas devem ser comprimidas, o que reduzirá seus tamanhos, possibilitando seu tráfego na rede de computadores de modo ágil.

A compressão pode ser com ou sem perdas. A compressão com perdas é o processo que reduz o espaço necessário para armazenamento de um arquivo, ocasionando em alguma perda de dados - quando descomprimida, a imagem derivada diferirá da original. No entanto, a informação descartada por esses algoritmos não é facilmente percebida pelo olho humano. Porém, compressões sucessivas num mesmo arquivo permitirão a percepção da perda de informação da imagem, o que torna a compressão com perdas inadequada para arquivos que precisem sofrer vários ciclos de retoques ou alterações. Já a compressão sem perdas utiliza um algoritmo reversível para reduzir o tamanho da imagem armazenada, preservando-a sem nenhuma alteração visível ou invisível.

Os formatos para compressão de arquivos mais difundidos são JPEG³⁶, GIF³⁷, PNG³⁸ e PDF³⁹, todos utilizados largamente por editores de imagens. As diferenças e semelhanças entre eles podem ser observadas no quadro (QUADRO 3).

QUADRO 3 - Paralelo entre os formatos JPEG, GIF, PNG e PDF

JPEG	GIF	PNG	PDF
<i>Joint Photographic Experts Group</i>	<i>Graphics Interchange Format</i>	<i>Portable Network Graphics</i>	<i>Portable Document Format</i>
Padrão desenvolvido por este grupo para a compressão de imagens paradas. Ele foi sancionado pela ISO-JPEG2000-1 ⁴⁰ . Lançamento em 1983, sendo sua última versão de 1998.	Muito utilizado na Internet foi desenvolvido por CompuServer ⁴¹ . Lançado em 1987.	Sucessor do GIF. Possui suas especificações publicadas pelo W3C ⁴² , que o recomenda. Lançado em 1996.	Criado pela Adobe, o formato permite que documentos eletrônicos imprimíveis tenham exatamente a mesma aparência dos documentos originais e preservem as fontes, as imagens, os elementos gráficos e o layout de qualquer arquivo de origem, independentemente do aplicativo e da plataforma que foram usados para criá-lo.
Formato de compressão com perda.	O formato do Gif (LZW) comprime a imagem sem perda. Ele limita a quantidade de cores simultâneas possíveis. Assim, para imagens com poucas cores – uma logomarca ou um ícone, por exemplo, este pode ser o melhor formato. Ele também pode ser usado para imagens em tons de cinza.	O formato PNG é um formato de compressão sem perdas. Não possui limitação da profundidade de cores, possui alta compressão (regulável).	Comprime o arquivo.
Formato livre.	Formato patenteado.	Formato livre.	Formato livre.

³⁶ JPEG - *Joint Picture Expert Group*. Disponível em: <<http://www.jpeg.org/>>.

³⁷ GIF - *Graphics Information File*.

³⁸ PNG - *Portable Network Graphic*. Disponível em: <<http://www.w3.org/Graphics/PNG/>>.

³⁹ PDF - *Portable Document Format*. Disponível em:

<<http://www.adobe.com/br/products/acrobat/adobepdf.html>>.

⁴⁰ Disponível em: <<http://www.jpeg.org/public/wg1n1807.txt>>.

⁴¹ Disponível em: <<http://www.compuserve.com/>>.

⁴² W3C – *The Word Wide Web Consortium*. Disponível em: <<http://www.w3.org>>.

As imagens monocromáticas estão presentes em grande parte dos acervos textuais e merecem considerações. Para este caso, a informação de cor é considerada irrelevante, limitando a escolha ao escaneamento bitonal ou o multitonal (tons de cinza). Isso é particularmente aplicável à digitalização de microformas (microfilmes, microfichas), em que a informação de cor foi perdida de qualquer maneira. Essa escolha, entretanto é muito importante, e afeta profundamente os resultados do trabalho de digitalização (VALLE, 2003).

As imagens bitonais costumam ter seu uso defendido em preferência às multitonais, pois possuem um alto contraste, tinta escura sob fundo claro. Entretanto, a imagem multitonal preserva melhor a informação, e em muitos casos, a torna mais legível, pois o texto pode ter variações de contraste, borrados e manchas irregulares. Imagens multitonais podem sofrer ajustes de brilho e contraste, o que facilita a leitura de páginas com variações de contraste original. Dessa forma, as imagens bitonais são especialmente recomendadas para uso de imagens de documentos regulares: tipografados ou impressos, com boa legibilidade, sem grande degradação e em que o contraste entre o texto e o fundo seja elevado e constante. As imagens multitonais precisam ser de apenas 256 tons de cinza, devido à capacidade de percepção do olho humano (GONZALEZ; WOODS, 1992).

Outro aspecto que deve ser considerado é a resolução. A resolução se refere tipicamente ao número de pixels horizontais e verticais que formam a imagem. Por exemplo, 512x768 é o mesmo que 512 pixels de lado por 768 na horizontal. DPI (*dots per inch*) são os pontos ou pixels capturados por polegada do material de origem. As resoluções maiores dão maior qualidade à imagem digital, porém o tamanho também aumenta. Sendo assim, dados os custos de armazenamento, procura-se utilizar a menor resolução que atenda com segurança os propósitos da imagem digital (VALLE, 2003). As resoluções das imagens derivadas devem ser adequadas ao seu propósito, desta forma, de modo geral, para impressão recomenda-se a resolução de 600 *dpi* para imagens bitonais, 300 *dpi* para imagens coloridas e multitonais. É recomendável fazer um teste para verificar se as resoluções aqui indicadas atenderão bem ao propósito, afinal, resoluções maiores podem ser necessárias para fotos com pequenos detalhes significantes. Já para visualização de imagens em tela e para acesso 72 *dpi* são suficientes para um bom resultado, colaborando para leveza do arquivo e agilidade de seu transporte na rede (QUADRO 4).

QUADRO 4 - Recomendações da Universidade de Columbia

Tipo de mídia	Método de conversão	Resolução	Formato de armazenamento do arquivo	Formato de apresentação na tela	Formato de impressão
Texto em branco e preto	Escaner de mesa ou Câmera Digital	1-bit, 600 dpi	TIFF w/CCITT Fax 4 Compressão	GIF, 4-bit, 120 to 200 dpi	Acrobat (PDF), 1-bit, 300 or 600 dpi
Ilustrações, Mapas, Manuscritos, etc.	Escaner de mesa ou Câmera Digital	8-bit escala de cinza ou 24-bit de cores, 200 a 300 dpi	TIFF	Múltiplo JPEG, 24-bit, 512x768, 1024x1536, 2048x3072, Nível de Qualidade 50-100	JPEG, 24-bit, 2048x3072, Nível de Qualidade 50-100
Objetos com 3-dimensões para serem representados em 2 dimensões	Câmera Digital	24-bit de cores, 200 a 300 dpi	TIFF	Múltiplo JPEG, 24-bit, 512x768, 1024x1536, 2048x3072, Nível de Qualidade 50	JPEG, 24-bit, 2048x3072, Nível de Qualidade 50-100
35mm Slide Preto & Branco & Colorido ou Negativo	PhotoCD ou Escaner de Slide	24-bit, 2048x3072	PhotoCD ou TIFF	Múltiplo JPEG, 24-bit, 512x768, 1024x1536, 2048x3072, Nível de Qualidade 50	JPEG, 24-bit, 2048x3072, Nível de Qualidade 50-100
Formato fotográfico de médio a grande, slide, negativo, transparência ou microficha colorida	ProPhotoCD ou Escaner de tambor	24-bit, 4096x6144	PhotoCD ou TIFF	Múltiplo JPEG, 24-bit, Nível de Qualidade 50	JPEG, 24-bit, 4096x6144, Nível de Qualidade 50-100
Preto & Branco Microfilme	Escaner de Microfilme	1-bit 600 dpi	TIFF w/ Fax 4	GIF, 4-bit, 120 to 200 dpi	PDF, 1-bit, 300 ou 600 dpi
		8-bit, 300 dpi	TIFF	GIF, 8-bit 120 to 200 dpi	PDF, 8-bit, 300 ou 600 dpi

Fonte: IMAGE, 1997.

Um cuidado importante no momento da digitalização é a realização de um gerenciamento de cores com a adoção de escalas-padrão para cores e para tons de cinza, a fim de assegurar a máxima fidelidade cromática das reproduções e o estabelecimento de um controle de qualidade. Cuidados com a calibração de monitores e impressoras também devem ser tomados, no mercado existem programas desenvolvidos com essa finalidade (KENNEY; RIEGER, 2000), (LEÃO, 2005).

3.2 Conversão de áudio

A importância da preservação do patrimônio audível tem sido reconhecida por instituições e nações, principalmente devido ao seu valor cultural. Estes acervos são formados por discos (de acetato ou vinil, por exemplo) e fitas (k7, micro-cassette, mini-cassette, por exemplo).

A gravação de áudio teve início em 1888 com Thomas Edison e sua comercialização se iniciou em 1894. Apesar da longa data, ainda é possível a extração do áudio dos gravadores analógicos para os digitais, porém, com o passar do tempo, essa possibilidade decaí. As digitalizações dessas mídias exigem programas e dispositivos específicos. Porém, para obtenção de uma ótima qualidade, um ouvido treinado também se faz essencial, pois no caso do áudio ele é o modo de se ter um controle de qualidade, o que naturalmente torna este processo muito subjetivo (COLORADO DIGITIZATION PROJETCT, 2005), (CLIR, 2006).

Após a verificação do estado de conservação da fita ou do disco e certificação de que o processo de digitalização não causará danos, o primeiro cuidado a ser tomado para digitalização é a limpeza (CDP, 2005), (CLIR, 2006). No caso dos discos, a limpeza dependerá da composição química do disco, podendo ser realizado através de métodos não abrasivos, como com um aspirador de pó, pano macio, espanamento delicado ou escovação anti-estática (CLIR, 2006). Para limpeza das fitas, observar o Capítulo 2.

O equipamento para reprodução do áudio deve ser selecionado com critério, tendo-se atenção, no caso dos discos, principalmente com o tamanho da agulha, o que irá determinar uma transmissão sonora mais pura. A agulha deve pousar levemente sobre o disco gerando um

sinal fiel e minimizando ruídos. A velocidade da transmissão sonora deve ser escolhida de acordo com o tipo de disco, ela pode variar de 33-1/3 rpm⁴³ a 78 rpm, sendo a última mais comum. Já em se tratando de fitas, a escolha do equipamento pode ser facilitada identificando-se o tipo de fita, essa informação pode estar disponível na própria caixa da fita. O ângulo de fundo da cabeça de leitura deve ser ajustado, a cabeça de leitura também deve ser limpa. A velocidade da leitura pode ser iniciada experimentalmente a 7-1/2 polegadas por segundo e corrigida se necessário. Cuidados especiais devem ser tomados com as fitas e discos que já sofreram deterioração, ainda que possível a extração do áudio de fitas e discos danificados, a ajuda de um profissional é recomendável (CLIR, 2006).

Os equipamentos para conversão basicamente incluem, além do equipamento analógico para reprodução de áudio, mencionado acima, um computador, programas de computador especializados e um conversor de sinal analógico para digital.

Os arquivos matrizes do som digitalizado devem utilizar um formato de arquivo de áudio sem compressão, como o WAV⁴⁴. Esse formato é utilizado nos CDs de música, onde é mascarado pela extensão “cda⁴⁵”. Ele é o formato padrão utilizado pela Microsoft e pela IBM para armazenamento de áudio em computadores e é também compatível com todos os aparelhos de som do mercado.

Para fins de acesso, o arquivo de áudio deve sofrer uma compressão para que seu tráfego na rede seja possibilitado e um desempenho melhor seja alcançado nas respostas de buscas. A compressão pode ser com ou sem perdas, sendo alcançado um tamanho de arquivo menor na compressão com perdas. Esta baseia-se nos modelos psico-acústicos, que retiram do arquivo sonoro informações que o ouvido humano não conseguiria perceber devido a fenômenos de mascaramento de sons e limitações da audição humana.

Os formatos mais comuns com compressão são MP3⁴⁶, WMA⁴⁷, Ogg Vorbis⁴⁸, Monkey’s Audio⁴⁹ e o FLAC⁵⁰. Dentre eles, o de maior destaque é o MP3 por sua alta taxa de

⁴³ RPM – Rotações por minuto

⁴⁴ WAV ou WAVE, formato de arquivo de áudio padrão adotado pela Microsoft e IBM para armazenamento de áudio em computadores.

⁴⁵ CDA – Disco Compacto de Áudio, do inglês *Compact Disc Audio*.

⁴⁶ MP3 – MPEG- ½ Audio Layer 3. Disponível em: <<http://www.mpeg.org/MPEG/audio.html>>.

⁴⁷ WMA – Windows Media Audio. Formato de áudio da Microsoft. Disponível em: <<http://www.microsoft.com/windows/windowsmedia/forpros/codecs/audio.aspx>>.

compressão e popularidade, porém após a requisição de permissão para sua utilização, alguns passaram a não o preferirem mais e a criação de novos formatos foi fomentada. Para formação do acervo digital é importante se considerar a utilização de formatos não proprietários com alto potencial de leitura, tendo atenção a prospecção tecnológica (QUADRO 5).

QUADRO 5 - Comparação de formatos de arquivos digitais de áudio

MP3 (MPEG-1/2 Audio Layer 3)	WMA (Windows Media Audio)	Ogg Vorbis	Monkey's Audio	FLAC (Free Lossless Audio Codec)
Patente controlada pela Thomson Consumer Eletronics..	Criado pela Microsoft.	Desenvolvido pela Xiph.Org Foundation.	Desenvolvido por Matthew T. Ashland.	Formato de código aberto.
Formato campeão de popularidade, vem da década de 90. Possibilita a transmissão por <i>streaming</i> ⁵¹ . Codificação de <i>bit-rate</i> estável.	Suportado por diversos tocadores portáteis. Possui opção de compactação especial para voz. Possibilita a transmissão por <i>streaming</i> . Tem a opção de proteção aos direitos autorais (DRM).	Formato isento de patentes, criado em resposta às cobranças de direitos do MP3. Codificação de <i>bit-rate</i> variável (VBR). Possibilita a transmissão por <i>streaming</i> . Formato ainda em uso minoritário, por ser recente – 2002.	Compressão sem perda de dados e de qualidade. O arquivo descompactado é igual ao original.	Compressão sem perdas, reflete com perfeição o conteúdo original, tornando-se idêntico ao original ao sofrer descompactação. Formato livre e com ferramentas de código aberto.
11:1 do WAV (em 128 Kbps); 21.7:1 (em 64 Kbps)	10.9:1 do WAV (em 128 Kbps); 22:1 (em 64 Kbps)	Seus arquivos são um pouco menores que o MP3 e possuem qualidade semelhante.	5:1 do WAV	Semelhante ao tamanho de arquivo conseguido pelo Monkey's Audio.

⁴⁸ Disponível em: <<http://xiph.org/>>.

⁴⁹ Disponível em: <<http://www.monkeysaudio.com/>>.

⁵⁰ FLAC – Free Lossless Audio Codec. Disponível em: <<http://flac.sourceforge.net/>>.

⁵¹ *Streaming*: Tecnologia que permite o envio de informação multimídia através de pacotes, utilizando redes de computadores, sobretudo a *Internet*.

Um cuidado crítico na conversão de áudio é a escolha da taxa de amostragem e da profundidade de *bit*⁵².

Como padrão para preservação, a taxa de amostragem de 96 KHZ foi aceita, devendo ser usada para os arquivos mestres. A taxa de 44.1 KHZ pode ser preferida para arquivos voltados para o acesso (CDP, 2005), (CLIR, 2006).

Já para a escolha da profundidade de *bits* (*bit depth*) é recomendável se assumir o padrão profissional de 24-*bits* de profundidade. Este valor porém é alto e não costuma ser tocado em todos os dispositivos de áudio, portanto, para os arquivos com finalidade de acesso pode-se adotar 16-*bits* de profundidade (CDP, 2005).

As observações sobre taxa de amostragem e profundidade de *bits* feitas acima também são válidas para os arquivos nato-digitais (CDP, 2005). O quadro (QUADRO 6) representa o tamanho do arquivo para uma hora de áudio digitalizado em diferentes taxas de amostragem e profundidade de *bits* ou resolução de *bits*:

QUADRO 6 - Tamanhos de arquivos para uma hora de áudio

Taxa de amostragem	Profundidade de <i>bit</i>	No. de canais	Tamanho do arquivo: Megabytes	Tamanho do arquivo: Gigabytes
44.1 kHz	16	2 (stereo)	605.6 MB	0.61 GB
44.1 kHz	16	1 (mono)	302.8 MB	0.30 GB
44.1 kHz	24	2 (stereo)	908.4 MB	0.91 GB
44.1 kHz	24	1 (mono)	454.2 MB	0.45 GB
96 kHz	24	2 (stereo)	1977.5 MB	1.98 GB
96 kHz	24	1 (mono)	988.8 MB	0.99 GB

Fonte:CDP, 2005.

⁵² Profundidade de bit: Em um áudio digital, a profundidade de *bit* descreve o número de degraus ou níveis usados para representar cada amplitude medida numa onda de som. Em geral, quanto mais *bits* houver, mais acurados serão os resultados do processamento dos dados.

4 METADADOS

O conceito de preservação envolve a aquisição, organização e distribuição de recursos a fim de impedir posterior deterioração ou renovar a possibilidade de utilização de um item. Desta forma, a ação de preservação é aplicada a um item com o objetivo de torná-lo acessível para uso democrático (CONWAY, 1997).

Metadados são informações estruturadas sobre ou representativas de um documento ou obra em geral, são dados sobre dados, usados no contexto das relações entre indivíduos e recursos. Os metadados podem denotar informação legível por máquina, e, para bibliotecas, ele pode significar qualquer forma de descrição de recurso aplicado a qualquer tipo de objeto, digital ou não⁵³ (KENNEY; RIEGER, 2000), (RODRIGUES, 2003).

O emprego de metadados vem possibilitar o acesso, o emprego de estratégias de preservação como migração e emulação, o estabelecimento de autenticidade aos conteúdos digitais e a obtenção de facilidades de gestão e uso de informações de um acervo.

Para possibilitar a facilidade de encontro de documentos e informações, existem os índices, que podem ser atribuídos através de programas de computador de indexação automática ou por intervenção humana. “A indexação é estabelecida através de metadados” (RODRIGUES, 2003). Além da indexação, existem os vocabulários controlados e tesouros. Através da indexação, termos de pesquisa (palavras chaves controladas ou não e códigos de classificação) são atribuídos ao acervo. Como avanço à indexação, está o emprego de vocabulários controlados, onde a escolha de indexadores se preocupa em evitar a polissemia e a sinonímia. Porém os termos num vocabulário controlado não estão relacionados entre si o que provoca a dissociação semântica. Tal relacionamento semântico entre termos é obtido através dos tesouros, que auxiliam na identificação de diferenças mínimas entre as palavras, através de uma lista de palavras com significados semelhantes e não necessariamente sinônimos (VALLE, 2003). Padrões de metadados comumente necessitam ou recomendam o uso de vocabulários controlados e tesouros para popular os esquemas de metadados, o que vem

⁵³ A catalogação MARC – catalogação legível por máquina é um padrão de esquema de metadados para bibliotecas. MARC, do inglês, *Machine Readable Cataloging*.

possibilitar a pertinência e relevância das respostas de buscas de informações (KENNEY; RIEGER, 2000), (BESSER, 2003).

“Metadados são capturados na forma de lista de elementos prescritos, ou campos, conhecidos como esquema de metadados” (BESSER, 2003, *tradução da autora*). Desta forma, esquemas de metadados podem ser empregados, entre outros objetivos para: facilitar o acesso às informações, para assegurar as informações das instituições, para promover a preservação e para gestão administrativa dos próprios metadados.

Uma vez que os metadados possibilitam o acesso através de bibliotecas digitais e da *Internet*, durante sua criação, cuidados com a proteção das coleções digitais não podem ser perdidos de vista. É preciso haver um equilíbrio entre a disponibilização de informações e a descrição de políticas complexas que definem direitos, restrições, e regras que ditam quem poderá fazer o que com os recursos digitais. O problema do gerenciamento de direitos e controle de acesso demanda um modelo de metadados que permita a descrição estruturada de políticas complexas, que facilite a aplicação dessas políticas em ambientes automatizados e promovam a interoperabilidade entre diferentes comunidades e domínios de aplicação (KENNEY; RIEGER, 2000).

Por vez, os dados associados a funções administrativas dos recursos digitais, e suas garantias de preservação são freqüentemente chamados de metadados para preservação ou metadados técnicos e administrativos. Esses são os metadados necessários para preservar a integridade e funcionalidade dos objetos digitais. Por esse motivo, esses metadados também são objeto de preservação assim como os recursos digitais que eles suportam (KENNEY; RIEGER, 2000), (NLA, 2003).

Os metadados para preservação englobam informações tais como os atributos do documento digital, o processo tecnológico para sua criação, dados sobre o fluxo de trabalho para criá-lo, quem o criou (quem detêm os direitos) e informação sobre a procedência e o contexto original, por exemplo (KENNEY; RIEGER, 2000), (NLA, 2003). A criação destes metadados pode considerar tanto uma preservação de curto prazo quanto a de longo.

Principalmente com a *Web*, diversas iniciativas foram desenvolvidas para estabelecimento de normas para criação de metadados a fim de facilitar a troca de informações entre instituições.

Sendo este um dos pontos fortes do OAIS, que é um modelo conceitual para metadados de preservação, e do NEDLIB, que o utiliza. Outros esquemas de metadados compartilhados são o Dublin Core⁵⁴, CDWA⁵⁵ e o NISO⁵⁶ *Data Dictionary*⁵⁷. Estas iniciativas alcançam êxito ao promover a interoperabilidade semântica, estrutural e sintática. Ou seja, as instituições passam a compartilhar do mesmo esquema de categorização e interpretação dessas categorias; assumem a mesma forma de estruturar a informação, adotando definições estritas para valores que possam ter múltiplas interpretações⁵⁸; e passam a adotar uma linguagem comum para expressar os metadados como *Hyper Text Markup Language* (HTML)⁵⁹, *Resource Description Framework* (RDF)⁶⁰, *Standard Generalized Markup Language* (SGML)⁶¹, *eXtensible Markup Language* (XML)⁶² e *eXtensible Hypertext Markup Language* (XHTML)⁶³ (KENNEY; RIEGER, 2000), (BESSER, 2003).

Algumas orientações mais específicas podem ser dadas sobre a criação de metadados para arquivos de imagens e som, como explanado a seguir.

4.1 Metadados para imagens

Boa parte de um acervo digital é formado por imagens, além de documentar a proveniência, as características da imagem e informações para preservação, os metadados são substanciais na promoção do acesso.

⁵⁴ Dublin Core define elementos semânticos para categorizar recursos na Web a fim de facilitar na recuperação de buscas. Disponível em: <<http://dublincore.org/>>.

⁵⁵ CDWA – *Categories for the Description of Works of Art*, do inglês, Categorias para descrição de trabalhos artísticos. Disponível em: <http://www.getty.edu/research/conducting_research/standards/cdwa/>.

⁵⁶ NISO - *National Information Standards Organization*, do inglês, Organização Nacional de Padrões Informacionais.

⁵⁷ NISO *Data Dictionary: Technical metadata for Digital Still Images*, do inglês, Dicionários de dados NISO: Metadados técnicos para imagens estáticas. Disponível em: <<http://www.niso.org/>>.

⁵⁸ Como exemplo, datas são lidas de formas diferentes em alguns países.

⁵⁹ HTML, do inglês, Linguagem para marcação de texto.

⁶⁰ RDF, do inglês, Sistema de descrição de recursos.

⁶¹ SGML, do inglês Padrão generalizado para linguagem de marcação.

⁶² XML, do inglês, Linguagem de marcação extensível.

⁶³ XHTML, do inglês, Linguagem de marcação extensível para hipertexto.

Para o caso das imagens, a atribuição de metadados pode ocorrer de forma mais automática, pois, existem tecnologias que auxiliam no processo de geração de metadados no momento da digitalização de páginas textuais, possibilitando a extração de informações textuais em forma de caracteres e fazendo com que o conteúdo da imagem resultante da digitalização seja pesquisável por ferramentas de busca (BOERES, 2005). Estas tecnologias de conversão de texto em imagens para texto codificado em caracteres, são o OCR – reconhecimento óptico de caracteres e o ICR – reconhecimento inteligente de caracteres⁶⁴. O OCR faz o reconhecimento das formas das letras a partir da imagem digitalizada. Já o ICR vai além, possibilitando o reconhecimento inclusive de textos problemáticos, com leituras difíceis como os impressos matriciais, tipografias antigas e manuscritos, onde a imagem do original não é tão legível, uma vez que o original já sofreu ações do tempo, por exemplo. Estas tecnologias ampliam as possibilidades de acesso, com a criação de metadados, e possibilitam que os usuários consultem o conteúdo do texto juntamente com a imagem do original de forma complementar (KENNEY; RIEGER, 2000).

O reconhecimento óptico de caracteres não acarreta custos tão onerosos, porém, não alcança a acurácia de um digitador humano, cabendo à instituição avaliar a necessidade de precisão da extração textual e o custo/benefício de seu emprego. Naturalmente, a acurácia obtida pelo OCR dependerá de fatores como o tamanho da fonte, o estado de conservação e a qualidade do objeto original. A acurácia é medida pela porcentagem de caracteres que são perfeitamente transcritos no texto resultante do processamento do OCR ou ICR. O maior nível de acurácia para textos obtidos de OCR, que sofreram intervenções de correção humana, comumente fica entre 99,995%, ou nada mais que um erro para 20 páginas impressas. Em certos casos, a digitação do conteúdo textual pode vir então a ser preferível, devido ao custo de um processo automatizado, acrescido de intervenção humana para correção. Há também, situações, em que o valor do arquivo não está significativamente em seu conteúdo textual, o que dispensaria a necessidade de transcrição do conteúdo para fins de leitura do usuário e onde uma avaliação da necessidade de extração de metadados por OCR seria necessária (KENNEY; RIEGER, 2000).

Alguns formatos de arquivos de imagens permitem a anotação de metadados textuais diretamente em seu cabeçalho, como é o caso do GIF, JPEG, PNG e do TIFF. No entanto,

⁶⁴ Do inglês, *Intelligent Character Recognition*.

estes recursos não são tão flexíveis e devem ser usados com parcimônia para evitar possíveis conflitos, pois nem todos os aplicativos dão suporte à leitura destes metadados (VALLE, 2003), (KENNEY; RIEGER, 2000).

Outro recurso sofisticado para extração de informações de imagens é a aplicação de um processamento para recuperação de imagens com base no seu conteúdo visual - CBIR⁶⁵. Este processo vai além da extração textual e possibilita o enriquecimento dos metadados, acrescentando maiores recursos às buscas. Através do CBIR, as imagens podem ser recuperadas tendo como critérios a cor, a forma, uma imagem exemplo ou a posição de elementos contidos na imagem (KENNEY; RIEGER, 2000), (BESSER, 2003).

4.2 Metadados para arquivos sonoros

Os metadados em arquivos sonoros são essenciais não apenas por possibilitar a sua recuperação através de ferramentas de busca, mas principalmente para preservação.

Metadados sucintos podem ser postos no cabeçalho de arquivos como WAV e MP3, com informações como o título do áudio, ano, direitos de cópia, tempo de duração e localização do original no arquivo. Entretanto, outros metadados devem ser criados separados do arquivo digital com informações que possibilitem sua futura migração ou transcrição, como informações técnicas sobre a mídia digitalizada, sua natureza, seu o processo de digitalização, entre outras. Para inserção de novos metadados, é possível se fazer uso de ferramentas auxiliares, como o METS⁶⁶, utilizado pela Biblioteca do Congresso Americano - *The Library of Congress* (LC) no seu Protótipo de Projeto de Preservação de Gravações Áudio-Visuais⁶⁷ (CDP, 2006), (CLIR, 2006).

⁶⁵ CBIR – *Content-Based Image Retrieval*.

⁶⁶ METS- *Metadata Encoding and Transmission Standard*, do inglês, Codificação de Metadados e Padrão de Transmissão. Disponível em: <<http://www.loc.gov/standards/mets/>>.

⁶⁷ LC *Audio-Visual Prototyping Project*. Disponível em: <<http://www.loc.gov/rr/mopic/avprot/avlcdocs.html>>.

São exemplos de instituições preocupadas com o desenvolvimento de esquemas padrões de metadados para áudio: o *Audio Engineering Society/European Broadcasting Union Standard Development*⁶⁸ e a Biblioteca do Congresso Americano.

⁶⁸ Sociedade de Engenharia de Áudio/ Desenvolvimento de Padrão de Transmissão da União Européia. Disponível em: <http://www.aes.org/standards/b_reports/b_meeting-reports/aes115-sc-06-06-report.cfm>.

5 DESENVOLVIMENTO DE SISTEMAS ORIENTADOS À PRESERVAÇÃO

O gerenciamento do acervo digital inclui a geração, organização e indexação, armazenagem, transmissão e uma contínua manutenção da integridade intelectual. Esse contexto envolve o uso de sistemas para promoção da gestão documental e do acesso (CONWAY, 1997).

Estes sistemas são essenciais para um projeto de preservação digital, devendo abarcar uma infra-estrutura tecnológica apropriada. De modo geral, eles objetivam (KENNEY; RIEGER, 2000), (NLA, 2003):

- Armazenar arquivos mestres e derivados, em bancos de dados. Muitas vezes, apenas uma referência à localização do arquivo fica armazenada no banco de dados, permitindo que as consultas e processamentos sejam mais dinâmicos. Há porém, casos, em que os arquivos digitais são guardados diretamente no banco de dados, como objetos binários, o que exige maior robustez do sistema gerenciador de banco de dados;
- Disponibilizar meios para armazenamento e relacionamento de metadados com os objetos digitais;
- Permitir a busca e recuperação automática dos arquivos digitais;
- Prover uma interface para acesso, freqüentemente via *Web*;
- Criar uma trilha de auditoria para rastrear alterações realizadas no acervo e em seus metadados.

Entretanto, não existe uma solução perfeita que possa ser utilizada para todo e qualquer projeto de preservação. Estes sistemas precisam ser especializados ao domínio de sua aplicação se adequando às especificidades da natureza do acervo, ao seu propósito e às políticas institucionais. Sendo freqüente, a combinação de sistemas, como por exemplo, sistemas para manipulação do acervo (digitalização, edição); para gestão documental, de uso restrito, e para promoção do acesso público.

Para desenvolvimento de um sistema adequado a cada instituição, é de rico valor a aplicação de um processo de engenharia de sistemas. E dentro desse, o emprego da engenharia de

requisitos para esclarecimento do que é necessário ser desenvolvido para melhor abraçar os propósitos do sistema. Pode-se, também, fazer uso de uma documentação de requisitos, especificada por instituições orientadoras e promotoras da preservação, como o projeto e-Arq da Conarq⁶⁹, tomando-se o cuidado de realizar as devidas adaptações necessárias. Neste capítulo, esclarecimentos sobre o desenvolvimento de sistemas são apresentados.

5.1 Gestão documental

Um sistema de informação é uma coleção de componentes inter-relacionados (maquinário, programas, procedimentos, pessoas, bancos de dados) que trabalham juntos para “coletar ou resgatar, processar, armazenar e distribuir informação” (LAUDON; LAUDON, 2003).

Dentro do contexto dos sistemas de informação, estão aqueles voltados para gestão documental. De acordo com o dicionário de Terminologia Arquivística, do Conselho Internacional de Arquivos, a gestão de documentos diz respeito a uma área da administração geral relacionada com a busca de economia e eficácia na produção, manutenção, uso e destinação final dos documentos (EVANS et al, 1984).

A norma ISO 15489:2001 orienta os arquivistas nas iniciativas básicas que devem ser empreendidas no desenvolvimento e reestruturação de sistemas de arquivos e particularmente de gestão documental. Há de se considerar, entretanto, que a gestão documental não se refere apenas a núcleos documentais dinâmicos, onde a rotatividade e o tempo de conservação dos documentos nos depósitos de banco de dados são determinantes para um uso apropriado de recursos (financeiros, de armazenamento em máquina, por exemplo). Ou seja, os sistemas de gestão documental, naturalmente, devem atender às necessidades da instituição a que se presta. Desse modo, em sua maioria são voltados para um domínio específico.

Para o cumprimento de um programa de gestão documental, a norma define os seguintes elementos básicos:

⁶⁹ Abordado no Capítulo 2.6.1.

- Definição de políticas e responsabilidades;
- Compreensão dos princípios da gestão de documentos;
- Concepção e implementação do sistema de arquivo, considerando suas características e metodologia de desenvolvimento;
- Definição dos processos de suporte ao sistema de arquivo, orientados ao controle documental desde sua concepção até o fim da vida útil;
- Monitorização do sistema;
- Formação.

A norma, entretanto, não traz considerações de como aplicar suas orientações, não esclarecendo sobre procedimentos de desenvolvimento destes tipos de sistemas.

5.2 Processo de desenvolvimento de sistemas

Fazer um sistema de informação envolve desde a percepção das necessidades levantadas pelo cliente até testes e migrações. Depois de enunciado o problema, é feito um projeto, contendo a proposta, a equipe para fazer o sistema, o orçamento e cronogramas. Dessa forma, um projeto representa a execução de um processo.

Segundo Paula (2000), processo é um conjunto de passos parcialmente ordenados construídos por atividades, métodos, práticas e transformações. Ele é usado para se atingir uma meta, que geralmente está associada a um ou mais resultados concretos ou finais (resultados da execução do processo).

Rocha (2001) ressalta que o processo de desenvolvimento de sistemas envolve métodos, ferramentas e pessoas. Para um processo funcionar satisfatoriamente (FIG. 4), ele deve possuir:

- Procedimentos e métodos que descrevam a relação entre as tarefas;
- Ferramentas e equipamentos que dêem suporte à realização das tarefas, simplificando e automatizando o trabalho;

- Pessoal com perfil adequado, treinado nos métodos e com as ferramentas para poderem realizar as atividades adequadamente.

Esse conjunto deve estar integrado harmoniosamente para funcionar de forma eficaz.

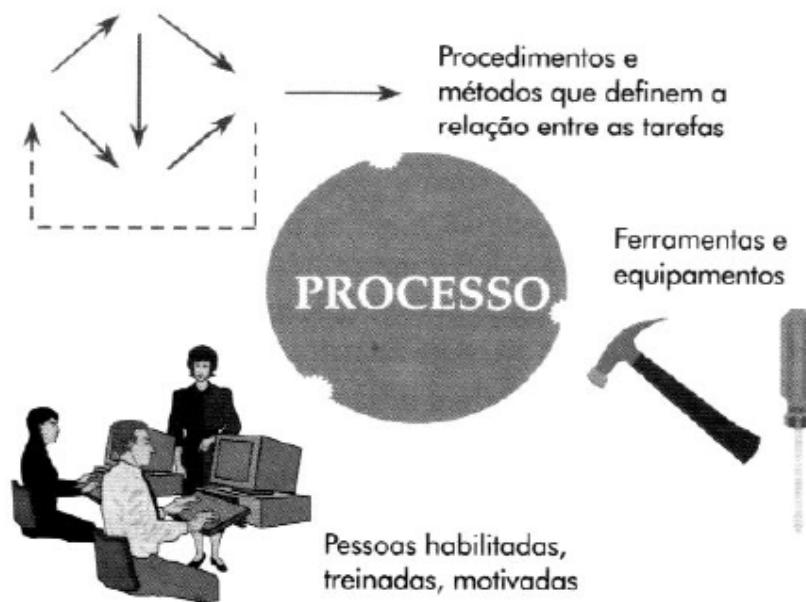


FIGURA 4 - Processo de desenvolvimento de sistemas.
Fonte: ROCHA, 2001.

Existem vários processos, a escolha de qual deles seguir deve estar relacionada à postura institucional, aos conhecimentos prévios, cultura e formação do pessoal, às ferramentas já familiarizadas, além do tamanho da equipe e da sua dedicação.

De acordo com Hix (1993), dentro de um processo é possível perceber que existem dois domínios: o construcional e o comportamental. Sendo o primeiro preocupado com a construção da interface e da aplicação, e o segundo, com o desenvolvimento da interação homem-máquina. Ambos os domínios produzem partes de um sistema interativo. No entanto, no desenvolvimento da interação estão os fatores humanos, determinantes da necessidade de metodologia específica. Os mesmos conceitos como planejamento, especificação, métricas, documentação e avaliação se aplicam, porém, de forma diferente.

Hix (1993) sugere o modelo de ciclo de vida em estrela apresentado na ilustração (FIG. 5) para o desenvolvimento de interfaces, considerando a necessidade de interação do usuário

durante todo o processo de desenvolvimento das mesmas. Este ciclo de vida não tem seus passos ordenados ou conectados em seqüência, o que permite ao desenvolvedor iniciar o processo pelo ponto mais conveniente.

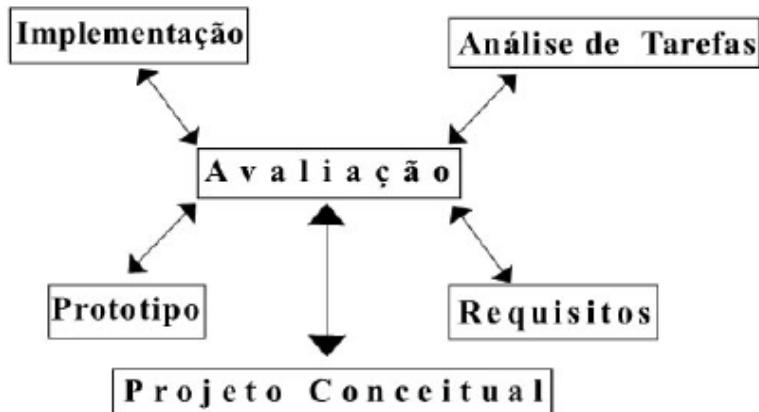


FIGURA 5 - Ciclo de vida para desenvolvimento de interfaces em estrela.

Fonte: HIX, 1993.

5.3 Engenharia de requisitos

Dentro do desenvolvimento de sistemas a engenharia de requisitos é a primeira e principal parte para obtenção de resultados satisfatórios. Ela consiste no entendimento de qual será o impacto do sistema sobre o negócio, do que a instituição cliente almeja e como os usuários finais irão interagir com o produto final desenvolvido. Portanto, a engenharia de requisitos é o elo entre um projeto e sua construção (PRESSMAN, 2006).

O processo de engenharia de requisitos abrange sete funções: Concepção, Levantamento, Elaboração, Negociação, Especificação, Validação e Gestão. Desta forma, tem-se um entendimento básico do problema, que é mais detalhado através de reuniões, possibilitando uma expansão, um refinamento do entendimento das necessidades, onde se pode definir o domínio do problema informacional, funcional e comportamental e a descrição das funções e restrições que estarão atreladas ao sistema. Após esses passos, os requisitos são verificados e as alterações eventuais são geridas (PRESSMAN, 2006). As atividades para alcance dessas funções são elucidadas no fluxo de Atividades de Requisitos conforme ilustração (FIG. 6).

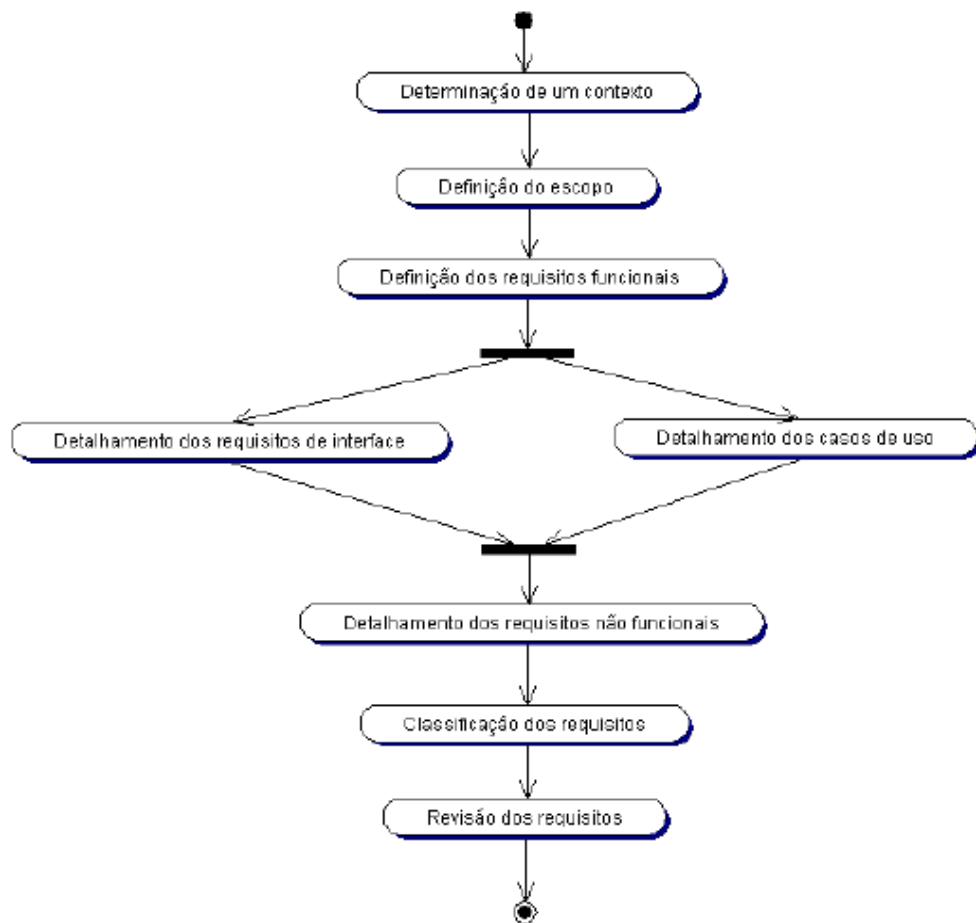


FIGURA 6 - Atividade de fluxo de requisitos.
Fonte: ÁLVARES, 2001.

5.4 Engenharia da Web

A *Web* é o maior meio para encontro de informações e por isso, o meio mais adequado para disponibilização de acesso ao acervo de uma instituição. Esse meio, no entanto, apresenta características específicas que repercutem no desenvolvimento de sistemas (PRESSMAN, 2006), (NIELSEN, 1993), (ÁLVARES, 2001):

- Maior importância de emprego de usabilidade, dado o maior número de usuários e diversidade entre eles;
- Desenvolvimento ágil de sistemas em ciclo de vida curtos;

- Desenvolvimento evolutivo ou incremental;
- Ênfase na estética;
- Voltadas para o conteúdo informacional;
- Ocorrência de modificações freqüentes.

Segundo a norma ISO 9241 (1998), usabilidade é a “capacidade que um sistema interativo oferece a seu usuário, em um determinado contexto de operação, para a realização de tarefas de maneira eficaz, eficiente e agradável”. Já segundo a norma ISO/IEC 9126 (1991), usabilidade é a facilidade com que um usuário pode aprender a operar, preparar entradas para interpretar as saídas de um sistema ou componente. Desta forma, a usabilidade trata da qualidade de interação homem-máquina através da adequação do contexto de utilização dos sistemas aos tipos de usuários envolvidos.

Dada todas essas características, o tempo empregado no levantamento de requisitos é muito precioso. Assim, essa fase do processo deve ser bastante detalhada, focando-se as formas de interação do usuário com o sistema e os testes.

Numa aplicação voltada para *Web*, os objetivos da coleta de requisitos passam a ser mais abrangentes, devendo-se identificar os requisitos funcionais e de conteúdo, e definir cenários de interação para as diversas classes de usuários. Essas classes, entretanto, devem abranger o menor conjunto possível, pois seu número é proporcional ao nível de complexidade do sistema (PRESSMAN, 2006).

Quatro diferentes tipos de análise são conduzidos durante a Engenharia da *Web* (PRESSMAN, 2006), (ÁLVARES, 2001):

- Conteúdo: Identifica o conteúdo das páginas *Web*, imagens, texto, vídeo e áudio;
- Interação: Descreve o modo de interação do usuário com a interface, definindo modelos navegacionais e formas de interação com as funcionalidades;
- Função: Define funções de processamento que poderão ser realizadas no sistema *Web*, como as buscas;
- Configuração: Descrição do ambiente e infra-estrutura necessária para hospedagem do sistema.

As realizações dessas análises culminam, portanto, em outro aspecto da Engenharia da *Web*, a maior interação entre desenvolvedores e clientes.

6 PRESERVAÇÃO DIGITAL NA PESQUISA GUIGNARD

Como mencionado no Capítulo 1, a Pesquisa Guignard tem como principal propósito caracterizar a obra de Alberto da Veiga Guignard (1896-1962), relacionando aspectos históricos, estéticos e técnicos. Pelo estudo sistemático das obras registradas, o projeto visa identificar os materiais utilizados por Guignard, a fim de formar um banco de dados referencial sobre a técnica e a maneira de execução das pinturas e desenhos, realizar também, uma revisão bibliográfica do material já publicado sobre o artista e sua obra e uma coleta de depoimentos orais com amigos, ex-alunos e críticos de arte, através da metodologia da História Oral. Este estudo inovador no país forma um acervo que pode ser considerado parte do patrimônio cultural de Minas Gerais e do Brasil.

A Pesquisa Guignard tem gerado um rico acervo digital dos objetos artísticos estudados, portanto, ela se enquadra no cenário de projetos e iniciativas que se preocupam com a preservação digital.

Este capítulo descreve um estudo de caso, explanando como a preservação digital tem ocorrido dentro dessa Pesquisa, descrevendo a formação do seu acervo, os cuidados com os suportes de armazenamento de informações, o respeito aos direitos e o desenvolvimento de sistemas computacionais.

6.1 Seleção para disponibilização de acesso

O acervo da Pesquisa possui alta consonância com a sua missão. Tem uma associação histórica cultural significativa e objetiva prover acesso a estudos de obras de Guignard, obras essas, dispersas em sua maioria em mãos de particulares.

Esse acervo tem caráter permanente e a preservação deve alcançar todo material em formato digital e seus metadados. Ocorre, pois, uma seleção para disponibilização de acesso à

comunidade. Uma vez que Guignard não deixou herdeiros, essa seleção é baseada na propriedade intelectual e nos direitos de imagem.

As obras de Guignard são emprestadas à Universidade com fins de estudo científico, via recibo de empréstimos⁷⁰. Os direitos sobre os resultados de tais estudos, enquanto obras literárias e artísticas, são apenas da Pesquisa, e não, dos proprietários das obras analisadas, que ao concederem o empréstimo já se encontraram cientes de sua finalidade, podendo inclusive discordar e não o concederem. Dessa forma, os resultados dos estudos são de propriedade exclusiva da Pesquisa, cabendo a ela a decisão de sua disponibilização. Porém, a divulgação das imagens das obras obtidas pela Pesquisa é autorizada pelos respectivos proprietários, via contrato de permissão. No caso da obra artística que retrata a imagem de uma pessoa, o retratado também concede autorização expressa para divulgação de sua imagem.

Os entrevistados pelo processo de História Oral, após revisarem sua entrevista ou depoimento autorizam formalmente a sua disponibilização. Quando são retiradas fotos dos entrevistados, estas também são autorizadas para divulgação⁷¹.

6.2 Formação do acervo digital

O acervo da Pesquisa Guignard abrange a documentação fotográfica, o registro sonoro de entrevistas e depoimentos, os registros textuais sobre os estudos da Pesquisa e referências bibliográficas.

A documentação fotográfica ocorre de modo convencional (com uso de filmes e processamento fotoquímico) empregando materiais e procedimentos de iluminação que possibilitam uma máxima fidelidade de cores entre as imagens originais e as registradas em

⁷⁰ Os recibos foram elaborados de acordo com a Lei de Direito Autoral (Lei N° 9610).

⁷¹ A Pesquisa recebeu orientações da Acessória Jurídica de Propriedade Intelectual da Coordenadoria de Transferência e Inovação Tecnológica (CT&IT) /UFMG. A CT&IT atua na gestão do conhecimento científico e tecnológico, exercendo, entre outras, atividades concernentes à disseminação da cultura de propriedade intelectual, ao sigilo das informações sensíveis, à proteção do conhecimento e à comercialização das inovações geradas na UFMG.

filme. Os materiais empregados são filmes diapositivos coloridos (*slides*) balanceados para luz do dia, nos formatos 35 mm e 120 (6x6 cm). Para melhor equilíbrio tonal e neutralidade de cores, é usado o filme diapositivo *KODAK Ektachrome Professional EPN*, com o uso de cartela de referência de cores *KODAK Color Separation and Grayscale Chart*. Para o registro em filmes preto-e-branco sob luz infra-vermelha, é usado o filme *KODAK High Speed Infrared Film* (GARBOCI, 2004).

As imagens provenientes dos cromos são digitalizadas num escaner dedicado para filmes de elevadas resolução óptica e fidelidade de cores, o *Scanner Nikon, super coolscan 4000 ED*, após calibração do escaner com slide padrão (*Q-60E3 Color Calibration Target- 35mm Ektachrome film (IT-8)*). Também são utilizados os programas *NikonScan 3 e Vue Scan 7.6* de calibração que proporciona qualidade de cor com definição superior ao programa do escaner, realizando a geração de perfis de gerenciamento de cores que permitem a sua reprodução fiel em dispositivos de saída, como monitores e impressoras (LIMA et al, 2003). O monitor é calibrado com o colorímetro *Color Vision Spyder 2 Pro Studio*.

Essas imagens têm incorporado em seus arquivos o perfil de caracterização gerado pelo escaner, permitindo a sua visualização correta através de sistemas integrados de gerenciamento digital de cores (GARBOCI, 2004) (LEÃO, 2005).

As imagens matrizes são geradas no formato TIFF, com a resolução máxima do escaner (4000 *dpi*). Essa opção se deve à possibilidade de visualização de características consideráveis para o âmbito da Pesquisa. As imagens derivadas da matriz são as para impressão (300 *dpi*) e para visualização em monitor de vídeo (72 *dpi*), ambas em formato JPEG (LIMA et al, 2003).

As imagens de detalhes das obras e dos estudos de materiais e técnicas são obtidas a partir da câmera de captura digital *Coolsnap-Pro_c Collor* da *Media Cybernetics*, acoplada no microscópio Olympus do Cecor/EBA/UFMG, fazendo-se uso do programa *Image-Pro Plus* versão 4.5 da *Media Cybernetics* para a captura das imagens a 100 *dpi*, formato TIFF. As imagens das radiografias são digitalizadas no escaner *ScanMaker 9800 XL* da Microteck a resolução de 300 *dpi*, formato TIFF. As imagens das obras do artista publicadas nas referências bibliográficas analisadas são digitalizadas no escaner *Color Page Vivid+* da Genius a 72 *dpi* formato JPEG.

As entrevistas e depoimentos são gravados em fitas cassete. Para digitalização foi utilizada uma placa de som profissional da marca *CREATIVE*⁷² e o seu programa *Creative Wav Studio*. Sendo gerados arquivos mestres em formato WAV com 96 KHz e 24 bits de profundidade e a partir destes arquivos, arquivos para disponibilização na *Web*, no formato MP3 com 44.1 KHz e 16 bits de profundidade, ambos mono, pois a gravação original é mono.

6.3 Cuidados com os meios digitais

Os meios de armazenamento utilizados pela Pesquisa Guignard são CD-R e discos rígidos, estes últimos, armazenam tanto arquivos mestres quanto derivados, além da produção textual. Nos CDs são guardados arquivos mestres.

A Pesquisa possui uma limitação de recursos que a impossibilita de tomar cuidados mais sofisticados com os suportes digitais, como a manutenção de uma aclimatação adequada e seu controle. São empreendidos, portanto, recursos simples como:

- Correto acondicionamento das mídias em posição vertical;
- Acomodação em ambiente limpo;
- Não exposição dos discos rígidos a campos magnéticos e flutuações de energia;
- Acomodação em ambiente pouco iluminado, evitando contato com água;
- Manuseio correto, evitando-se depósitos de sujidades e impressões digitais;
- Acondicionamento dos CDs em embalagens apropriadas, de forma isolada;
- Minimização de mudanças ambientais bruscas;
- Não exposição a solventes orgânicos;
- Não aplicação de etiquetas nos discos.

As implicações climáticas, no entanto, são minimizadas graças ao clima da cidade de Belo Horizonte que não sofre grandes alterações, não provocando aclimatações bruscas.

⁷² <http://br.creative.com/>

6.4 Sistemas

Os sistemas de informação vêm possibilitar a organização, o armazenamento, transmissão e gerenciamento do acervo e seus metadados.

O desenvolvimento de sistemas para a Pesquisa Guignard seguiu duas vertentes: um voltado para gestão documental e outro para consultas, endereçado à comunidade.

Para o desenvolvimento do sistema de informação, houve uma customização do processo de Engenharia de Software PRAXIS⁷³, visando atender à necessidade de uma documentação básica, focando mais na especificação de requisitos para o sistema que dá acesso aos resultados obtidos pela Pesquisa Guignard e no atendimento da tipicidade do ambiente do sistema, a *Web* (PAULA, 2000).

6.4.1 Sistema de Obtenção de Dados

Inicialmente, a documentação dos resultados da Pesquisa era registrada apenas em planilhas impressas criteriosamente criadas pelos pesquisadores. As referências da bibliografia levantadas sobre Guignard e sua obra eram registradas no banco de dados *File Maker Pro*. Esses foram, portanto, os insumos para o desenvolvimento de um sistema apropriado às necessidades de gestão documental da Pesquisa (FIG. 7).

⁷³ PRAXIS - PRocesso para Aplicativos eXtensíveis InterativoS.

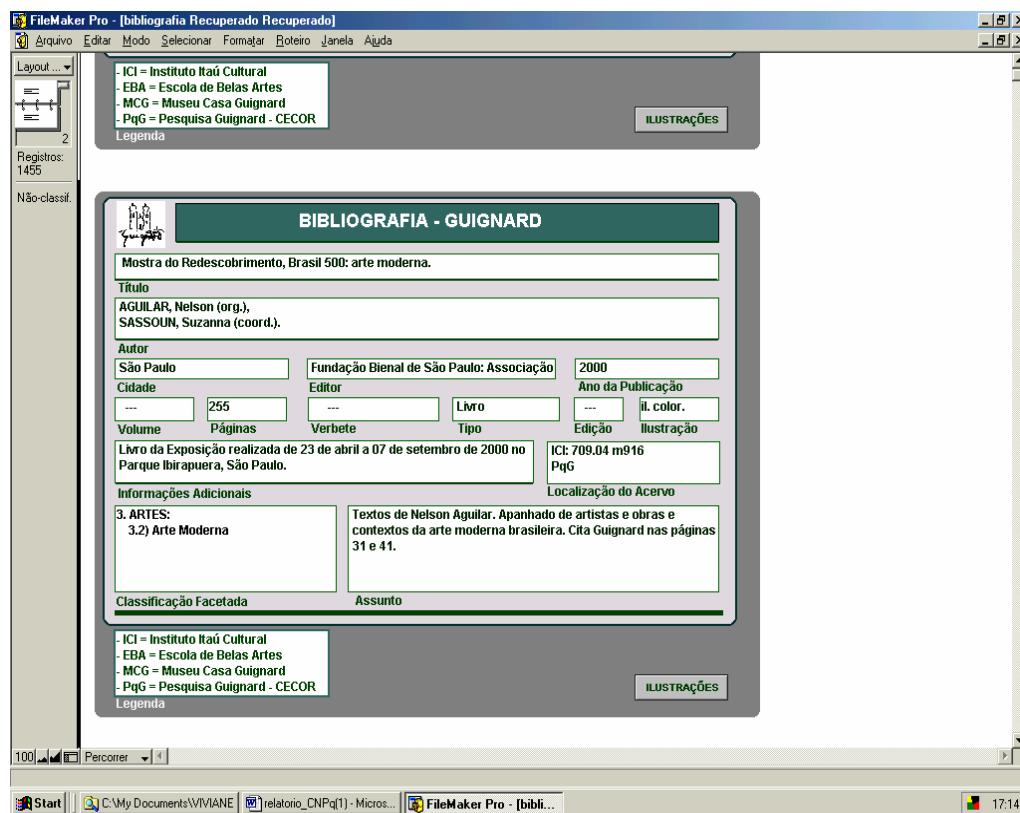


FIGURA 7 - Tela do File Maker Pro, sistema de banco de dados utilizado inicialmente pela Pesquisa Guignard para documentação bibliográfica.

Fonte: PRATES, 2004.

Para o desenvolvimento do sistema de gestão documental da Pesquisa, Sistema de Obtenção de Dados – SOD, ocorreram diversas reuniões para levantamento de requisitos e definição do escopo do sistema (FIG. 8 e FIG. 9). Em paralelo, deu-se a formação de um vocabulário controlado e tesouros específicos para o domínio de aplicação na Pesquisa Guignard, baseados em normas ABNT e no código de Catalogação Anglo Americano.

O SOD se divide nas partes coligadas referentes a:

- Imagens – Captura Digital e Registro Fotográfico: onde o fluxo de trabalho, desde a captura de imagens, sua digitalização e tratamento são cuidadosamente documentados;
- Obras: registro de informações resultantes dos estudos científicos das obras de Guignard;
- Bibliografia: cadastro de referências bibliográficas;
- História Oral: registro de entrevistas, depoimentos e fitas.

- Gestão: criação de novos usuários do sistema, definição de permissões, automatização de cópias de segurança e geração de relatórios.

No desenho do banco de dados do SOD, procurou-se evitar problemas como redundância de dados, inconsistências de dados redundantes, dispêndio de recursos para transferir informação de uma parte do sistema para a outra.



FIGURA 8 - Tela do Sistema de Obtenção de Dados – Bibliografia.

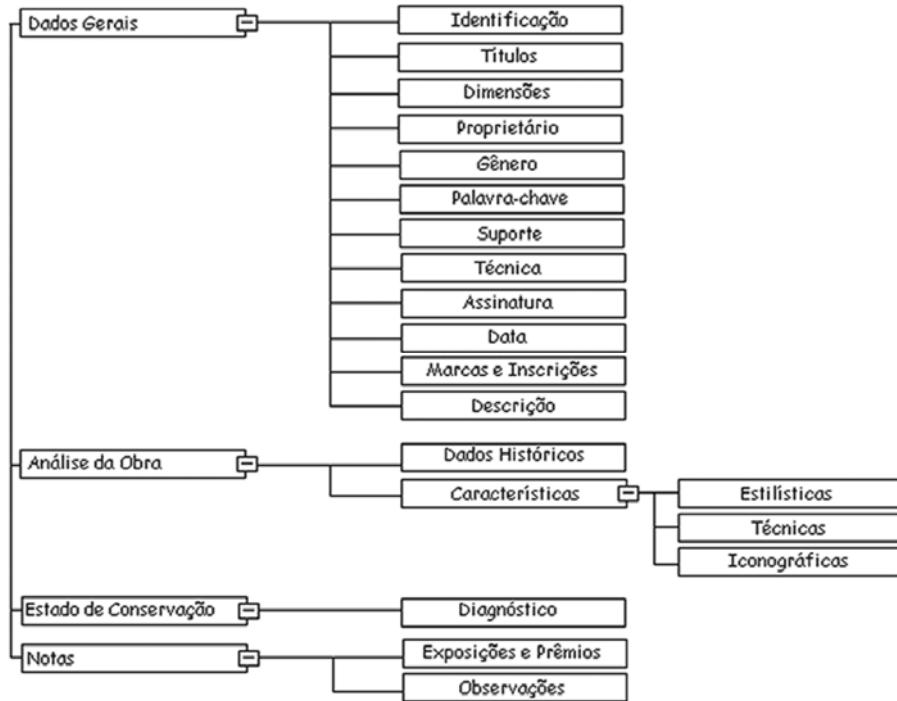


FIGURA 9 - Diagrama esquemático da representação da informação sobre Obras.
Fonte: LIMA et al, 2003.

Como um dos requisitos do SOD, está o acesso restrito e dedicado aos seguintes perfis de usuários:

- Administrador: Responsável por designar permissões de consulta, preenchimento e alteração das subdivisões do SOD. Perfil, este não implementado.
- Arquivista: Responsável pelo preenchimento de formulários específicos do sistema, podendo realizar consultas;
- Especialista: Responsável pela verificação do preenchimento dos formulários do sistema, podendo alterá-los.

Para o desenvolvimento do sistema foram utilizados recursos independentes de plataforma e não proprietários, como Servidor Web Apache⁷⁴ e as linguagens *HTML*, *JavaScript* e *PHP*⁷⁵ e o banco de dados *MySQL*⁷⁶.

⁷⁴ Apache. Disponível em: <<http://www.apache.org/>>.

⁷⁵ PHP - *Hipertext PreProcessor*. Disponível em: <<http://www.php.net/>>.

⁷⁶ MySQL. Disponível em: <<http://www.mysql.org/>>.

Os dados inseridos no banco de dados *File Maker Pro* foram migrados com sucesso para o banco *MySQL*, após o desenvolvimento do programa Bibliografia do SOD, através de uma ferramenta desenvolvida na Pesquisa para transformação do formato (*.fp3) para um compatível com o *MySQL* (BARBOSA, 2004).

O Sistema de Obtenção de Dados foi testado e tem alimentado o banco de dados da Pesquisa, com as informações textuais e visuais. Porém, o SOD necessita ser finalizado com o desenvolvimento de sua parte administrativa (Gestão), a criação de alguns campos e possíveis conexões entre as partes de Imagem, Obras, Bibliografia e História Oral.

6.4.2 Sistema de busca

Para disponibilização do acervo da Pesquisa na *Web* foi proposto o desenvolvimento de um sistema específico para este fim, que fará consultas às bases de dados alimentadas pelo SOD. Portanto, novas reuniões ocorreram com a equipe da Pesquisa visando à realização da engenharia de requisitos.

O processo de engenharia de sistema adotado foi o PRAXIS. Este processo é desenhado para suportar projetos de seis meses a um ano de duração, realizados individualmente ou em equipe (PAULA, 2000). No entanto, foram realizadas adaptações no processo para adequação das especificidades da engenharia da *Web*, cobrindo os requisitos como a análise de conteúdo, de interação, função e configuração.

A notação gráfica UML⁷⁷ foi utilizada para descrição do projeto de sistema a fim de propiciar melhor entendimento sobre os aspectos estruturais e comportamentais do sistema através de representações diagramáticas.

A missão do sistema foi definida como: disponibilizar informações sobre a Pesquisa Guignard e resultados de seus estudos, provendo recursos de buscas às bases de dados do Sistema de

⁷⁷ UML – *Unified Modeling Language*.

Obtenção de Dados da Pesquisa Guignard, bem como recurso de busca interna no próprio Sistema de Acesso, via *Internet*, para uso do público em geral.

As funcionalidades e casos de uso relacionados identificados são os descritos no quadro (QUADRO 7) e representados no diagrama de casos de uso (FIG. 10). Os casos de uso, constituem uma abordagem simples para identificar os requisitos funcionais e não-funcionais específicos de cada interação entre os atores⁷⁸ e sistema, refletindo o comportamento do sistema sob diversas condições, apontando como o sistema responde aos atores.

QUADRO 7 - Funções do produto

Casos de uso	Descrição
Emitir Conteúdo para Impressão	Retirar elementos gráficos do layout da página para imprimi-la.
Alterar Tamanho de Letras	Alterar tamanho de caracteres de texto, os aumentando e os diminuindo.
Exibir Definição de Termo	Exibir o significado de uma palavra ao se pousar o mouse sobre a mesma.
Cadastrar Colaboradores	Cadastrar colaboradores (pessoa que possua obra(s) de Guignard e/ou conheceu o artista) para a Pesquisa. Envio de notificação de cadastro para membro da Pesquisa.
Pesquisar Bibliografia	Buscar referencia bibliográfica da Pesquisa de modo dirigido através da escolha do ano de publicação ou através do uso de palavras chaves.
Pesquisar de Forma Dirigida por Gênero ou Cronologia	Buscar obras estudadas pela Pesquisa Guignard por gênero ou cronologia sem entrada de informação. Operação de busca para auxiliar o usuário leigo no encontro de informações sobre obras de Guignard.
Buscar por Imagem Exemplo	Recuperar uma imagem semelhante à imagem dada como exemplo.
Buscar por Resultados de Estudos de Obras	Recuperar resultados de estudos de obras obtidos pela Pesquisa, de modo avançado, através da entrada de texto.
Buscar no Sítio da Pesquisa	Buscar conteúdo do sítio da Pesquisa.
Contactar Pesquisa	Disponibiliza um formulário simples para o usuário escrever e enviar dúvidas e comentários a Pesquisa Guignard, através de correio eletrônico.

Fonte: LIMA, 2007.

Os atores identificados foram (QUADRO 8):

⁷⁸ Usuários humanos ou outros sistemas (sistemas externos) que se comunicam com o sistema do projeto.

QUADRO 8 - Descrição de atores

Atores	Definição
Consulente	Pessoa que consulta as informações da Pesquisa Guignard. Utilização do esporádica do sistema.
Gerenciador de e-mails	Gerenciador de e-mails responsável por administrar os e-mails que são enviados e os que são recebidos.
SOD	Sistema de Obtenção de Dados – Sistema responsável pela gestão documental da Pesquisa Guignard.
Pesquisador	Membro da Pesquisa Guignard. Freqüente utilizador do sistema.

Fonte: LIMA, 2007.

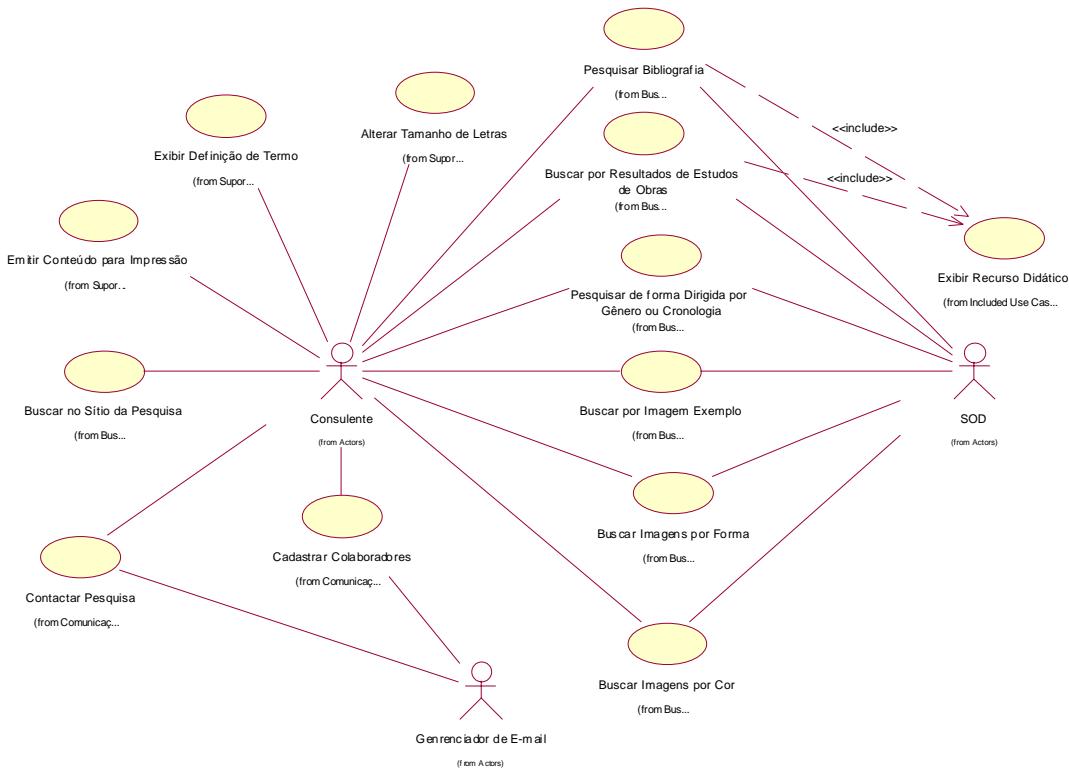


FIGURA 10 - Diagrama de caso de uso para o sistema de acesso ao acervo da Pesquisa.
Fonte: LIMA, 2007.

A configuração do sistema é a mesma adotada no SOD, Servidor Web Apache e a linguagens *HTML, JavaScript e PHP* e o banco de dados *MySQL*. Devendo o sistema ser acessado nos navegadores *Mozilla Firefox 1.5.0.7* ou superior, *Internet Explorer 6.0.2800* ou superior.

Como parte da definição da análise de interação, o projeto de interface foi desenvolvido (FIG. 12, 13 e 14). Este projeto descreve a estrutura e organização da interface com o usuário, incluindo uma representação do layout do usuário rica e suficiente para transparecer os modos de interação e os mecanismos de navegação, atendendo também a um projeto estético ou gráfico, com definições de cores, tamanhos de fonte, diagramação e layout esquemático (FIG. 11).

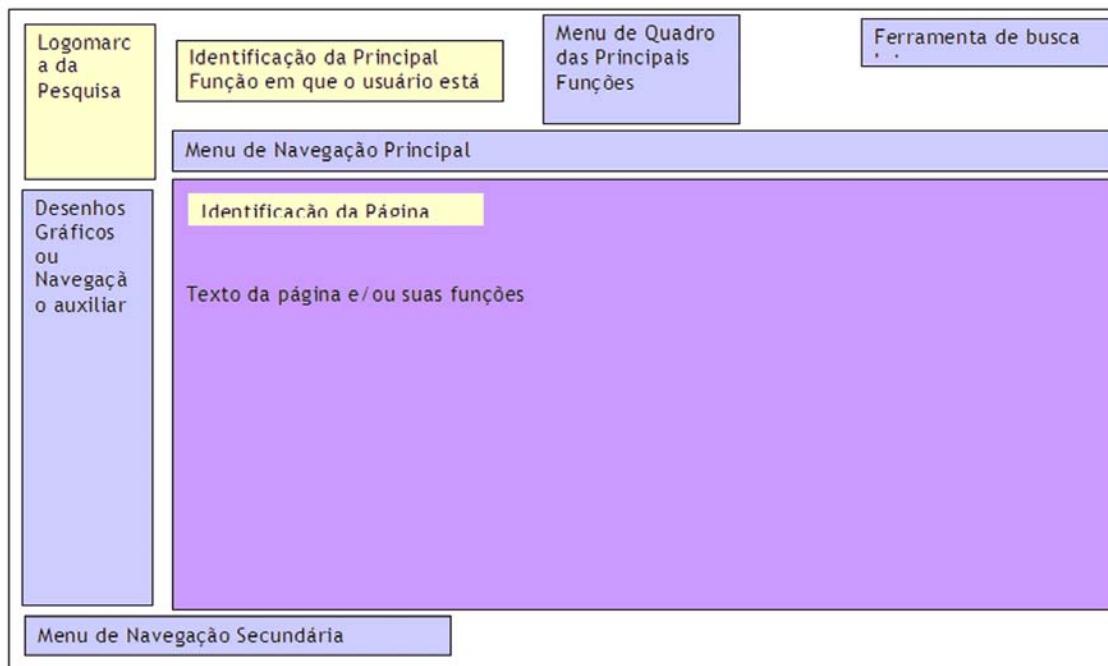


FIGURA 11 - Leiaute esquemático.
Fonte: LIMA, 2007.

 Guignard	Pesquisa Guignard	Buscas	Entrevistas e Depoimentos	ok
Entrevistas e Depoimentos				
Entrevista e Depoimentos por Entrevistados				
Tamanho da letra Aa				
Navegar por:				
Assunto				
Entrevistados				
Ivone Luzia	 PDF	 áudio		
Metodologia de ensino de Guignard	 PDF	 áudio		
Técnicas do artista	 PDF	 áudio		
A obra de Guignard	 PDF	 áudio		
Jarbas Juarez	 PDF	 áudio		
Metodologia de ensino de Guignard	 PDF	 áudio		
Contexto da época	 PDF	 áudio		
O mestre Guignard	 PDF	 áudio		
Gêneros das obras de Guignard	 PDF	 áudio		
Gelson Antunes	 PDF	 áudio		
Guignard em Ouro Preto	 PDF	 áudio		
A obra de Guignard	 PDF	 áudio		

FIGURA 12 - Tela Entrevistas e Depoimentos.
Fonte: LIMA, 2007.

Guignard · pesquisas · Buscas

Bibliografia | Trabalho Artístico | Obras Estudadas

Pesquisa | Vida do Artista | Buscas | Entrevistas e Depoimentos

Busca no site

FIGURA 13 - Tela Busca por Resultados de Estudos de Obras.
Fonte: LIMA, 2007.

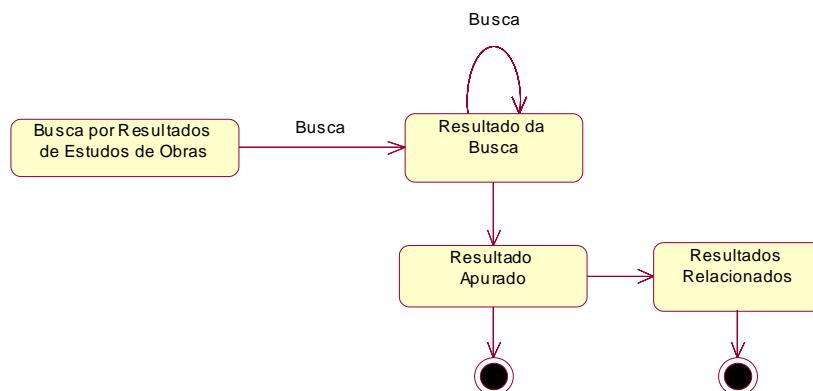


FIGURA 14 - Diagrama de estados para a Tela Busca por Resultados de Estudos de Obras.
Fonte: LIMA, 2007.

Desta forma, o projeto de interface visou atingir um alto nível de pregnância⁷⁹ e identificação com a Pesquisa Guignard, instilando no usuário uma sensação de controle sobre as ações, através da visualização ampla de suas opções de navegação e comunicação do estado de atividades empreendidas. O projeto possui consistência, seus elementos gráficos auxiliam na navegação do usuário indicando sua localização no sistema de forma coerente em todas as páginas *Web*. O usuário possui flexibilidade para realizar tarefas de forma direta ou através da exploração do sistema de modo aleatório. Assim, todo desenho propicia um aprendizado rápido em um projeto intuitivo, que organiza conteúdo e funcionalidade em categorias que são óbvias para o usuário.

O projeto de conteúdo visou à melhor comunicação de idéias por meio da organização do conteúdo nas páginas do sistema, estabelecendo-se uma arquitetura em estrutura hierárquica, onde se é possível alcançar um conteúdo através de um fluxo de controle horizontal e, através desse, ao longo de ramos verticais da estrutura, ou seja, o conteúdo apresentado mais à esquerda da hierarquia pode ter ligações de hipertexto que levem ao conteúdo existente no meio, ou no ramo mais a direita da estrutura.

6.4.3 Ferramenta de busca pelo conteúdo

Mesmo com a organização dos metadados aplicando o uso de vocabulários controlados e tesouros, foi percebida a necessidade de mecanismos de busca mais complexos, que facilitassem o processo de recuperação de imagens para usuários com pouco ou nenhum conhecimento prévio sobre Guignard e sua obra. Em obras artísticas, o conteúdo textual referente a uma imagem pode ser incerto ou polêmico, sendo classificado diferentemente por pesquisadores distintos (é possível a existência de referências a uma mesma obra de Guignard com títulos diferentes, por exemplo).

⁷⁹ Pregnância ou pregnância das formas: Qualidade que determina a facilidade com que percebemos figuras. Percebemos mais facilmente as boas formas, ou seja, as simples, regulares, simétricas e equilibradas.

Dante desse quadro, como parte do sistema de acesso, foi desenvolvida uma ferramenta para recuperação de imagens com base no conteúdo, que faz consulta a partir de uma imagem-exemplo. Ou seja, sua operação consiste no fornecimento de uma imagem como insumo para a consulta e no retorno de imagens semelhantes à imagem dada, seguindo as características definidas para extração de conteúdo da imagem. A ferramenta opera em duas fases: Na fase *offline* há a extração de assinaturas⁸⁰ para a composição de um banco de assinaturas. Na fase *on-line* há a geração de uma assinatura para a consulta e a comparação entre assinaturas (FIG. 15). A ferramenta utiliza imagens no formato JPEG com resolução de 100 dpi (BARBOSA, 2006). Ela deverá ser disponibilizada no sistema de acesso para a comunidade, adotando o leiaute do sistema (FIG. 16 e 17) e operando com imagens com 72 dpi de resolução.

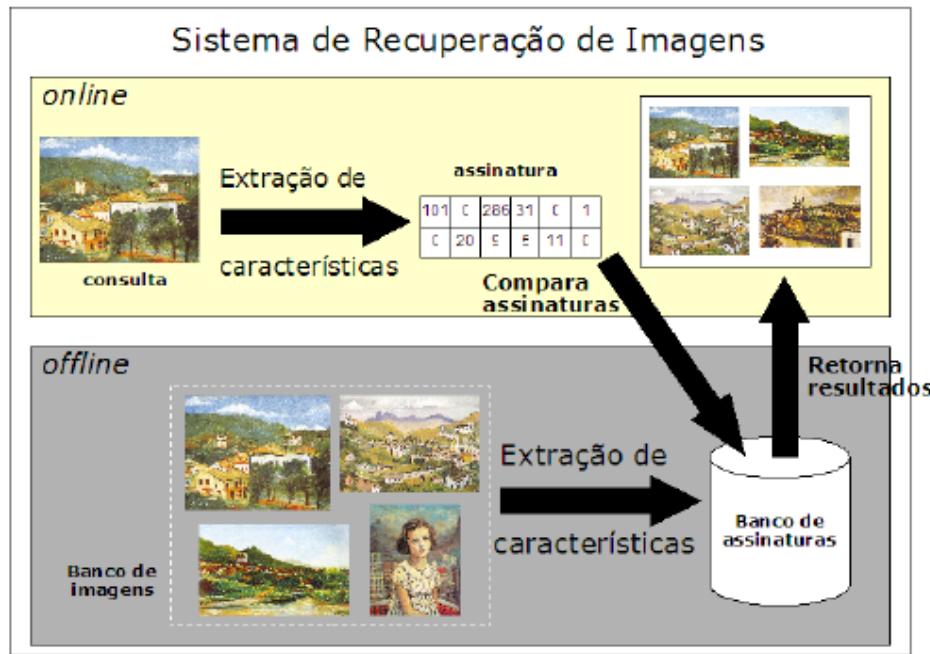


FIGURA 15 - Diagrama de funcionamento do sistema de recuperação de imagens com base em imagem exemplo.
Fonte: BARBOSA, 2006.

⁸⁰ A assinatura de uma imagem corresponde às características relevantes (porcentagem de cor, por exemplo) extraídas da imagem e organizadas num vetor.

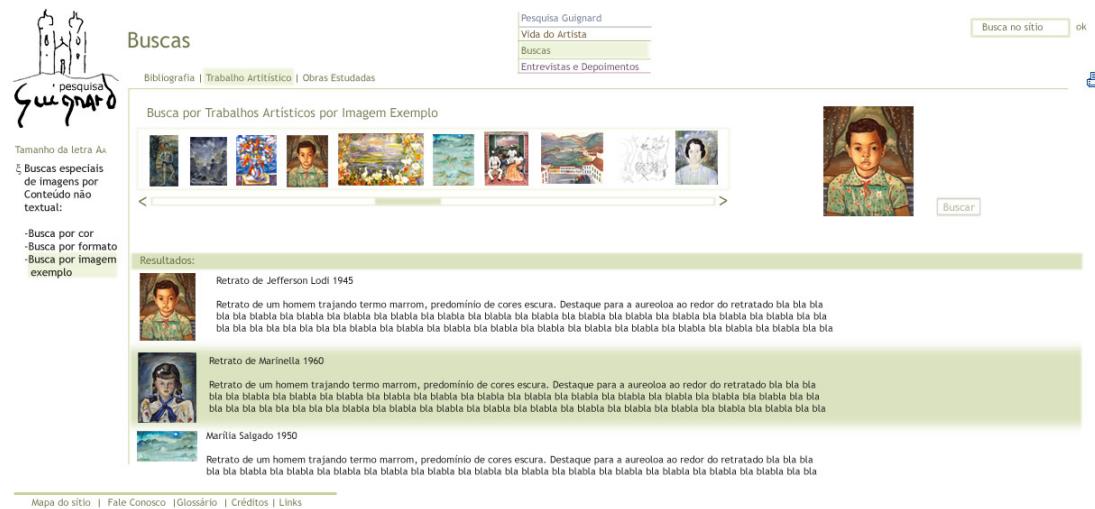


FIGURA 16 - Tela Busca por Imagem Exemplo aos Trabalhos Artísticos Estudados pela Pesquisa.
Fonte: LIMA, 2007.



FIGURA 17 - Diagrama de estado para a Tela Busca por Imagem Exemplo aos Trabalhos Artísticos Estudados pela Pesquisa.
Fonte: LIMA, 2007.

6.4.4 Considerações finais

A experiência adquirida na Pesquisa Guignard com o uso do Sistema de Obtenção de Dados e o emprego da engenharia de sistemas e da *Web*, para o levantamento de requisitos do sistema de acesso, possibilitou a percepção da necessidade de finalização completa do SOD. Para tanto, a estrutura do banco de dados da Pesquisa passará por pequenos aperfeiçoamentos, porém significativos. Desta forma, a implementação das ferramentas de busca textuais do sistema de acesso está condicionada às alterações no SOD. O desenvolvimento evolutivo do

sistema está iniciado, ele advém da maturidade da Pesquisa e faz parte do ciclo de vida dos sistemas.

6.5 Propostas futuras

A Pesquisa Guignard tem se esmerado, em seus poucos anos de existência⁸¹, na preservação digital de seu acervo, estando atenta à necessidade de desenvolvimento de sistemas apropriados ao seu domínio de aplicação e a melhor manutenção de seu acervo a fim de atingir uma maior longevidade.

Ao longo desses anos, o emprego de estratégias de preservação não foi tão necessário. Ocorreu apenas uma migração de dados entre bancos de dados distintos, sem perdas ou corrupção da informação. Os equipamentos utilizados pela Pesquisa desde seu início permanecem os mesmos, seus manuais, licenças e documentação são conservados. Não obstante, a constante atenção à evolução tecnológica deve continuar sendo empreendida.

A Pesquisa tem realizado cópias em CDs e em disco rígido do seu acervo, arquivos mestres e documentos. Porém, um nível de redundância maior da informação distinguida como prioritária é recomendável para uma maior garantia contra falhas de uma cópia ou componente, o que inclui a aquisição de sistemas de cópia de segurança apropriados. Principalmente, para as imagens capturadas no microscópio.

As mídias da Pesquisa recebem um manuseio adequado e estão dentro de condições regulares de armazenamento, pois não é realizado um controle climático, o que é amenizado com a não exposição das mídias a variações climáticas bruscas. Entretanto, a execução de um controle de qualidade periódico das mídias com uso de programa de computador especial para avaliação do nível de deterioração progressiva se faz recomendável a fim de possibilitar a transcrição da informação em tempo, dentro das necessidades.

⁸¹ A Pesquisa Guignard teve início no ano 2000.

Portanto, por possuir uma limitação de recursos, a Pesquisa deve prezar pela prevenção e evitar se expor a eventuais riscos contra o seu patrimônio e à necessidade de investimento em recuperação de perdas.

7 CONCLUSÃO

A tecnologia tem progredido de modo muito ágil, impelindo uma grande transitoriedade de mídias de armazenamento, formatos de arquivos, dispositivos, computadores, protocolos, sistemas, programas e suas versões, e meios de difusão de informação. As incertezas advindas de constantes prognósticos tecnológicos mostram a dificuldade de se definir estratégias consoantes com o futuro e acenam a urgência de se assegurar a longevidade de uma herança cultural, histórica e científica digital. Dentro desse quadro, nasce a necessidade de uma gestão capaz de assegurar a manutenção a longo prazo de uma cadeia de *bytes*, constituintes da informação, provendo acesso contínuo ao seu conteúdo.

Através deste trabalho, foi possível participar orientações sobre o desenvolvimento de um plano de preservação e averiguar a proposta adotada pela Pesquisa Guignard. Em seu discorrer, foi possível observar a importância da definição de responsabilidades e emprego de políticas organizacionais apropriadas para o emprego da preservação digital.

Como início do trabalho, houve a caracterização do patrimônio digital, sua origem, fragilidade e extensão. Foram explorados os cuidados que devem ser tomados para evitar a perda desse patrimônio, com a exploração de estudo sobre conservação e manuseio de mídias.

Questões ligadas à seleção de materiais para serem preservados também foram tratadas. Dentro do contexto da preservação, as instituições precisam gerir seus recursos para melhor empreendimento de uma política adequada. Diante de uma grande massa digital é cabível a priorização de parte do acervo em detrimento de outras, devendo-se para isso adotar critérios de seleção do patrimônio que deve ser mantido para as gerações futuras.

Foram consideradas as principais e mais difundidas estratégias de preservação: refrescamento, migração, emulação e arqueologia digital. Este tópico configura grande importância para o auxílio de decisões que visem perdurar os conteúdos digitais, ainda que não garantindo sua longevidade. A aplicação das opções de estratégias de preservação, contudo, dependem do orçamento, tipicidade do acervo, ambiente computacional e programas. É cabível se apregoar que o melhor alcance de resultados está condicionado a uma combinação de tais estratégias.

No contexto das instituições promotoras e orientadoras da preservação digital, foi possível a verificação de medidas que visam estimular e disciplinar a organização de diretrizes, através do desenvolvimento de sistemas, documentos de requisitos de sistemas e estabelecimento de padrões, que objetivam melhor possibilitar uma maior integração e comunicação entre as diversas instituições patrimoniais no mundo.

Como parte de uma política de preservação digital, está a digitalização de materiais analógicos e a conversão de formatos nato-digitais. Foi observado que para tanto, métodos devem ser empregados para assegurar a qualidade do acervo digital, como o emprego da limpeza do material a ser digitalizado, o uso de cartelas de cores e calibração de equipamentos, de acordo com a pertinência. Durante este processo, as instituições devem procurar restringir a variedade de formatos adotados, pois uma maior diversidade implicará em maiores custos e trabalho para preservação. Um especial cuidado deve ser adotado com os arquivos mestres, sendo eles, sempre uma escolha dentro dos critérios de seleção. No entanto, o descarte dos originais não deve ocorrer, pois estes podem fazer-se essenciais nos casos de perdas catastróficas no acervo digital.

Dentro do conceito de preservação, está a disponibilização do acervo para acesso e como consequente, a geração de metadados e seus esquemas, como promotores de resultados apurados na recuperação de informação. Para definição de um padrão de metadados a ser utilizado dentro de um escopo, faz-se importante o uso de vocabulários controlados e tesouros. Também, é interessante observar padrões de metadados já estabelecidos por modelos como OAIS a fim de haver uma consonância com outras instituições possibilitando maiores interações. Outro importante aspecto é a criação e conservação de metadados para preservação, estes poderão ser utilizados para recuperação de informações e possibilitarão a utilização de estratégias de preservação.

É, também, fundamental na promoção do acesso, o desenvolvimento de sistemas de informação para veicular a disponibilização do acervo à comunidade interessada, através de buscas automatizadas, e para gerir o acervo digital e seus metadados relacionados. Muitas instituições são caracterizadas pela particularidade de seu acervo, o que as orienta para o desenvolvimento de sistemas mais especializados para o domínio de sua aplicação. Principalmente, nestes casos, onde os sistemas ou a documentação de requisitos disponíveis

no mercado não são adequados ou precisam sofrer adaptações para maior acomodação de necessidades específicas, a engenharia de sistemas deve ser aplicada e um processo de desenvolvimento de sistemas deve ser escolhido e muito comumente adaptado para o escopo do projeto, a fim de se atingir resultados satisfatórios. Sempre que possível, deve-se fazer a opção por sistemas, plataformas e padrões abertos, a fim de evitar entraves na preservação, com o uso de ferramentas proprietárias.

Como demonstração do emprego da preservação digital, houve o estudo de caso da proposta adotada pela Pesquisa Guignard, formadora de patrimônio cultural artístico digital. O método adotado pela Pesquisa corresponde ao de uma instituição detentora de poucos recursos, se alicerçando mais na prevenção. Como proposta de trabalho futuro, foi recomendada a adoção de maior redundância de cópias de parte do acervo, como das imagens capturadas no microscópio, pois elas são nativo-digitais. Aconselha-se, também, a aquisição de um programa para checagem do estado de conservação das mídias, a fim possibilitar a transcrição da informação nelas contidas, antes que o limite de vida útil seja atingido. O relato da experiência da Pesquisa e considerações de melhorias podem ser norteadores para outras instituições com custódia de acervos digitais.

Finalmente, pode-se concluir com este trabalho, que lidar com a preservação digital é lidar com constante controle de riscos. O reconhecimento de que perdas são inevitáveis dá relevância aos modos de evitá-las. Não existe consenso quanto às melhores estratégias e metodologias a serem empregadas num plano de preservação digital, pois nenhuma medida singular é capaz de atender todos os requisitos de todos os tipos de materiais digitais. No entanto, quanto mais a gestão de recursos digitais estiver conforme com padrões e boas práticas, maior probabilidade haverá da sustentabilidade do patrimônio. Embora o conhecimento e a experiência que existem hoje não sejam capazes de garantir a durabilidade dos materiais digitais, a adoção de planos de preservação digital, mesmo que básicos, se faz necessária.

REFERÊNCIAS

ALLETT, C. Dennis *et al.* *User Involvement in The Design Process: Why, When & How?* In *Interchi 93*. Disponível em: <<http://www.ieee.org>>. Acesso em: 14 abr. 2007.

ÁLVARES, Patrícia Marques Rodrigues. *WebPraxis – um processo personalizado para projetos de desenvolvimento para a Web*. Dissertação (mestrado) - Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, Belo Horizonte, 2001. 105p. Disponível em: <http://www.wppf.uaivip.com.br/pesquisa/orientacoes_recentes.htm>. Acesso em Jun. 2006.

APACHE. Disponível em: <<http://www.apache.org>>. Acesso em: 11 abr. 2007.

BARBOSA, Maria Cecília Ribeiro. *Sistema de Informação Multimídia para a Pesquisa Guignard*. Belo Horizonte: Relatório parcial de bolsista de iniciação científica, 2004. 12p.

BARBOSA, Maria Cecília Ribeiro. *Protótipo para Recuperação de Imagens com base no conteúdo aplicado à Pesquisa Guignard*. Relatório Final de Projeto Orientado de Computação (bacharelado em Ciência da Computação). 2006. 16 f. Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, Belo horizonte, 2006.

BEARMAN, David. *Reality and Chimeras in the Preservation of Electronic Records*. D-Lib Magazine, Apr.-1999. Disponível em: <<http://www.dlib.org/dlib/april99/bearman/04bearman.html>>. Acesso em: 2 abr. 2007.

BESSER, Howard. *Introduction to imaging*. Technical report, Getty Information Institute, 2003. Disponível em: <http://www.getty.edu/research/conducting_research/standards/intro_images/index.html>. ISBN 978-0-89236-733-7. Acesso em: 10 abr. 2007.

BOGART, John W.C. Van. *Armazenamento e Manuseio de Fitas Magnéticas*: Um guia para bibliotecas e arquivos. Coord. Ingrid Beck, Trad. José Luiz Pedersoli Júnior. 2. ed. Rio de Janeiro: Arquivo Nacional, 2001. 46 p. (Trad. de *Magnetic Tape Storage and Handling – A Guide for Libraries and Archives*). Disponível em: <http://siarq02.siarq.unicamp.br/cpba/pdf_cadtec/42.pdf>. Acesso em: 2 nov. 2006.

BOGART, John W.C. Van; MERZ John. *St. Thomas Electronic Records Disaster Recovery Effort*. National Media Laboratory, 1995. 14 p. Disponível em: <http://www.imation.com/government/nml/pdfs/AP_NMLdoc_StThomasElectronic.pdf>. Acesso em: 10 mar. 2007.

BOSTON, George; SCHÜLLER, Dietrich. Magnetic Carriers. *UNESCO 00*. Disponível em: <http://webworld.unesco.org/safeguarding/en/all_magn.htm>. Acesso em: 10 abr. 2007.

BOSTON, George; SCHÜLLER, Dietrich. Optical Carriers. *UNESCO 99*. Disponível em: <http://webworld.unesco.org/safeguarding/en/all_opti.htm>. Acesso em: 10 abr. 2007.

BRAIN, Marshall. How Hard Disks Work. *How Stuff Works*, HowStuffWorks Inc., 2007. Disponível em: <<http://computer.howstuffworks.com/adding-a-hard-drive1.htm>>. Acesso em: 15 mar. 2007.

BRASIL. Lei de Introdução ao Código Civil Brasileiro, Lei no. 10.406, de 10 de Janeiro de 2002. Disponível em: <<http://www.planalto.gov.br/CCIVIL/leis/2002/L10406.htm#secaoicomodato>>. Acesso em: 24 abr. 2007.

BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Diário Oficial [da República Federativa do Brasil], Brasília, 19 fev. 1998. Disponível em: <<http://www.planalto.gov.br/CCIVIL/Leis/L9610.htm>>. Acesso em: 10 mar. 2007.

BYERS, Fred R.. *Care and Handling of CDs and DVDs: A Guide for Librarians and Archivists*. Washington, D.C.: Council on Library and Information Resources, National Institute of Standards and Technology, 2003. 49 p. Disponível em: <<http://www.clir.org/pubs/reports/pub121/pub121.pdf>>. Acesso em: 11 nov. 2006.

CAMiLEON. Disponível em: <<http://www.si.umich.edu/CAMILEON/about/aboutcam.html>>. Acesso em: 12 abr. 2007.

CATEGORIES FOR THE DESCRIPTION OF WORKS OF ART - CDWA. Disponível em: <http://www.getty.edu/research/conducting_research/standards/cdwa/>. Acesso em: 7 abr. 2007.

CCSDS – CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. *Reference Model for an Open Archival Information System (OAIS), Recommendation for Space Data System Standards, CCSDS 650.0-B-1*. Blue Book. Washington, D.C.: Consultative Committee for Space Data Systems, 2002. 148p. Disponível em: <<http://public.ccsds.org/publications/BlueBooks.aspx>>. Acesso em: 30 jan. 2007.

CDP - COLORADO DIGITIZATION PROJECT. *Digital Audio Best Practices*, Version 2.0 (November 2005), http://www.cdpheritage.org/digital/audio/documents/CDPDABP_1-2.pdf. Acesso em: 29 abr. 2006.

CEDARS - CURL EXEMPLARS IN DIGITAL ARCHIVES. Disponível em: <<http://www.leeds.ac.uk/cedars/>>. Acesso em: 7 abr. 2007.

CHEN, P. M. et al. RAID: High-Performance, Reliable Secondary Storage. *ACM Computing Surveys*, v.26, n.2, p.145-185, 1994. Disponível em: <<http://citeseer.nj.nec.com/chen94raid.html>>. Acesso em: 17 jan. 2007.

CLIR - CONCIL ON LIBRARY AND INFORMATION RESOURCES. *Authenticity in a Digital Environment*. CLIR. Report, 2000a. Disponível em: <<http://www.clir.org/pubs/reports/pub92/contents.html>>. Acesso em: 18 mar. 2007.

CLIR - COUNCIL ON LIBRARY AND INFORMATION RESOURCES. *Capturing Analog Sound for Digital Preservation: Report of a Roundtable Discussion of Best Practices for Transferring Analog Discs and Tapes.* 2006. Disponível em <<http://www.clir.org/pubs/reports/pub137/pub137.pdf>>. Acesso em 20 mar. 2007.

CLIR - COUNCIL ON LIBRARY AND INFORMATION RESOURCES. *Guides to Quality in Visual Resource Imaging,* Research Libraries Group, 2000b. Disponível em: <<http://www.rlg.org/visguides/index.html>>. Acesso em: 26 jan. 2007.

COCKBURN, Alistair. *Escrevendo Casos de Uso Eficazes – Um guia prático para desenvolvedores de software.* Trad. Roberto Vedoato. Porto Alegre: Bookman, 2005.

CONARQ - CONSELHO NACIONAL DE ARQUIVOS. *Carta para Preservação do Patrimônio Arquivístico Digital.* CONARQ, 2004. Disponível em: <http://www.arquivonacional.gov.br/conarq/cam_tec_doc_ele/download/CartaPreservPatrimArqDigital-Conarq-2004.PDF>. Acesso em: 12 dez. 2006.

CONARQ - CONSELHO NACIONAL DE ARQUIVOS. *Modelo de requisitos para sistemas informatizados de gestão arquivística e-ARQ.* CONARQ, 2006. Disponível em: <<http://www.conarq.arquivonacional.gov.br/Media/publicacoes/earqbrasilv1.pdf>>. Acesso em: 5 abr. 2007.

CONWAY, Paul. *Preservação no universo digital.* Coord. Ingrid Beck, Trad. Olga Marder. Rio de Janeiro: Arquivo Nacional, 1997. 24p. (Tradução de *Preservation in the digital world*). Disponível em: <http://siarq02.siarq.unicamp.br/cpba/cadtec/cadtec_52.htm>. Acesso em: 26 jan. 2007.

DOLLAR, Charles M. *Authentic Electronic Records: Strategies for Long-term Access* (Chicago, Ill.: Cohasset Associates, 1999). See also, Paul Wheatley, "Migration—a CAMILEON discussion paper," *Ariadne* 29, September 2001.

DUBLIN CORE. Disponível em: <<http://dublincore.org/>>. Acesso em: 8 abr. 2007.

DURANTI, Luciana. Reliability and Authenticity: The Concepts and Their Implications. *Archivaria*, n.39, 1995. 6p. Disponível em: <<http://journals.sfu.ca/archivar/index.php/archivaria/article/viewFile/12063/13035>>. Acesso em: 22 mar. 2007.

EVANS, Frank B.; HIMLY, François J.; WALNE, Peter Walne (Comp.). *DICTIONARY of archival terminologya english and french; with equivalents in dutch, german, italian, russiaan and spanish = Dictionnaire de terminologie archivistique.* München, New York, London, Paris: ICA, 1984. 226 p. (ICA Handbooks series, v.3)

FEDROA. Disponível em: <<http://www.fedora.info>>. Acesso em: 26 jan. 2007.

FOWLER, Martin. *UML Essencial – Um breve guia para linguagem-padrão de modelagem de objetos.* 3.ed. Trad. João Tortello. Porto Alegre: Bookman, 2005.

FREE LOSSLESS AUDIO CODEC - FLAC. Disponível em: <<http://flac.sourceforge.net/>>. Acesso em: 3 abr. 2007.

GARBOCI, Renata Silva. *Pesquisa Guignard* - Documentação Fotográfica/ Digitalização, tratamento e arquivamento das imagens. Belo Horizonte: Relatório Técnico-científico de final de bolsa BDTI/FAPEMIG, 2004. 3p. Relatório.

GARZON, Ana Rosa Lemos da Cunha. *Direitos autorais* – busca do equilíbrio. 2006. 114 f. Dissertação (mestrado) - Escola de Belas Artes, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

GERTZ, Janet. *Selection Guidelines for Preservation* (paper presented at the Jprint RLG and NPO Preservation conference, Guidelines for Digital Imaging, September 28-30, 1998). Disponível em: <www.rlg.org/preserv/joint/gertz.html>. Acesso em: 3 abr. 2007.

GLADNEY, H. Use and misuse of OAIS. *Digital Document Quarterly*, v.1, n.3. Disponível em: <http://home.pacbell.net/hgladney/ddq_1_3.htm>. apud (HOWELL, 2001). Acesso em: 3 abr. 2007.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. Addison-Wesley Publishing Company, 1992.

HEDSTROM, Margaret; LAMPE, Clifford. *Emulation Vs. Migration: Do Users Care?* RLG DigiNews 5.6, 2001. Disponível em:< <http://www.rlg.org/preserv/diginews/diginews5-6.html#feature1>>. Acesso em: 3 abr. 2007.

HIX, D.; HARTSON, H.R. *Developing User Interfaces*. New York: John Wiley & Sons, 1993.

HOWELL, Alan. Perfect one day – digital the next: challenges in preserving digital information. *AARL – Australian Academic & Research Libraries*, p.121-141, dez. 2000.

HOWELL, Alan. Preservation Solutions. Australia. *AARL – Australian Academic & Research Libraries*, 76p., jun. 2001.

HUMMELEN, Ijsbrand; SILLÉ, Dionne (Ed.). *Modern art: who cares?* An interdisciplinary research project and an international symposium on the conservation of modern and contemporary art. Amsterdam: The Foundation for the Conservation of Modern Art and the Netherlands Institute for Cultural Heritage, 1999. 445p.

INNARELLI, Humberto Celeste. *Preservação de documentos digitais*: Confiabilidade de mídias CD_ROM e CD-R. Dissertação (mestrado) - Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, 2006. 170p. Disponível em: <<http://libdigi.unicamp.br/document/?code=vtls000384479>>. Acesso em: 22 mar. 2007.

IQWGAC - IMAGE QUALITY WORKING GROUP OF ARCHIVESCOM, a joint Libraries/AcIS committee. *Technical Recommendations for Digital Imaging Projects*,

American Columbia: 1997. Disponível em: <<http://www.columbia.edu/acis/dl/imagespec.html>>. Acesso em: 20 jun 2006.

ISO. *ISO 9241-11*. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability, 1998.

ISO. *ISO/IEC 9126*. Information Technology. Software product evaluation. Quality characteristics and guidelines for their use - International standards ISO9126. 1.ed., 1991.

ISO. *ISO/TC46*. Information et documentation. Sous-comité SC 11, Archives/records management – International standards ISO 15489-1: information and documentation: records management. Genève: ISO, 2001.

JOINT PICTURE EXPERT GROUP. Disponível em: <<http://www.jpeg.org/>>. Acesso em: 10 abr. 2007.

KENNEY, Anne R.; RIEGER, Oya Y. *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. Mountain View, CA: Research Libraries Group, 2000. 188p.

LANZELOTTE, R. S. G.; MARQUES, M. P.; PENNA, M. C. G.; Portinari, J. C.; Ruiz, I. D.; SCHWABE, D. The Portinari project: Science and art team up together to help cultural projects. In: *Proceedings of the 2nd International Conference on Hypermedia and Interactivity in Museums (ICHIM'93)*, Cambridge, UK, 1993, 13p. Disponível em: <<http://www.portinari.org.br/pps/ppacervo/vg/ichim93.pdf>>. Acesso em: 9 jan. 2007.

LAUDON, K.; LAUDON, J. *Sistemas de informação*. LTC: Rio de Janeiro, 1999. 389p.

LEÃO, Alexandre Cruz. *Gerenciamento de cores para imagens digitais*. 2005. 135 f. Dissertação (mestrado) – Escola de Belas Artes, Universidade Federal de Minas Gerais, Belo Horizonte, 2005. Disponível em: <<http://www.npd.ufrj.br/>>. Acesso em: 18 dez. 2006.

LILLEY, Chris. *PNG (Portable Network Format)*. World Wide Web Consortium, 2003. Disponível em: <http://www.w3.org/Graphics/PNG/>. Acesso em: 5 abr. 2007.

LIMA, C.C. et al. Sistema de obtenção de dados multimídia para a pesquisa Guignard. *Anais do X Brazilian Symposium on Multimedia and the Web – WebMedia [Anais eletrônicos]*, Salvador-BA, ISBN no. 8588442728, 2003, p. 539-542. 1 CD-Rom (Tools and Demos Track)

LIMA, Clarissa Costa. *Especificação de Requisitos de Software Guignard 2.0*. Belo Horizonte, 2007. 68p.

MANNIS, José Augusto. O futuro do passado. *Jornal da UNICAMP*, Campinas, 11 jul. 2005, Sala de Imprensa, v. 294. Disponível em: <http://www.unicamp.br/unicamp/unicamp_hoje/ju/julho2005/ju294pag02.html>. Acesso em: 3 abr. 2007.

MARTINS FILHO, Plínio. Copyright and Internet. *Scielo Brazil*, 1998, v.27, n.2. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651998000200011>. Acesso em: 3 abr. 2007. ISSN 0100-1965

METS. Disponível em: <<http://www.loc.gov/standards/mets/>>. Acesso em: 3 abr. 2007.

MONKEYS AUDIO. Disponível em: <<http://www.monkeysaudio.com/>>. Acesso em: 3 abr. 2007.

MPEG. Disponível em: <<http://www.mpeg.org/MPEG/audio.html>>. Acesso em: 3 abr. 2007.

MySQL. Disponível em: <<http://www.mysql.org/>>. Acesso em: 3 abr. 2007.

NEDLIB. *NEDLIB Contribution to the review of OAIS*. 2000. Disponível em: <http://nedlib.kb.nl/results/OAISreviewbyNEDLIB.html>. Acesso em: 7 abr. 2007

NIELSEN, Jakob. *Usability Engineering*. Academic Press: Boston (MA), 1993.

NISO. Disponível em: <<http://www.niso.org/>>. Acesso em: 3 abr. 2007.

NLA - NATIONAL LIBRARY OF AUTRALIA. *Guidelines for preservation of digital heritage*. Paris: UNESCO, 2003. 177p. Disponível em: <<http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>>. Acesso em: 16 nov. 2006.

OCLC/RLG - Online Computer Library Center/ Research Libraries Group. Disponível em:<<http://www.oclc.org/research/pmwg/>>. Acesso em: 3 abr. 2007.

PÁDUA, Clarindo Isaías Pereira da Silva e. *Engenharia de Usabilidade*. Belo Horizonte, 2007. Disponível em: <<http://www.dcc.ufmg.br/~clarindo/>>. Acesso em: 16 abr. 2007. Apostila (documento em elaboração – uso interno somente),

PANDORA. Disponível em: <<http://pandora.nla.gov.au/index.html>>. Acesso em: 3 abr. 2007.

PAULA, Wilson de Pádua Filho. *Engenharia de Software: Fundamentos, Métodos e Padrões*. LTC: Rio de Janeiro, 2000.

PDF - PORTABLE DOCUMENT FORMAT. Disponível em: <<http://www.adobe.com/br/products/acrobat/adobepdf.html>>. Acesso em: 3 abr. 2007.

PHP. Disponível em: <<http://www.php.net/>>. Acesso em: 3 abr. 2007.

PORTINARI, João Cândido. *Projeto Portinari. Estud. av. [online]*. 2000, v.14, n.38, p.369-400. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142000000100021&lng=pt&nrm=iso>. ISSN 0103-4014. doi: 10.1590/S0103-40142000000100021. Acesso em: 23 fev. 2007.

PRATES, Viviane Soares. *Preservação digital da obra de Guignard em Minas Gerais*. Belo Horizonte: Relatório final de bolsista DTI, 2004. 12p.

PRESMAN, Roger S. *Engenharia de Software*. 6.ed. Trad. Rosângela Delloso Penteado. São Paulo: McGraw-Hill, 2006.

ROCHA, Maldonado Weber. *Qualidade de Software: Teoria e Prática*. São Paulo: Prentice-Hall, 2001.

RODRIGUES, Maria de L. T.. *Preservação Digital de Longo Prazo*. 2003. 206 f. Dissertação (mestrado em Estudos de Informação e Bibliotecas Digitais) – Departamento de Ciências e Tecnologias de Informação, Instituto Superior de Ciências do Trabalho e da Empresa, Lisboa, 2003.

ROTHENBERG, Jeff. An Experiment in Using Emulation to preserve Digital Publications. *RAND-Europe*. Den Haag: Koninklijke Bibliotheek, 2000. NEDLIB Report series; 1. Disponível em: <<http://nedlib.kb.nl/>>. ISBN 90-62-59-1442. Acesso em: 7 dez. 2006.

ROTHENBERG, Jeff. *Avoiding technological quicksand: finding a viable technical foundation for digital preservation*. CLIR - Concil on Library and Information Resources, 1998. 35p. Disponível em: <<http://www.clir.org/pubs/reports/rothenberg/contents.html>>. Acesso em: 2 mar. 2007.

SADASHIGE, Koichi. *Data Storage Technology Assessment - 2002 Projections through 2010*. National Media Laboratory, 2003. 80p. Disponível em: <http://www.imation.com/government/nml/pdfs/AP_NMLdoc_DSTAssessment.pdf>. Acesso em: 20 abr. 2007.

SALDANHA, Leopoldo. FHOXSP. Disponível em: <http://www.fhxosp.com.br/07/materias_01.php>16/março/2007. Acesso em: 29 mar. 2007.

SIMÕES, S. S. Storage: vamos proteger nosso capital mais valioso. *Mundo da Imagem*, São Paulo, n. 51, p.3-9, mai/jun. 2002.

SITTS, Maxine K. (Ed). *Handbook for Digital Projects: A Management Tool for Preservation and Access*. Andover, MA, 2000: Northeast Document Conservation Center, 2000. 182 p. Disponível em: <<http://www.nedcc.org/digital/dman.pdf>>. Acesso em: 11 nov. 2006.

SLATTERY, O.; LU, R.; ZHENG, J.; BYERS, F.; TANG, X. Stability Comparison of Recordable Optical Discs: a study of error rates in harsh conditions. *Journal of Research of the National Institute of Standards and Technology*, Gaithersburg, v.109, n.5, p.517-524, 2004. Disponível em: <http://www.rootsecure.net/content/downloads/pdf/cd_lifetime_study_nist.pdf>. Acesso em: 11 nov. 2006.

ST. LAURENT, Gilles. *Guarda e manuseio de materiais de registro sonoro*. Coord. Ingrid Beck, trad. José Luiz Pedersoli Júnior. Rio de Janeiro: Projeto Conservação Preventiva em Bibliotecas e Arquivos: Arquivo Nacional, 1997.

TANSELLE, Thomas G. *Reproductions and Scholarship*. Disponível em: <<http://etext.virginia.edu/etcbin/toccer-sb?id=sibv042&images=bsuva/sb/images&data=/texts/english/bibliog/SB&tag=public&part=2&division=div>>. Acessado em: 3 mar. 2007.

VALLE JR., Eduardo Alves. *Sistemas de informação multimídia na preservação de acervos permanentes*. Dissertação (mestrado) - Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, Belo Horizonte, 2003. 128p. Disponível em: <<http://www.npdऒdcc.ufmg.br>>. Acesso em Jun. – 2006.

VIEIRA, Ivone Luzia. *Guignard em Minas Gerais: a modernidade de uma obra*. Belo Horizonte, 2000. Disponível em: <<http://www.pesquisaguignard.eba.ufmg.br>>. Acesso em: 10 abr. 2007.

WATERS, Donald; GARRETT, John. *Preserving digital information: report of the Task Force on Archiving of Digital Information* commissioned by the Commission on Preservation and Access and the Research Libraries Group. Washington: Commission on Preservation and Access - CPA, 1996. 71p. Disponível em: <<http://www.rlg.org/legacy/ftp/pub/archtf/final-report.pdf>>. Acesso em: 25 mar. 2007.

WERF-DAVELAAR, Titia. Long-term preservation of electronic publications – The NEDLIB project. *D-Lib Magazine*, v.5, n.9, set. 1999. Disponível em: <<http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/september99/vanderwerf/9vanderwerf.html>>. Acesso em: 15 jul. 2006.

WINDOWS MEDIA AUDIO. Disponível em: <<http://www.microsoft.com/windows/windowsmedia/forpros/codecs/audio.aspx>>. Acesso em: 12 abr. 2007.

XIPH.ORG FOUNDATION . Disponível em: <<http://xiph.org/>>. Acesso em: 11 abr. 2007.

ZANINI, Walter (org). *História geral da arte no Brasil*. São Paulo: Instituto Walter Moreira Salles / Fundação Djalma Guimarães, 1982. 2 v.