

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO**

TAMIE ALINE LANÇA

**MODELO DE REFERÊNCIA PARA DIAGNÓSTICO DA OBSOLESCÊNCIA
TECNOLÓGICA E DESENVOLVIMENTO DE UMA SOLUÇÃO COMPUTACIONAL
VISANDO A PRESERVAÇÃO DIGITAL DO CONHECIMENTO CIENTÍFICO: CASO
BDTD/UFSCAR.**

SÃO CARLOS

2010

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO**

TAMIE ALINE LANÇA

**MODELO DE REFERÊNCIA PARA DIAGNÓSTICO DA OBSOLESCÊNCIA
TECNOLÓGICA E DESENVOLVIMENTO DE UMA SOLUÇÃO COMPUTACIONAL
VISANDO A PRESERVAÇÃO DIGITAL DO CONHECIMENTO CIENTÍFICO: CASO
BDTD/UFSCAR.**

**Trabalho de Conclusão de Curso apresentado
ao Departamento de Ciência da Informação da
Universidade Federal de São Carlos – UFSCar,
como pré-requisito para obtenção do título de
Bacharel em Biblioteconomia e Ciência da
Informação.**

***Orientação: Prof.Dr. Roniberto Morato do
Amaral.***

SÃO CARLOS

2010

FICHA CATALOGRÁFICA

Lança, Tamie Aline

Modelo de referência para diagnóstico da obsolescência tecnológica e desenvolvimento de uma solução computacional visando a preservação digital do conhecimento científico: caso BDTD/UFSCar. / Tamie Aline Lança. – São Carlos: UFSCar, 2010.

64 f.

Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia e Ciência da Informação) – Universidade Federal de São Carlos, 2010.

Acompanha DVD-ROM.

1. Preservação digital. 2. Biblioteca digital aberta e interoperável. 3. Obsolescência tecnológica. I. Título.

DEDICATÓRIA

À minha mãe, que me colocou na escola da vida e na vida da escola. Por tudo que sempre me proporcionou e por fazer o meu mundo melhor sendo espírito de luz, perfeita como ela só. Obrigada por toda força, respeito e principalmente a paciência de sempre.

“Aceite com sabedoria o fato de que o caminho está cheio de contradições. Há momentos de alegria e desespero, confiança e falta de fé, mas vale à pena seguir adiante.”

COELHO (2003)

RESUMO

Na contemporânea sociedade do conhecimento, as universidades são cada vez mais valorizadas por aquilo que produzem e divulgam: ganham credibilidade à medida que externalizam o conhecimento produzido, disponibilizando seu conteúdo na web. Grande parte deste conhecimento hoje se encontra registrado em documentos digitais, devido à facilidade de criação e principalmente de acesso. Com aumento da geração de informação no formato digital e sua rápida obsolescência tecnológica, surge um novo desafio e requisito fundamental neste contexto: a preservação digital. Dada a sua importância para as bibliotecas universitárias, o objetivo geral desta pesquisa é desenvolver uma solução computacional para a preservação digital do conhecimento científico, com base em um diagnóstico da obsolescência tecnológica da Biblioteca Digital de Teses e Dissertações da Biblioteca Comunitária da Universidade Federal de São Carlos (BDTD/UFSCar), objeto de estudo da pesquisa. Pesquisa-ação foi o método utilizado e compreendeu atividades de diagnóstico da obsolescência tecnológica, com base em uma estrutura de requisitos de referência necessários a preservação digital, construída a partir de artigos que versam sobre a temática. Após o diagnóstico foi elaborada uma solução computacional, baseada em software livre para a preservação digital do conhecimento científico produzido pelos Programas de Pós Graduação da UFSCar. Os resultados alcançados foram: 1) a formatação e aplicação de uma estrutura de referência para apoiar o processo de diagnóstico do *status* da obsolescência tecnológica da BDTD/UFSCar; 2) desenvolvimento de uma solução computacional para a preservação digital do conhecimento científico. A estrutura de referência e a solução computacional apresentadas contribuem para as iniciativas de preservação digital nas bibliotecas universitárias, ao disponibilizarem uma sistemática para avaliação da obsolescência tecnológica e uma solução envolvendo o processo de preservação denominado *Refrescamento*. Os resultados garantem que a contribuição cultural, histórica e científica da comunidade científica da UFSCar, expressa cada vez mais por manifestações de documentos digitais, possa ser continuamente acessada, em termos de conteúdo e funcionalidades através dos recursos tecnológicos disponíveis na época, tendo garantida sua integridade, autenticidade e confiabilidade.

Palavras-chave: Preservação digital. Biblioteca digital aberta e interoperável. Obsolescência tecnológica. Repositórios institucionais.

LISTA DE FIGURAS

Figura 1.	Ciclo do conhecimento.....	19
Figura 2.	Contexto das necessidades de preservação digital da informação....	33
Figura 3.	10 mandamentos da preservação digital.....	34
Figura 4.	Ciclo da Pesquisa-ação.....	41
Figura 5.	Tratamento dos dados para aplicação do MX da família ISIS.....	46
Figura 6.	Geração da base de dados com o aplicativo MX.....	47
Figura 7.	Indexação da base de dados no WINISIS.....	47
Figura 8.	Aplicação do software GENISIS para desenvolver as interfaces.....	48
Figura 9.	Implementação de melhorias nas interfaces através do Macromedia Dreamweaver.....	49
Figura 10.	Processo de instalação do DVD-ROM 'BDTD Lingüística'.....	53
Figura 11.	Interface de busca da 'BDTD Lingüística'.....	53
Figura 12.	Interface da lista de resultados na 'BDTD Lingüística'.....	54
Figura 13.	Interface do resultado detalhado na 'BDTD Lingüística'.....	55
Figura 14.	Acesso ao texto completo em PDF na 'BDTD Lingüística'.....	55

LISTA DE QUADROS

Quadro 1.	Matriz com as abordagens e estratégias de preservação propostas por diferentes autores.....	39
Quadro 2.	Processo de desenvolvimento da solução computacional.....	45
Quadro 3.	Da conformidade da BDTD/UFSCar aos requisitos necessários à preservação digital.....	51

LISTA DE SIGLAS

BCo	Biblioteca Comunitária da UFSCar
BDTD	Biblioteca Digital de Teses e Dissertações
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CTC	Comitê Técnico-Consultivo
CONARQ	Conselho Nacional de Arquivos
DSI	Disseminação Seletiva da Informação
DSPACE	Institutional Digital Repository System
Finep	Financiadora de Estudos e Projetos
IBICT	Instituto Brasileiro de Ciência e Tecnologia
ISO	International Organization for Standardization
LMPL	Linguagem de Marcação da Plataforma Lattes18
MEC	Ministério da Educação
METS	Metadata Encoding and Transmission Standards
MTD-BR	Padrão Brasileiro de Metadados para Teses e Dissertações
NDLTD	Networked Digital Library of Theses and Dissertations
OAI	Open Archives Initiative
OAIS	Open Archival Information System
PADI	Preserving Access to Digital Information

LISTA DE SIGLAS

PPG	Programa de Pós-Graduação
PUC - Rio	Pontifícia Universidade Católica do Rio de Janeiro
SEER	Sistema Eletrônico de Editoração de Revistas
SGML	Standard Generalized Markup Language
TEDE	Sistema de Publicações Eletrônica de Teses e Dissertações
TIC	Tecnologias de Informação e Comunicação
UFSC	Universidade Federal de Santa Catarina
UFSCar	Universidade Federal de São Carlos
UNESCO	United Nations Educational, Scientific and Cultural Organization
USP	Universidade de São Paulo
XML	Extensible Markup Language

SUMÁRIO

1. INTRODUÇÃO	11
2. ACESSIBILIDADE E PRESERVAÇÃO DIGITAL	15
2.1 O Papel das Bibliotecas Universitárias	15
2.2 O Conhecimento Gerado nos Programas de Pós-Graduação	17
2.3 Formas de Armazenamento.....	20
2.4 Repositórios Digitais	23
2.4.1 Biblioteca digital aberta e interoperável.....	24
2.4.2 Ferramentas disponibilizadas pelo IBICT	28
2.5 Preservação Digital	29
2.5.1 Preservação estrutural e operacional	34
3. MÉTODO E DESENVOLVIMENTO	39
3.1 Abordagem e método da pesquisa	39
3.2 Desenvolvimento da pesquisa	43
4. RESULTADOS	49
4.1 Formatação e Aplicação de uma Estrutura de Referência para Apoiar o Processo de Diagnóstico do Status da Obsolescência Tecnológica.....	49
4.2 Solução Computacional para a Preservação Digital	54
5. CONSIDERAÇÕES FINAIS	58
REFERÊNCIAS.....	60

1. INTRODUÇÃO

Na contemporânea sociedade do conhecimento, as universidades são cada vez mais valorizadas por aquilo que produzem e divulgam. Ganham credibilidade à medida que externalizam o conhecimento produzido, disponibilizando seu conteúdo na web. Atualmente, grande parte deste conhecimento produzido nas universidades brasileiras é registrada em documentos digitais, devido à facilidade de criação e principalmente de acesso. A iniciativa coordenada pelo Instituto Brasileiro de Informação Científica e Tecnológica (IBICT), denominada de Biblioteca Digital Brasileira de Teses e Dissertações – BDTD, por exemplo, pode ser acessada a partir de qualquer computador conectado a internet e disponibiliza textos completos e os dados referenciais (metadados) dos trabalhos de conclusão de curso de mestrado e doutorado desenvolvidos nas universidades brasileiras consorciadas (BDTD, 2010). Torres e Manzoni (2004) apontam as possibilidades que este novo espaço digital, criado pelas Tecnologias de Informação e Comunicação - TIC proporciona ao atendimento das necessidades de informação da sociedade.

A evolução das TIC exerce profundas influências nas bibliotecas universitárias. A internet, principal ferramenta dessa evolução, disponibiliza à comunidade universitária uma diversidade de mecanismos de interação com o conhecimento em formato digital, proporcionando autonomia aos usuários no seu acesso e compartilhamento do conhecimento científico, além de remover barreiras geográficas e temporais. As TIC aumentam a importância dos documentos digitais, pois são ferramentas responsáveis pela facilidade de armazenamento e acesso.

É crescente a preocupação, por parte de arquivistas e bibliotecários, com a questão da preservação dos documentos digitais. Faz-se necessário trabalhar o problema da preservação dessas fontes de conhecimento nos planejamentos das bibliotecas universitárias. Este processo irá garantir, segundo Sant’anna (2001), a preservação e a acessibilidade das informações contidas em documentos criados no ambiente tecnológico anterior. Tradicionalmente preservar significa manter imutável e intacto, entretanto no ambiente digital preservar significa, na maioria dos casos, mudar, recriar, renovar: mudar formatos, renovar mídias, hardware e software. Por

um lado, deseja-se manter o conhecimento registrado intacto, como ele foi gerado. Por outro, é necessário acessá-lo dinamicamente e com as mais avançadas ferramentas.

Dada a importância da preservação digital para as bibliotecas universitárias, o objetivo geral desta pesquisa é desenvolver uma solução computacional para a preservação digital do conhecimento científico. Este objetivo geral pode ser dividido em objetivos específicos, relativos à fundamentação teórico-conceitual e à pesquisa empírica:

- Diagnóstico da obsolescência tecnológica;
- Seleção de uma solução tecnológica;
- Desenvolvimento da solução visando à preservação digital;
- Divulgação dos resultados.

A pesquisa visa otimizar os recursos utilizados para a divulgação e preservação do conhecimento gerado através dos Programas de Pós-Graduação *Strictu Sensu* – PPG e garantir que a contribuição cultural, histórica e científica da comunidade científica, expressa cada vez mais por manifestações de documentos digitais, possa ser continuamente acessada, em termos de conteúdo e funcionalidades através dos recursos tecnológicos disponíveis na época, tendo garantida sua integridade, autenticidade e confiabilidade.

A estrutura da pesquisa foi dividida em outras cinco seções:

- Seção 2 – Acessibilidade e preservação digital: esta seção descreve o papel das bibliotecas universitárias; aborda a importância do conhecimento gerado nos PPG; formas de armazenamento e acesso à informação nas bibliotecas; repositórios digitais; biblioteca digital e definições e técnicas de preservação digital:
 - 2.1 - Papel das bibliotecas universitárias: nesta seção é realizada uma exposição do papel das bibliotecas universitárias, seus recursos e as mudanças que estão ocorrendo e as que virão a ocorrer devido à evolução das TIC;
 - 2.2 - O conhecimento gerado nos PPG: o valor do conhecimento na contemporânea na sociedade da informação, o papel dos PPG e

importância da disponibilização do conhecimento científico são tratados nesta seção;

- 2.3 Formas de armazenamento – impresso x digital: vantagens e desvantagens dos dois tipos de suporte de informação.
- 2.4 Repositórios digitais: aborda a temática dos repositórios institucionais e digitais que utilizam a tecnologia de arquivos abertos (*Open Archives Initiative - OAI*), responsabilidades adquiridas a se construir uma biblioteca digital, objetos digitais e ferramentas disponibilizadas pelo IBICT:
 - 2.4.1 Biblioteca Digital aberta e interoperável: são discutidos a estrutura necessária para seu desenvolvimento, níveis de interoperabilidade, pessoas e sistema (padrões, normas, formatos, protocolos e metadados);
 - 2.4.2 Ferramentas disponibilizadas pelo IBICT: apresentação de ferramentas que auxiliam no processo de geração de repositórios digitais, como DSpace (Institutional Digital Repository System), Serviço Eletrônico de Editoração de Revistas - SERR e o Teses e Dissertações Eletrônicas - TEDE utilizado para se desenvolver bibliotecas digitais.
- 2.5 Preservação digital: são temas discutidos nessa seção a responsabilidade da preservação digital nos quesitos acesso, autenticidade e integridade da informação, questões sociais e as estratégias estruturais e operacionais de preservação digital.
- Seção 3 – Método e desenvolvimento: é discutido a escolha do método pesquisa-ação e o objeto de estudo da pesquisa, a Biblioteca Comunitária da Universidade Federal de São Carlos – BCo/UFSCar. Também é apresentado o desenvolvimento de uma solução tecnológica para a preservação digital.
- Seção 4 – Resultados: apresentação e discussão dos resultados obtidos: [1] Formatação e aplicação de uma estrutura de referência para apoiar o processo de diagnóstico do *status* da obsolescência tecnológica da biblioteca digital implantada pela organização estudada; e [2] Desenvolvimento da

solução computacional para a preservação digital do conhecimento científico gerado pelos PPG da UFSCar.

- Seção 5 – Conclusão: concluí-se que a estrutura de referência para o diagnóstico da obsolescência tecnológica e a solução desenvolvida são de fácil aplicação e baixo custo e não há necessidade de trabalhos terceirizados. O modelo de referencia pode ser aplicado em qualquer tipologia de biblioteca e permite o acompanhamento da evolução do *status* obsolescência tecnológica do repositório digital. A solução computacional promove a liberação do espaço físico da biblioteca, provem o acesso nos períodos de instabilidade da rede e do servidor e adiciona uma rotina mais segura de *backup*.

2. ACESSIBILIDADE E PRESERVAÇÃO DIGITAL

Recentemente, Lawrence demonstrou que artigos disponíveis *on-line* são mais citados do que os não disponíveis. A disponibilidade de trabalhos acadêmicos em bibliotecas digitais beneficia os pesquisadores acadêmicos. Este benefício, no entanto, depende do nível de disponibilidade e facilidade de acesso oferecidos. (LAWRENCE apud PACHECO e KERN, 2001, p.65).

2.1 O Papel das Bibliotecas Universitárias

As bibliotecas universitárias fornecem insumo para a produção de novos conhecimentos e armazenam o conhecimento produzido a partir dele. São peças-chave do processo de criação e gestão do conhecimento, centro da produção científica. Elas estão inseridas num contexto maior, que é a universidade. Nos últimos anos ocorreram diversas mudanças estruturais e financeiras nas universidades que conseqüentemente refletiram em suas bibliotecas.

Independentemente do suporte da informação, a biblioteca foi durante séculos e ainda é o principal local de preservação e disponibilização do conhecimento (CUNHA, 2000). No contexto atual, em plena revolução das TIC, onde esse conhecimento existe em diversos formatos e é distribuído na rede mundial, ele deixa de ser privilégio de poucos dentro da universidade. A universidade pública vem se preocupando, nos últimos anos, com a fixação de metas voltadas para a consolidação de possíveis aplicações de TIC na divulgação do conhecimento por elas gerado, no ambiente acadêmico e na sociedade brasileira. Com o crescimento do ciberespaço, aumenta a importância das bibliotecas digitais, que facilitam aos estudantes das modalidades de ensino a distância e presencial o acesso às informações necessárias ao seu desenvolvimento acadêmico e profissional.

Morigi e Pavan (2004), afirmam que “as bibliotecas como instituições sociais são partes integrantes da sociedade. Como tais, também acompanham os processos de desenvolvimento econômico, social e tecnológico.” O acervo dessas

unidades deve ser tratado, armazenado e disseminado de tal maneira que satisfaça as necessidades de seus usuários e cumpra o objetivo da instituição mantenedora. As TIC influenciam toda a sociedade e com as bibliotecas universitárias não tem sido diferente, pois elas vêm incorporando as TIC em suas práticas cotidianas e grande parte já disponibiliza informações armazenadas em suporte eletrônico além da versão impressa. Segundo Morigi e Pavan (2004), isso confere um caráter “híbrido” ao acervo. O suporte impresso então coexiste com o suporte eletrônico, pois o segundo facilita o acesso a informação, mas nem sempre substitui o primeiro. Muitas obras procuradas nas bibliotecas não se encontram digitalizadas e talvez jamais sejam. Em contrapartida muitos periódicos começam a ser publicados/disponibilizados somente na versão eletrônica.

Em todas as épocas, bibliotecas sempre foram dependentes da tecnologia da informação. A passagem dos manuscritos para a utilização de textos impressos, o acesso a base de dados bibliográficos armazenados nos grandes bancos de dados, o uso do CD-ROM e o advento da biblioteca digital, no final dos anos 90, altamente dependente das diversas tecnologias de informação, demonstram que, nos últimos 150 anos, as bibliotecas sempre acompanharam e venceram os paradigmas tecnológicos (CUNHA, 2000, p.65).

Com o uso crescente da Internet, atriz principal da evolução das TIC, como fonte de informação, a Disseminação Seletiva da Informação - DSI passa a ter um papel crucial, pois um indivíduo sozinho jamais será capaz de acompanhar o enorme volume de dados que são disponibilizados diariamente. Ela é necessária para que haja uma seleção das informações de acordo com necessidades e interesses dos usuários. De acordo com Cunha (2000), a DSI não ficará restrita aos documentos tradicionais, livros e artigos. Poderá colaborar com outras necessidades informacionais diárias de seus usuários, incluindo noticiários, jornais, rádio, mercado de valores, programações cinematográficas entre outros.

“A transmissão eletrônica da informação dá novo sentido à biblioteca, cujo propósito é tornar o conhecimento acessível aos usuários finais, integrando múltiplas tecnologias disponíveis...” (DRABENSTOTT, BURMAN e MACEDO, 1997).

Obstáculos cada vez maiores colocados sobre a alocação de recursos financeiros para as universidades, segundo Cunha (2000), criam o seguinte dilema: crescem e são mais valorizadas as áreas do conhecimento úteis ao mercado global. Estas têm mais prestígio dentro do campus enquanto as humanidades, incluindo aí as bibliotecas, ficam em desvantagem. Bibliotecas são centros de custos e possuem dificuldades na obtenção de recursos externos a universidade. Cunha (2000) ainda alerta para o fato de que a universidade pode se tornar um mero instrumento da política econômica nacional.

Segundo Vicentini (1997), as bibliotecas universitárias deverão estar preparadas para mudanças, redimensionando seus espaços, trabalhos, serviços e produtos e acompanhando a evolução tecnológica disponível, principalmente voltada para um usuário cada vez mais conhecedor de tecnologias de comunicação de dados.

O esforço para recuperar as funções básicas da biblioteca universitária necessitará compreender que, de fato, mais e mais usuários estão resolvendo suas demandas informacionais por meio do ciberespaço. Além disso, a criação de acervo digital será um dos caminhos a serem trilhados com a re-alocação de recursos para projetos colaborativos (CUNHA, 1997, p.267).

2.2 O Conhecimento Gerado nos Programas de Pós-Graduação

A biblioteca universitária é um sistema de informação que é parte de um sistema de informação acadêmico, onde a geração do conhecimento é o objetivo da “vida” universitária (FUJITA, 2005). Um ciclo do conhecimento (Figura 1) é formado pela geração do conhecimento, que é armazenado, organizado e divulgado, e servirá de insumo para geração de novos conhecimentos, que por sua vez serão

armazenados, organizados e divulgados e assim sucessivamente. A universidade atua como geradora, transmissora e receptora de conhecimentos.

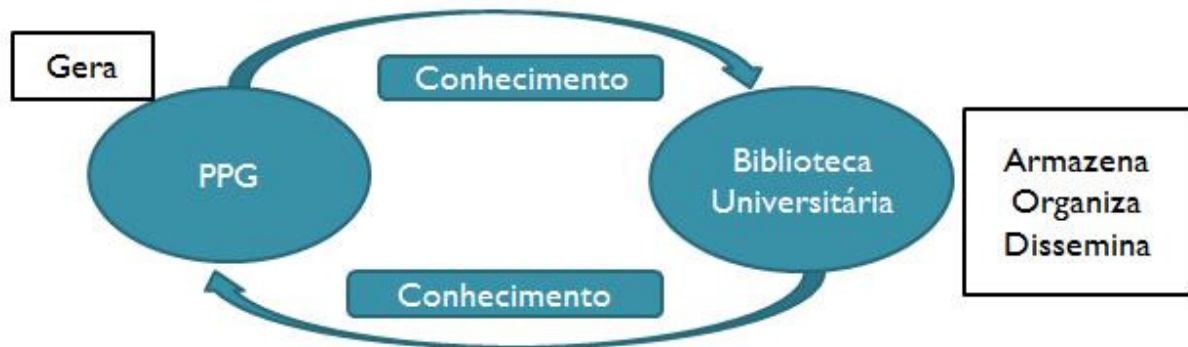


Figura 1 – Ciclo do conhecimento.

Fonte: Adaptado de Fujita (2005).

Para Cunha (2000), hoje o conhecimento tem valor financeiro. A economia está cada vez mais ligada à produção, tratamento, distribuição e uso da informação. Os PPG constituem uma importante peça nesse contexto. Eles devem apresentar originalidade para serem bem avaliados e bem sucedidos. Novas tecnologias e os mais importantes estudos e descobertas em todas as áreas do conhecimento advêm das pesquisas realizadas pelos PPG e são orientadas e validadas por docentes credenciados. Os principais objetivos dos PPG são a produção de conhecimento e a formação de pesquisadores criativos: atuam como formadores de cientistas e professores de nível superior que tem como dever garantir ao país um bom potencial de aprendizagem, produção de conhecimento e tecnologia que beneficiem toda a sociedade (BOTOMÉ e KUBO, 2002).

Botomé e Kubo (2002) ressaltam que “a responsabilidade específica dos mestrados e doutorados é desenvolver capacidade de pessoas para produzirem conhecimento e torná-lo acessível à sociedade”. O uso do conhecimento novo depende da capacidade dos cientistas de educar a sociedade para tal.

Vários fatores contribuem para que os alunos de pós-graduação stricto sensu (mestrado e doutorado) alcancem esse objetivo. Segundo instrumento de pesquisa desenvolvido por Alencar (1995) apud Alencar (2001) para investigar a

percepção de pós-graduandos quanto ao estímulo a criatividade, constatou-se os seguintes fatores: turmas menores que as de graduação, maior estimulação a curiosidade, mais tempo aos alunos para pensar e desenvolver novos pontos de vista e incentivo a desenvolver questões sobre o assunto. O professor também exerce um importante papel nesse contexto universitário: deve deixar de ser apenas transmissor do conhecimento e passar a ser estimulador da criação, da inovação, favorecendo as habilidades criativas dos pós-graduandos. De acordo com Alencar (2001), cultivar o gosto pela descoberta e pela busca de novos conhecimentos, motivar os alunos a pensar e a raciocinar, estimular a produção de novas idéias, desenvolver habilidades de análise crítica e valorizar as idéias dos alunos certamente contribuem para a formação do pesquisador.

Ao disponibilizar sua produção por meio de bibliotecas digitais bem aparelhadas, a pós-graduação brasileira tem a oportunidade de expor a qualidade do que produz para o grande público e, assim, promover sua integração com o restante da sociedade (PACHECO e KERN, 2001, p.65).

As bibliotecas digitais de teses e dissertações, cujos acervos digitais em rede são o testemunho das atividades acadêmicas e de pesquisa da universidade, contribuem para que a universidade seja mais valorizada, agregando maior projeção nacional e internacional a sua contribuição. Desta forma as bibliotecas digitais se tornam cada vez mais um elo importante na preservação dos estoques de conhecimento no formato digital, que constituem testemunho das atividades dos PPG da universidade a qual a biblioteca esta inserida (MARCONDES e SAYÃO, 2001 apud AMARAL, 2009).

2.3 Formas de Armazenamento

Atualmente os formatos impresso e digital coexistem e sua coexistência é completamente necessária e coerente, pois apesar da grande facilidade de acesso às informações digitais, grande parte da informação ainda não está disponível neste formato e talvez nunca sejam digitalizadas, pois não são consideradas prioritárias (DRABENSTOTT, BURMAN e MACEDO, 1997). Existem pontos positivos e negativos, facilidades e também restrições em cada tipo de formato, mas o que realmente importa são o desempenho e a contribuição de cada um deles no acesso à informação. Segundo Drabenstott, Burman e Macedo (1997), o princípio orientador é usar a tecnologia apropriada para cada propósito particular.

De um lado temos o formato impresso, que tem um alto custo de aquisição, esta que ademais pode ser mal planejada e ineficaz, gerando ainda mais gastos desnecessários. A demora no recebimento do material adquirido também é agravante em muitas bibliotecas. O volume de exemplares impressos ocupa espaço em estantes físicas, que devem ser adquiridas e mantidas. Portanto há necessidade contínua de aquisição de materiais, estantes e espaço para armazená-los. Existem ainda despesas com materiais para reparos e encadernações dos exemplares, pois há deterioração devido ao uso e vandalismo por parte de alguns usuários. A deterioração também acontece no edifício da biblioteca. E não menos importante, deve-se manter um espaço físico adequado para os usuários da biblioteca tradicional (CUNHA, 1997). Porém entre estas questões, talvez a mais importante seja a explosão bibliográfica que tornou quase impossível adquirir e encontrar espaço físico para atender todo o interesse dos usuários.

O que interessa é informação nova. Para Hawkins (1994) apud Cunha (2000), torna-se cada vez mais difícil para a biblioteca prover acesso à totalidade da informação demandada por seus usuários. Uma das exigências do mundo globalizado é a maior agilidade de acesso às informações através de vários mecanismos. Deste lado temos então o formato digital, que permite a busca e consulta de dados em catálogos online. Ele é perfeitamente apropriado à contemporânea sociedade do conhecimento: informações muito mais acessíveis, a

um preço de aquisição e conservação muito mais razoável e com todas as vantagens dos novos recursos que a WEB 2.0 proporciona (MANESS, 2007).

No quesito acessibilidade, o formato digital possibilita acesso universal, atemporal e simultâneo a diversos usuários, por meio de computadores em rede. Com isso, “há disponibilidade infinita das coleções, conteúdos e arquivos transferíveis quantas vezes for necessário” (DRANBENSTOTT, BURMAN e MACEDO, 1997). A informação acadêmica é gerada cada vez mais rapidamente, o que cria a necessidade de mecanismos mais rápidos e eficientes para administrá-la e publicá-la e para que os usuários possam consultá-la no menor espaço de tempo possível. Tammaro e Salarelli (2008) dão as seguintes características ao documento digital: flexibilidade, simulação, reprodutibilidade, conservação e transmissibilidade. A versão digital é capaz de gerar cobertura universal do conhecimento, atendendo mais completamente as necessidades dos pesquisadores.

Os periódicos estão cada vez mais sendo disponibilizados e adotados pelas bibliotecas tradicionais na versão digital, devido ao alto custo do material impresso e o largo espaço que esse material ocupa no acervo físico. Outra vantagem desse hábito é que se evita artigos repetitivos: a agilidade e facilidade da disponibilização, atualização e revisão na versão eletrônica possibilitam aos estudiosos publicarem mais e mais rapidamente. Isso faz com que as revistas eletrônicas reduzam a proliferação de publicações supérfluas. Os periódicos impressos estão sendo substituído pelos artigos eletrônicos disponíveis nas bases de dados online.

Na biblioteca digital não há necessidade de espaço físico como a tradicional. É necessário somente um espaço com equipamentos e pessoal profissionalizado para mantê-la, o que elimina gastos com materiais de conservação do acervo, obras de expansão e conservação do edifício e também a necessidade de manter ambientes adequados aos usuários. De acordo com Lastres (1999) a conversão para a forma digital das diferentes formas de informação impressa e gravada (sonora e/ou visual), oferece a possibilidade de utilizar uma quantidade mínima de matéria e assim reduzir sua alta dependência dos custos associados tanto ao consumo de recursos físicos, quanto a tempo e espaço. Salienta ainda que as TIC permitem tratar as formas visuais e sonoras quase da mesma maneira que a

informação escrita, o que levará a transformações ainda mais radicais sobre a atual posição relativa da informação escrita, não no sentido de sua superação, mas sim na progressão do entrelaçamento entre todas as formas. Segundo Tammaro e Salarelli (2008), um dos pontos fortes da versão digital é que pode ser formalmente manipulada, desmontada e remontada, situação ilustrada da seguinte forma pela autora:

Cortar um livro antigo a tesouradas é reprovável. Recortar um documento digital é aplaudível, como sugere um ícone colocado bem à vista na barra de ferramentas do meu processador de texto (TAMMARO e SALARELLI, 2008, p.13).

Bax (1997) define bibliotecas digitais como entidades capazes de vencer as limitações naturais e espaço-temporais impostas a objetos físicos (livros, estantes, salas, prédios), permitindo novas práticas de trabalho e oportunidades. A biblioteca digital, segundo Dranbenstott, Burman e Macedo (1997), prioriza o acesso ao invés do acúmulo de coleções, atuam sobre o novo paradigma '*access over ownership*'. Nessa mudança, acredita-se que o acervo de uma biblioteca deve ser definido pelo acesso à informação, não pelo volume. O interessante são informações novas e integridade da informação significa preservação por longo tempo.

Algumas vantagens da biblioteca digital sobre a tradicional podem ser identificadas, como por exemplo, na maioria das vezes a versão digital é uma opção mais econômica em termos de custo e também sai na frente no quesito rapidez de acesso. Existem recursos que os documentos impressos jamais serão capazes de contemplar. Hipertexto, integração de diversos formatos, como fotos, imagens virtuais em movimento, informações por satélites, sons, diversos tipos de formatação e interação, maior autonomia do usuário e ainda a comunicação informal que acontece nas redes, uma das formas mais importantes de comunicação e fonte de informação da atualidade (CUNHA, 2000).

O surgimento de novas tecnologias não implica que as anteriores devam ser abolidas. Esse argumento é usado por autores que se asseguram de que

as duas versões, impressa e digital, coexistirão ainda por muito tempo. Baseiam essa teoria no fato de que há altíssima quantidade de material impresso e a conversão desse material para a forma digital ainda é pouco realizada. Dranbenstott e Burman (1997) colocaram que ainda não existiam instrumentos cabíveis para esse tipo de conversão e que se isso fosse possível, seriam priorizados os materiais com mais alta demanda, enquanto os de baixa demanda conservariam a importância da biblioteca tradicional (forma impressa). Hoje já existem alguns projetos de digitalização no Brasil. A Universidade de São Paulo (USP) adquiriu, por R\$ 1,5 milhão, dinheiro conseguido com verba da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), um scanner robotizado para digitalizar as obras doadas por José Mindlin à biblioteca Brasileira USP. A máquina terá a capacidade de digitalizar 2500 imagens por hora. O acervo a ser digitalizado tem 30 mil volumes e, com a ajuda do scanner, será disponibilizado gratuitamente na internet (KATO, 2009).

É necessário que os usuários das bibliotecas digitais possuam competências para navegação; o usuário deve saber principalmente o que quer buscar. A infra-estrutura de padrões e normas também deve ser desenvolvida para que o intercâmbio entre geradores e receptores das informações seja eficiente (CUNHA, 2000).

2.4 Repositórios Digitais

De acordo com Souza (1997), “vivemos um momento de exaltação a palavra digital no qual a meta é prover acesso às publicações eletrônicas”. Isto ligado a utilização cada vez mais intensa das TIC, leva ao surgimento das bibliotecas digitais e a novas formas de sociabilidade entre profissionais e usuários da informação. Um novo sentido é dado às bibliotecas com a transmissão digital da informação, cujo objetivo é integrar múltiplas tecnologias disponíveis tornando o conhecimento acessível aos usuários finais e fazendo com que o acesso a

informação seja compatível com o modo de vida da população, cada vez mais digital, eliminando assim barreiras de tempo e espaço.

2.4.1 Biblioteca digital aberta e interoperável

Biblioteca digital é um repositório de informações em formato digital que é composto por um sistema aberto, de múltiplas interligações e múltiplos subsistemas, envolvendo um ambiente organizacional, profissionais especializados, recursos informacionais, usuários, tecnologia de informação, padrões e protocolos. As bibliotecas digitais surgiram na década de 90 e as pesquisas nos diferentes, porém complementares, aspectos da área (projeto, implementação, desenvolvimento e avaliação) passam por um expressivo crescimento (SAYÃO, 2007).

Dranbenstott, Burman e Macedo (1997) complementam que ao se levar em conta outras características e mecanismos do que se denomina biblioteca digital, encontram-se termos complementares, tais como acessibilidade universal, conexão eletrônica, transparência das informações, uso de computadores pessoais e portáteis, instituições, entre outros. Ao se considerar uma biblioteca digital como objeto de estudo, deve-se enxergar os vários aspectos que a engloba: arquiteturas, sistemas, ferramentas e tecnologias; conteúdos e coleções digitais; metadados; interoperabilidade; normas e padrões; organização do conhecimento; usuários e usabilidade; aspecto legal, organizacional, econômico e social (SAYÃO, 2007).

As bibliotecas digitais podem ser repositórios temáticos ou institucionais, como a BDTD/UFSCar, objeto de estudo deste trabalho. Repositórios institucionais são sistemas de informação que servem para armazenar, preservar e difundir a produção intelectual de uma instituição. Elas contribuem para aumentar o impacto das pesquisas desenvolvidas na instituição, aumentando a sua visibilidade e acessibilidade, facilitando a gestão da informação de sua produção científica e preservando sua memória intelectual em suporte digital (RCAAP, [s.d.]).

A disponibilização do conhecimento científico em “arquivos abertos” na Internet permite a qualquer utilizador ler, descarregar (*download*), copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral dos documentos digitais sem

barreiras financeiras, legais ou técnicas. Deve-se compreender que o compartilhamento da informação e dos recursos não é mais uma opção, é uma necessidade. A implementação de bibliotecas digitais abertas e interoperáveis contribui para “romper o círculo vicioso que oculta à produção científica nacional da sociedade e, muitas vezes, da própria comunidade científica” (PACHECO E KERN 2001).

Bibliotecas digitais com “arquivos abertos” ampliam a disseminação dos trabalhos científicos, introduzindo inovações no processo de construção do conhecimento: geram facilidade e diminuição no custo de transmissão. Com a criação dos repositórios institucionais abertos, tornou-se possível o auto-arquivamento da produção científica e os pesquisadores passaram a dispor tanto de uma ferramenta de difusão de seus trabalhos como também de um veículo de comunicação com seus pares, conhecidos e desconhecidos (ARELLANO, 2003).

As bibliotecas digitais abrigam itens chamados na literatura de ‘objetos digitais’. Ferreira (2007) define objetos digitais como qualquer objeto de informação que possa ser representado através de uma seqüência de dígitos binários. Eles podem ser de dois tipos: objetos produzidos como representação/substitutos de materiais em alguma forma analógica ou objetos originalmente digitais – fotografias digitais, *e-books*, bases de dados científicas, web sites, etc. Um objeto digital pode estar completo em um único arquivo ou pode ser constituído de múltiplos arquivos vinculados por hiperlinks ou ainda consistir de múltiplos arquivos unificados por metadados estruturais (SAYÃO, 2007).

Objetos digitais adequados seguem alguns princípios. Devem manter as qualidades de interoperabilidade e reusabilidade; permanecer acessível independentemente das mudanças tecnológicas; ser digitalizado de acordo com padrões e melhores práticas; ser designado por um identificador único e persistente; deve apresentar autenticidade, integridade, proveniência e contexto. Podem ainda ser de vários formatos, como texto, áudio, imagem, vídeo. Os textos digitais são os mais comuns nas bibliotecas digitais de teses e dissertações, objeto de estudo neste trabalho. A codificação de caracteres acontece através de um código computacional para cada letra do documento, o que ocorre no momento da sua criação (SAYÃO, 2007). A importância dos padrões para que esse texto digital se torne interoperável é

inegável. Existe uma codificação universal padronizada pela ISO (*International Organization for Standardization*), o Unicode. Essa codificação fornece um número para cada caractere, independente do programa, plataforma ou da língua, permitindo que as escritas possam ser codificadas sem ambigüidades (ISO, 2006 apud SAYÃO, 2007).

O pleno compartilhamento e disseminação desses objetos digitais denomina-se interoperabilidade. Pode ser entendida, segundo Miller (1998) apud Sayão (2007) como o processo contínuo de assegurar o intercâmbio e reuso das informações digitais. Devem ser observados os diversos aspectos de interoperabilidade: semântica, que é a adoção de ferramentas comuns; política/humana, equipe responsável tornar a informação amplamente disponível; legalidade, implicações legais de tornar os itens livremente disponíveis; e internacional, contornando a diversidade de padrões e normas além das barreiras lingüísticas. A interoperabilidade possui níveis. O mais amplo, nível de *federação*, permite a forma mais potente de interoperabilidade e também gera maior custo para os participantes. Este nível permite, através de uma única interface, o acesso uniforme a uma variedade de fontes de informação. O grau de dificuldade apresentado pelo nível de *federação* fez com que fosse desenvolvido um nível menos complexo, adotado por mais organizações, chamado de *harvesting*. Esse nível consiste de uma colheita automática de metadados estabelecido pelo OAI-PHM² no compartilhamento de alguns serviços básicos (SAYÃO, 2007).

A interoperabilidade entre bibliotecas digitais depende da definição de padrões que permitam o intercâmbio de informações. Eles estabelecem regras pelas quais os objetos são descritos, identificados e preservados e os sistemas aos quais estão inseridos se comuniquem (SAYÃO, 2007). A busca por interoperabilidade entre arquivos abertos, segundo Sena (2000), “visa a transformar cada um dos arquivos em parte de um arquivo global para a realização de pesquisas on-line”.

Usando as definições de ISO (2006) apud Sayão (2007), entendem-se *padrões* como documentos que provêm uso comum, regras, diretrizes ou características de natureza não obrigatória. *Protocolos* são conjuntos de padrões que contém as regras que determinam as funções de comunicação numa rede:

descrevem o formato da mensagem e a maneira pela qual ela deve ser transmitida. Segundo Tamaro e Salarelli (2008), eles estruturam um vínculo ativo entre um ponto e outro no universo documental. Já *normas* têm natureza obrigatória e visam garantir a qualidade dos produtos e serviços, segurança e uniformidade. Por último temos o *formato*, configuração na qual a informação digital pode ser armazenada. Sayão (2007) ressalta que o objetivo é que dois ou mais diferentes sistemas possam ‘compreender’ um determinado objeto digital por meio da unificação da representação, manipulação e transmissão, constituindo a base da interoperabilidade. No Brasil tem-se o exemplo da ontologia LMPL (Linguagem de Marcação da Plataforma Lattes¹⁸) para informações sobre currículos, que consiste de um vocabulário e gramática definidos pelo consenso de nove universidades, prestando-se à integração de aplicativos e bases de dados nesta área.

Outra estrutura necessária à interoperabilidade são os metadados. Metadados são informações estruturadas que descrevem, explicam, localizam, ou ainda possibilitam que um recurso informacional seja fácil de recuperar, usar ou gerenciar. O termo significa dados sobre dados ou mesmo informação sobre informação (NISO, 2004 apud SAYÃO, 2007). Nas bibliotecas digitais os metadados mais usados são os descritivos, que incluem elementos como título, autor, palavras-chave e resumo. O MARC21 é um formato bibliográfico de metadados que é a base da maioria dos catálogos usados em bibliotecas, formando assim mais uma estrutura padronizada, favorecendo a interoperabilidade.

Enquanto as linguagens SGML (*Standard Generalized Markup Language*) e XML (*Extensible Markup Language*) são consensuais quanto linguagens de marcação, ainda existe mais uma preocupação em relação aos textos digitais: aparência e *layout*. Existe um acordo quanto a padronização acerca da utilização do formato PDF (*Portable Document Interface*), da empresa Adobe, gerado a partir do software Acrobat, destinado a criação, edição e exibição de textos digitais. Apesar de pertencer a uma empresa privada, sua especificação se encontra aberta e disponível (SAYÃO, 2007).

As instituições mantenedoras das bibliotecas digitais tem como responsabilidade, além de garantir o acesso, a integridade e a interoperabilidade do objeto digital, a preservação a longo prazo desse material. Existe um

consenso de que, na medida em que os criadores, disseminadores e proprietários de informação digital aceitam a responsabilidade inicial de arquivarem seus objetos digitais, terão também que incorporar preservação digital nos seus processos (SAYÃO, 2007, p.32).

Enfim, para atender às necessidades dos usuários, os fundadores da biblioteca digital devem cumprir alguns atributos, como a compilação com o formato OAI, sustentabilidade tecnológica e financeira, sistemas de segurança e responsabilidades administrativas e processuais. Cumprirão assim duas tarefas expostas por Cunha (1997), que são objetivos das bibliotecas digitais: estabelecimento de um repositório de material eletrônico para a pesquisa e implementação de mecanismos para a sua plena utilização.

2.4.2 Ferramentas disponibilizadas pelo IBICT

O IBICT (2010) disponibiliza ferramentas que possibilitam disseminação, compartilhamento e permitem a interoperabilidade da informação digital. Entre eles existem o DSpace, desenvolvido para possibilitar a criação de repositórios digitais com funções de captura, distribuição e preservação da produção intelectual, permitindo sua adoção por outras instituições em forma consorciada federada, gerenciando a produção científica em qualquer tipo de material digital. (www.dspace.com.br). O SEER, software desenvolvido para a construção e gestão de uma publicação periódica eletrônica e que contempla ações essenciais à

automação das atividades de editoração de periódicos científicos, é outro exemplo de ferramenta disponibilizada pelo IBICT. Ambas geram maior visibilidade e garantem a acessibilidade da informação ao longo do tempo (IBICT, 2010).

Outra ferramenta disponibilizada é o Sistema TEDE. Esse sistema tem como objetivo proporcionar a implantação de bibliotecas digitais de teses e dissertações nas instituições de ensino e pesquisa e sua integração à Biblioteca Digital de Teses e Dissertações – BDTD nacional. Distribuído gratuitamente, em um pacote contendo o Sistema TEDE já implementado, o MTD-BR (Padrão Brasileiro de Metadados para Teses e Dissertações) e a camada do Protocolo OAI-PMH, metodologia de implantação, manuais operacionais e de usuário, documentação e treinamento (IBICT, 2010).

A busca pela interoperabilidade levou à criação e adoção de padrões tecnológicos abertos que permitissem o intercâmbio de informações entre sistemas de informações. A preservação digital pode ser facilitada pela adoção, em larga escala, de padrões tecnológicos abertos para criação, armazenamento e recuperação de documentos digitais (SANT'ANNA, 2001). Em 1997, Conway já se questionava: as bibliotecas digitais virão a se transformar em algo de tão alta qualidade e tão extraordinário valor para as universidades, para o país ou outra instituição que cada participante assumirá o compromisso de colaborar com a preservação? (CONWAY, 1997).

2.5 Preservação Digital

A tecnologia da escrita passou por uma constante evolução desde a antiguidade até o presente. Hoje somos capazes de armazenar imagens de centenas de livros, compostas de minúsculos fragmentos no limitado espaço de um disco ótico. Essa capacidade de registrar e armazenar informações dá origem ao seguinte dilema: nossa capacidade de registrar informações aumentou continuamente enquanto a longevidade do suporte decresceu de modo equivalente.

Observa-se, por exemplo, que os papiros egípcios remontam aproximadamente 4500 anos e ainda são legíveis enquanto as fitas magnéticas tornam-se ilegíveis em apenas 30 anos (CONWAY, 1997). Segundo Lastres (1999), novas formas de gestão do tempo subordinam-se ao tempo tecnológico e dado o caráter praticamente imaterial do documento digital, sua duração é quase igual a sua processabilidade e limitada pela obsolescência tecnológica.

Gracio e Fadel (2008) expõem a falta de conhecimento quando o assunto é preservação digital, além de pessoal com preparação nas instituições de ensino superior. Observaram que o tema é um assunto novo na maioria das instituições que tratam a informação científica e tecnológica, e apontam para necessidade de conexão entre as instituições com o objetivo de tornar mínimo os esforços de preservação e possibilitar uma melhor interoperabilidade dos objetos digitais.

Reprodutibilidade e conservação são características dos documentos digitais e são intimamente ligadas: a sobrevivência do documento digital é diretamente proporcional a sua reprodutibilidade e inversamente proporcional à instabilidade dos suportes (TAMMARO e SALARELLI, 2008).

Sant'ana (2001) coloca que desde o início da utilização em massa de computadores nas organizações, a preocupação com a preservação dos documentos digitais resumia-se em ter um plano de cópias de segurança (*backup*). Se houvesse problema na cópia mais atualizada, recorria-se a segunda mais atualizada. Cada problema encontrado significava um acréscimo no esforço de reconstrução dos registros perdidos. Essa reconstrução normalmente baseava-se nos registros armazenados paralelamente em outro suporte, normalmente papel.

Com o crescimento da importância e aumento da criação de informação no formato digital e sua rápida obsolescência tecnológica, surge um novo desafio e requisito fundamental neste contexto: a preservação digital. O planejamento e a aplicação de métodos e de tecnologias de preservação são necessários para assegurar que a informação digital tenha um valor contínuo, remanescente, acessível e usável. Preservação digital é definida então, segundo Ferreira (2007) como a

capacidade de garantir que a informação digital permaneça acessível e com qualidade e autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação (FERREIRA, 2007, p.20).

Conway (1997) já definia preservação digital como aquisição, organização e distribuição de recursos a fim de que venham impedir posterior deterioração ou renovação da possibilidade de utilização dos objetos digitais. Para ele a essência da gestão da preservação está na alocação de recursos. Pessoas, recursos e materiais devem ser requeridos e organizados para assegurar a proteção adequada às fontes de informação. Para Arellano (2004), a questão da preservação é muito mais um problema social e institucional do que um problema técnico, porque depende de instituições que passam por mudanças de direção, missão, administração e fontes de financiamento. Essas instituições ainda levam em consideração aspectos legais e culturais e adotam determinada infra-estrutura tecnológica.

Uma política de preservação digital deve definir um conjunto de propriedades e ações que preservem cada tipo de objeto digital e o conjunto de suas características, seu ciclo de vida e sua autenticidade, garantindo a perenidade do seu conteúdo (GRACIO e FADEL, 2008). Ela deve envolver, segundo Ferreira (2006), todos os aspectos de um objeto digital, como criação de uma política de avaliação e seleção do material, definição de metadados de preservação, estratégias para cada classe de objeto, política de continuidade, financiamento sustentável, enfim, objetivos a nível social e organizacional.

A preservação digital começa na criação do objeto digital e engloba questões de recursos, pessoas, tecnologia e adequação de normas e padrões, atentando para os aspectos legais e tendo como compromisso a garantia de autenticidade, integridade, confiabilidade e plena acessibilidade a informação digital a longo prazo. Arellano (2006) propõe a Figura 2 para representar o contexto de uma biblioteca digital.

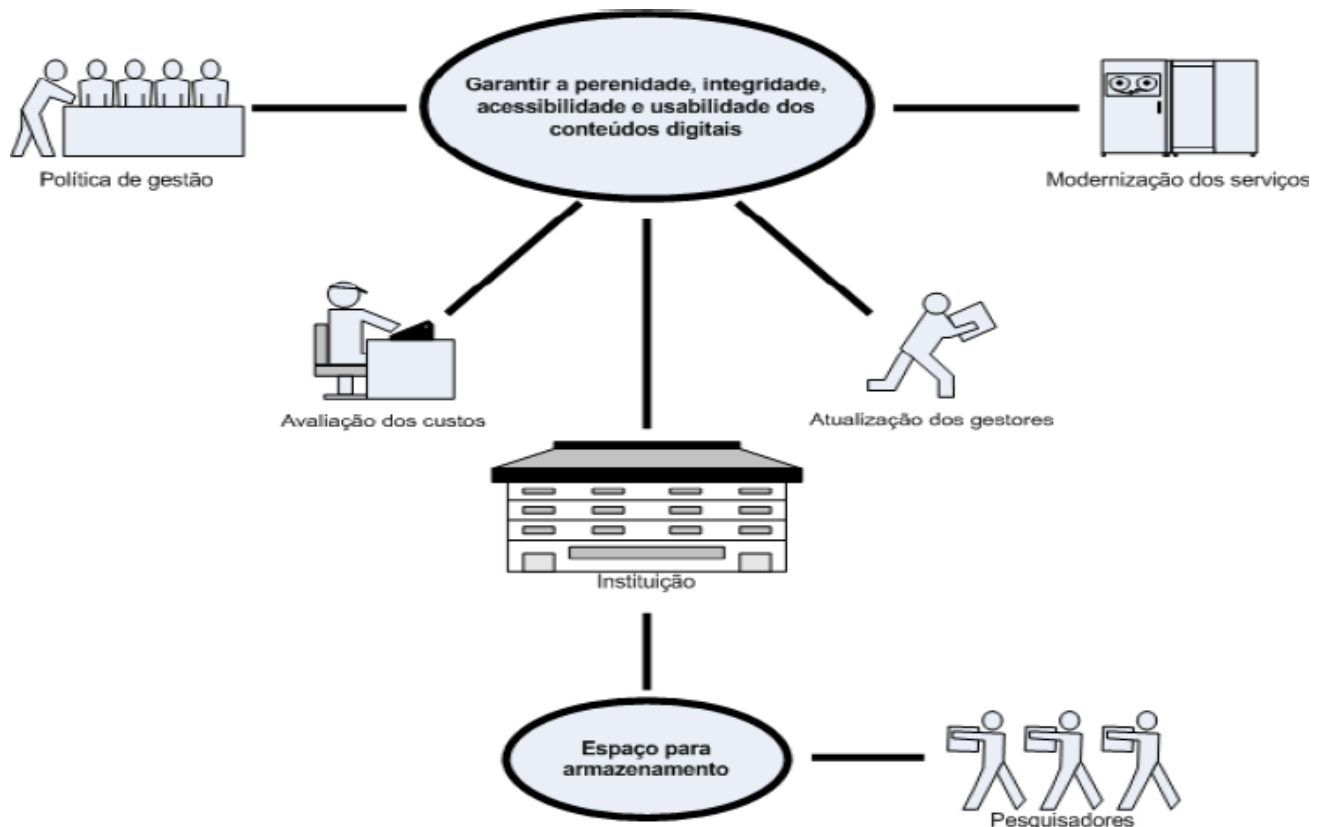


Figura 2 - Contexto das necessidades de preservação digital da informação científica.
Fonte: (ARELLANO, 2006).

Entre os especialistas, existe um consenso acerca de alguns procedimentos que facilitam o processo de preservação digital, tal como definição de um conjunto gerenciável de padrões (preferencialmente abertos), aplicação de padrões na criação dos objetos digitais, acompanhamento da obsolescência tecnológica, migração dos formatos assim que estejam consolidados, armazenamento em ambiente estável e controlável, geração de cópias de preservação, entre outras. Um grande problema dos objetos digitais, especificamente os da Internet, é a forma dinâmica em que as coisas acontecem nesse meio. Um documento citado hoje em um trabalho científico pode não mais estar disponível para consulta amanhã. Isso afeta a integridade de documentos que o referenciavam em *hyperlinks* (SANT'ANNA, 2001).

Innarelli (2007) ilustra os requisitos de preservação digital como os “10 mandamentos da preservação digital” (Figura 3).



Figura 3 – 10 mandamentos da preservação digital.
Fonte: (INNARELLI, 2007).

Há várias formas de se classificar as estratégias de preservação digital. Arellano (2004) propõe organizá-las em física e lógica. A preservação física está focalizada nos conteúdos armazenados em suportes magnéticos e ópticos. Já a lógica procura novas aplicações de hardware e software que mantenham em atividade os seus *bits*. Ferreira (2007) complementa que, no extremo oposto da preservação física/lógica, existe a preservação conceitual. O objeto digital possui o nível físico (suporte), nível lógico (modo como está inserido e como será interpretado pelo hardware/software) e nível conceitual, que é onde o objeto assume formas familiares aos seres humanos, é a imagem que se forma na mente do receptor quando vê a informação (FERREIRA, 2007). Esse é o nível do objeto digital que deve ser primordialmente preservado.

As estratégias de preservação digital neste trabalho são divididas em duas grandes abordagens, que são consenso na literatura da área: estrutural e operacional (as preservações física, lógica e conceitual estão inseridas dentro dessas abordagens). Na primeira encontram-se adoção de padrões, metadados de preservação e montagem de infra-estrutura: as mesmas características citadas necessárias à interoperabilidade são fundamentais no processo de preservação. Na operacional podem ser realizada conservação de software/hardware, migração, emulação e demais estratégias operacionais (ARELLANO, 2006).

O próximo tópico descreve as duas principais abordagens mais detalhadamente, visto que o resultado da pesquisa consiste em elaborar o diagnóstico da obsolescência tecnológica da BDTD/UFSCar, determinada através dos requisitos sociais, estruturais e operacionais para preservação digital.

2.5.1 Preservação Estrutural e Operacional

As publicações científicas de universidades e instituições de pesquisa estão adotando os arquivos abertos porque a informação que elas produzem está na sua maioria em formato digital, o que facilita a sua transferência, a vencer o problema da obsolescência da informação científica, e a criação de um espaço para a produção e armazenamento dos documentos digitais dentro da própria instituição (MORENO e ARELLANO, 2005, p.84).

Por preservação estrutural entende-se o investimento inicial por parte das instituições que vão implementar o processo de preservação. Há extrema importância em se observar este nível, pois este fornece requisitos, através de normas, padrões e estruturas, que promovem e facilitam as estratégias operacionais. Os repositórios digitais que adotam padrões abertos, por exemplo, são mais qualificados para uma preservação a longo prazo, especialmente porque permitem o acesso aos formatos dos arquivos, tornando mais fácil o desenvolvimento de

ferramentas de migração em caso de um formato se tornar obsoleto (FERREIRA, 2007).

Os recursos em ambiente digital sofrem transformações cujos resultados nem sempre são fáceis de controlar. Os metadados neste contexto são informações de apoio aos processos associados à preservação digital e formam um histórico de mudança ao longo do tempo. Eles subdividem-se em descritivos (que se destinam ao acesso), administrativos (que documentam atos de gestão) e estruturais (que descrevem o enquadramento tecnológico), sendo mais relevantes à preservação digital os dois últimos. Metadados contêm informação técnica e administrativa sobre decisões e ações de preservação, registram os efeitos das estratégias de conversão de dados e informação acerca de gestão de coleções e de direitos, asseguram a autenticidade dos recursos digitais ao longo do tempo e ainda fornecem informação acerca dos próprios metadados. A inclusão dos metadados deve acompanhar todo o ciclo de vida do objeto digital (SARAMAGO, 2004).

Entre os sistemas de metadados de âmbito mais genérico e que procuram adaptar-se à preservação digital, o esquema de metadados METS (Metadata Encoding and Transmission Standards) é uma norma para codificação de metadados descritivos, administrativos e estruturais de recursos digitais que utiliza a linguagem XML (SARAMAGO, 2004, [s.p.]).

Existe um modelo que fornece uma estrutura útil de desenvolvimento de metadados para a preservação que vai ao encontro dos requisitos necessários a atividade de preservação digital. Aprovado pela ISO em 2003, o OAIS (*Open Archival Information System*) consiste numa organização de pessoas e sistemas que aceitaram a responsabilidade de preservar a informação e torná-la disponível. Esse modelo visa identificar os componentes funcionais de um sistema de informação objetivando a preservação a longo prazo da informação digital. Ele “descreve um enquadramento conceptual para um repositório digital genérico, aberto a todas as comunidades com garantias de confiabilidade” (SARAMAGO, 2004). A tecnologia do modelo definiu uma terminologia própria que facilitou a comunicação entre os intervenientes envolvidos na preservação digital (FERREIRA, 2007).

Da mesma forma os metadados de preservação devem ser pensados em função das vocações e objetivos das comunidades, ou seja, partindo de uma base comum a vários parceiros, os conjuntos de metadados devem sofrer as necessárias adaptações (SARAMAGO, 2004, [s.p.]).

Estratégias operacionais são medidas concretas aplicadas aos objetos digitais. Sua aplicação é facilitada quando as estratégias estruturais já se encontram desenvolvidas na unidade informacional. As mais conceituadas são: ausência de solução, soluções parciais, encapsulamento, preservação da tecnologia, refrescamento, emulação, migração, conversão e Pedra de Rosetta digital, detalhadas a seguir.

Por *ausência de solução* entendem-se duas atitudes: preservar os documentos ativos e deixar que os outros desapareçam ou então deixar para o futuro, arqueologia digital, considerando que haverá tecnologias mais baratas futuramente (ARELLANO, 2005; SAYÃO, 2007).

Seguida pela ausência de solução, existem as *soluções parciais*. Estas englobam atividades como extrair e salvar somente o 'conteúdo essencial' do documento, salvar metadados que facilitem a interpretação do documento ou mesmo salvar os códigos-fontes dos softwares que processam esses documentos, para realização de engenharia reversa no futuro.

O *Encapsulamento* é uma estratégia parcial e consiste em preservar o objeto digital e toda informação necessária para permitir o futuro desenvolvimento de conversões ou emuladores (SANT'ANNA, 2001; SAYAO, 2007).

A abordagem mais utilizada para se preservar informações de documentos digitais de uma possível ausência das tecnologias envolvidas é, por mais irônico e contraditório que possa parecer, a reprodução em papel. Ocorre que os papéis e métodos de impressão utilizados também têm vida útil limitada. Além disso, a impressão de documentos multimídia é teoricamente impossível (ROTHENBERG, 1998 apud SANT'ANNA, 2001, p.130).

Na *Preservação da tecnologia*, realiza-se a preservação dos formatos lógicos/físicos originais e a utilização da tecnologia associada. Trata-se da criação

de museus de tecnologia, pois preservam o objeto na sua forma original. É uma estratégia inadequada a longo prazo, pois há dificuldade em relação à gestão do espaço físico, manutenção e custo (ARELLANO, 2005; SAYÃO, 2007; FERREIRA, 2007).

Refreshamento é a técnica na qual a informação é transferida para um suporte físico mais estável antes que o primeiro se deteriore. Esta será a estratégia de preservação sugerida como solução à obsolescência diagnosticada no objeto de estudo deste trabalho. (SANT'ANNA, 2001 et al).

A utilização de um software chamado emulador, capaz de reproduzir o comportamento de uma plataforma de hardware/software numa outra que seria incompatível, trata-se da estratégia de *Emulação*. É estratégia utilizada quando se quer preservar o ambiente tecnológico original ou quando os objetos digitais não podem ser convertidos para novos formatos. (SANT'ANNA, 2001 et al).

Migração e conversão são as técnicas mais aceitas entre os autores, e mais realizadas e consolidadas atualmente. Alguns autores as consideram juntas, outros as separam em migração de suporte e conversão de formatos. Uma definição coerente pode ser: a estratégia que consiste na transferência periódica do material digital de uma configuração de software/hardware para outra de nova geração de tecnologia; transferir para novos formatos preservando a integridade da informação. Esta estratégia se concentra em preservar o conteúdo intelectual, não conservando os originais. O que se tem é a cópia da cópia e assim sucessivamente. A migração é subdividida ainda em variantes, tais quais: migração para suportes analógicos, atualização de versões, conversão para formatos concorrentes, migração a-pedido e migração distribuída (CONWAY, 1994 et al).

Ferreira (2007) apresenta uma estratégia para quando todas as estratégias de preservação falharem: a *Pedra de Rosetta digital*. Nesta estratégia são recolhidas amostras de objetos que seriam representativas do formato o qual se quer recuperar, formando um conjunto de referências, de onde seria possível extrair as regras para se interpretar e converter o objeto em questão.

O Quadro 1 apresenta as estratégias de preservação operacionais e estruturais, segundo autores que versam em seus trabalhos a temática da preservação digital.

Abordagem e estratégias		Autores						
		1994 Conway	2001 Sant'Anna	2004 Saramago	2004/05 Arellano	2005 Moreno	2007 Sayão	2007 Ferreira
Operacional	Ausência de solução				X		X	
	Soluções parciais (Ex: Encapsulamento)		X				X	X
	Formalização				X			
	Refrescamento		X		X		X	X
	Preservação da tecnologia (conservação hardware/software) ‘Museus de tecnologia’				X		X	X
	Migração (de suporte)	X	X		X		X	X
	Conversão (de formato)	X			X			
	Emulação		X		X		X	X
Estrutural	Metadados de preservação		X	X	X		X	X
	Padrões abertos (OAI)		X	X	X	X	X	X
	Linguagem comum (ex.:XML)			X	X	X	X	X
	Protocolo OAI-PMH			X	X	X	X	X
	Padrão na criação do objeto digital			X	X		X	X
	Padrão de formato do objeto digital			X	X		X	X

Quadro 1: Matriz com as abordagens e estratégias de preservação proposta por diferentes autores.
 Fonte: Adaptado de Conway (1994), Sant'Anna (2001), Saramago (2004), Arellano (2004, 2005), Moreno e Arellano (2005), Sayão (2007), Ferreira (2007).

3. MÉTODO E DESENVOLVIMENTO

3.1 Abordagem e método da pesquisa

A presente pesquisa pode ser classificada como teórica aplicada, pois não gera somente o conhecimento pelo conhecimento, mas realiza uma aplicação prática da teoria, visando melhorias no objeto de estudo. Tem abordagem qualitativa, na medida em que se realiza a interpretação do objeto de estudo por meio da observação, bem como delimita as fronteiras do trabalho por meio de indução (TERENCE e ESCRIVÃO, 2006).

Pesquisa-ação foi o método utilizado, definido por Gomes (2007) como um processo metodológico capaz de acompanhar e ao mesmo tempo transformar a realidade que estuda. A pesquisa-ação visa resolver ou pelo menos esclarecer problemas da situação observada. O método possui um duplo objetivo: transformar a realidade e produzir conhecimentos relativos a essas transformações (GOMES, 2007). Tripp (2005) complementa dizendo que a pesquisa-ação se diferencia da prática por si só, e embora seja pesquisa, se distingue da pesquisa científica tradicional, porque ao mesmo altera o que está sendo pesquisado e é limitada pelo contexto da prática.

Ainda segundo Tripp (2005), a pesquisa-ação é uma forma de investigação que utiliza técnicas de pesquisa para informar a ação que se decide tomar para melhorar a prática. É uma forma na qual os práticos, através de pesquisa científica, buscam efetuar transformações em suas próprias práticas. Tripp (2005) propõe o diagrama, apresentado na Figura 4, pra representar o ciclo da pesquisa-ação.

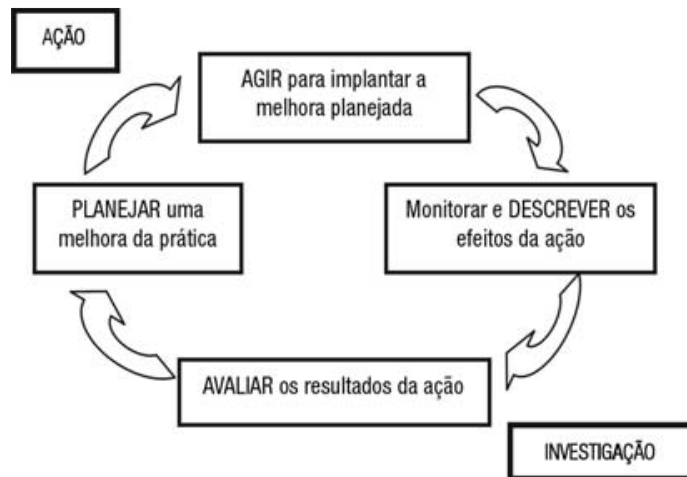


Figura 4 - Ciclo da Pesquisa-ação.
Fonte: (TRIPP, 2005).

A maioria dos processos de melhoria segue este ciclo (Figura 4). “A solução de problemas começa com a identificação do problema, o planejamento de uma solução, sua implementação, seu monitoramento e a avaliação de sua eficácia” (TRIPP, 2005). Outra questão levantada é que o relacionamento pesquisa e prática não gera apenas o aprimoramento da prática, mas também uma melhor compreensão dela por meio de sua melhora, de modo que melhora seu contexto.

Seguindo as características da pesquisa, expõe-se agora o objeto de estudo, a Biblioteca Digital de Teses e Dissertações da UFSCar – BDTD/UFSCar, escolhida porque, além da experiência de estágio da pesquisadora no local e do interesse da organização na construção de soluções para a preservação digital do seu acervo, trata-se de uma instituição pública, segundo Gomes (2007), local propício para realização da pesquisa-ação.

O IBICT criou uma solução para a integração de bases de dados de teses e dissertações, denominada nesta pesquisa de BDTD/IBICT, nos padrões de metadados e de transferência MTD-BR, utilizado para descrever eletronicamente os metadados de teses e dissertações, viabilizando o intercâmbio entre o sistema BDTD/IBICT e outros sistemas de repositórios. A BDTD/IBICT é compatível com *Dublin Core* e *NDLTD (Networked Digital Library of Theses and Dissertations)* e protocolo OAI-PMH para exportar metadados, através do TEDE (BDTD, 2009). A BDTD/IBICT foi desenvolvida com apoio da Financiadora de Estudos e Pesquisas (Finep). Possui

um Comitê Técnico-Consultivo (CTC), instalado em abril de 2002, constituído por representantes do IBICT, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), MEC (Ministério da Educação), Finep e das três universidades que participaram do grupo de trabalho e do projeto-piloto (USP, Puc-Rio (Pontifícia Universidade Católica do Rio de Janeiro) e UFSC (Universidade de Santa Catarina)). O CTC é um colegiado e atua na especificação de padrões a serem adotados no âmbito do sistema da BDTD/IBICT.

A integração à BDTD/IBICT não obriga que as instituições de ensino e pesquisa façam uso de sistemas específicos para o desenvolvimento de suas bibliotecas digitais institucionais. Somente é necessária a adoção do MTD-BR e o padrão ETD-MS da NDLTD, assim como a implementação da camada do Protocolo OAI-PMH. As instituições que já possuem BDTD usando sua própria tecnologia, o IBICT apóia na implementação do protocolo, disponível no site (www.ibict.br), para que opere sobre o repositório local e gere registros de metadados em XML/MTD-BR, promovendo interoperabilidade com os demais serviços compatíveis com o protocolo (BDTD, 2009).

Para as instituições que não possuem BDTD, o IBICT desenvolveu o TEDE, com o objetivo de proporcionar a implantação de BDTDs nas instituições. O sistema gera BDTDs institucionais e é gratuitamente distribuído em um pacote (hardware e software) contendo o Sistema TEDE já implementado, o MTD-BR e a camada do Protocolo OAI-PMH, assim como a metodologia de implantação, os manuais operacionais e de usuário, documentação e treinamento. Para utilizar esse serviço, duas ferramentas estão disponíveis para *download* com aplicações específicas: TEDE Modular, que requer a existência de infra-estrutura de integração entre o curso de pós-graduação, o autor da dissertação e a biblioteca da instituição; e o TEDE Simplificado, que possibilita a publicação eletrônica da tese ou dissertação diretamente pela biblioteca, com a autorização do autor da tese ou dissertação (BDTD, 2009). O objeto de estudo utiliza a ferramenta TEDE Simplificado.

O IBICT promoveu ainda a integração do Consórcio Brasileiro de Teses e Dissertações à NDLTD, iniciativa internacional que disponibiliza textos completos de teses e dissertações publicadas em instituições distribuídas em vários países.

Assim, a produção nacional de teses e dissertações é também disponibilizada internacionalmente (BDTD, 2009). No site do IBICT (www.bdttd.ibict.br) tem-se acesso ao campo de busca que recupera, com nível *harvesting* de interoperabilidade, as referências dos trabalhos de todas as BDTDs integrantes. Feita a busca, ele fornece um link de redirecionamento para a BDTD institucional detentora trabalho buscado, para que o usuário tenha acesso ao texto completo em PDF. Os arquivos em PDF das teses e dissertações ficam hospedados no banco de dados da Universidade. Os usuários também têm a possibilidade de fazer a busca diretamente na página da BDTD institucional. Essa iniciativa do IBICT estimula o registro e a publicação de teses e dissertações em meio eletrônico.

As instituições de ensino e pesquisa atuam como provedores de dados e o IBICT opera como agregador, coletando metadados de teses e dissertações dos provedores, fornecendo serviços de informação sobre esses metadados e expondo-os para coleta por outros provedores de serviços (BDTD, 2009, [s.p.]).

A BCo contém outras iniciativas em repositórios digitais, como por exemplo, o Portal do Conhecimento e o Fundo Florestan Fernandes, mas o presente trabalho analisa somente a BDTD/UFSCar, responsável pelo tratamento do conhecimento científico produzido pelos PPGs da UFSCar. A BDTD/UFSCar conta hoje com cerca de 2800 trabalhos (teses e dissertações) disponíveis online. No setor responsável também ocorre a confecção da Ficha Catalográfica das teses e dissertações defendidas na instituição, seguida da geração da versão final desses trabalhos em formato PDF. É no momento da confecção que o autor é convidado a disponibilizar seu trabalho online na BDTD/UFSCar. Para a disponibilização é necessário um documento de autorização disponibilizado pela BCo, no qual o autor autoriza disponibilizar seu trabalho sem ônus e em texto completo (formato PDF). O bibliotecário responsável mantém uma cópia do trabalho no computador do setor e arquivava o documento de autorização. Em seguida o estagiário, em posse do documento, salva outra cópia do trabalho em questão, esta protegida contra cópia e alteração, e disponibiliza o trabalho online na BDTD/UFSCar. Esse processo é formado por três etapas: digitação dos metadados, disponibilização da cópia/arquivo

PDF protegido e uma confirmação final de disponibilização. Na rotina de disponibilização também havia um procedimento de backup em mídia de CD-ROM, onde a bibliotecária encarregada pelo setor realizava cópias individuais dos trabalhos em CDs-ROM fornecidos pelos autores. Devido à má qualidade das mídias e a falta de espaço para o seu armazenamento, optou-se por eliminar este procedimento.

3.2 Desenvolvimento da pesquisa

Com base na literatura apresentada na seção 2, “Acessibilidade e preservação digital”, selecionou-se um conjunto de requisitos necessários ao processo de diagnóstico do *status* da preservação digital na organização estudada. Os requisitos e resultado da sua aplicação podem ser visualizados no Quadro 3, encontrado na Seção 4 – Resultados.

Após o diagnóstico foi desenvolvida uma solução para a preservação digital dos trabalhos científicos produzidos pelos PPG da UFSCar. A solução proposta permite o armazenamento, recuperação e acesso aos trabalhos em texto completo (PDF) disponibilizados em uma mídia de DVD-ROM. Esta estratégia é denominada, segundo a literatura, *Refreshamento*, técnica na qual a informação é transferida para um suporte físico mais estável antes que o primeiro se deteriore. No caso desta pesquisa, os CDs-ROM que já estão sofrendo deterioração e contém apenas um documento por mídia, serão substituídos pela solução desenvolvida em DVD-ROM, que compreende uma base de dados, um conjunto de documentos e um sistema de busca para recuperação dos mesmos. Proporciona-se então refreshamento de mídia e otimização do uso do espaço físico. No Quadro 2 é apresentada uma síntese do processo de desenvolvimento da solução, que contou com a aplicação de alguns softwares da família ISIS, de acesso livre, disponibilizados no Brasil pela BIREME (2010). É importante ressaltar que todo o desenvolvimento da pesquisa contou com uma forte interação entre a pesquisadora

e os membros da organização, objeto de estudo desta pesquisa, e esta interação é parte essencial do método pesquisa-ação.

Passos	Descrição das atividades
1) Coleta BDTD/UFSCar	Coleta dos trabalhos em PDF disponíveis online que não possuíam cópia no computador do setor. Realizada através de uma listagem com os dados referenciais dos trabalhos.
2) Recuperação de dados da BDTD	Recuperação dos dados referenciais sobre Teses e dissertação armazenadas na BDTD.
3) Tratamento dos dados (Figura 5)	Formatação dos dados em um arquivo texto necessários à geração de uma base de dados ISIS.
4) Geração da base de dados ISIS	Geração da base de dados com o aplicativo MX da família ISIS, disponibilizado pela BIREME (2010) - (Figura 6).
	Construção de uma interface de gestão com o aplicativo WWINISIS, disponibilizado pela BIREME (2010) - (Figura 7).
	Indexação da base de dados: por autor, PPG, palavras-chave, entre outros campos descritivos - (Figura 7).
5) Geração de uma interface de acesso (Figura 8)	Aplicação do programa GENISIS para desenvolver as interfaces necessárias à recuperação e manipulação dos dados e dos textos completos das Teses e dissertações depositadas na BDTD/UFSCar.
6) Personalização da interface gráfica (Figura 9)	Implementação de melhorias nas interfaces de através do software Macromedia Dreamweaver: inserção logo UFSCar e da BDTD/UFSCar, configuração das cores entre outras melhorias.
7) Gravação DVD-ROM	Gravação de uma série de DVD abordando o conhecimento científico produzido por PPG, visando a disponibilização dos dados bibliográficos e do seu texto completo em formato PDF.
8) Avaliação	Avaliação da solução pelos trabalhadores da organização - em andamento.

Quadro 2 – Processo de desenvolvimento da solução computacional.

A solução compreende uma amostra de trabalhos do PPG Linguística, recuperados em texto referencial e completo na BDTD/UFSCar. O tratamento dos dados (Passo 3) foi realizado estabelecendo os requisitos para a execução do aplicativo MX. Os dados referenciais sobre os trabalhos científicos ficaram dispostos no seguinte formato (Figura 5):

Autor|Título|Ano|Palavras-chave|Cobertura|Resumo.

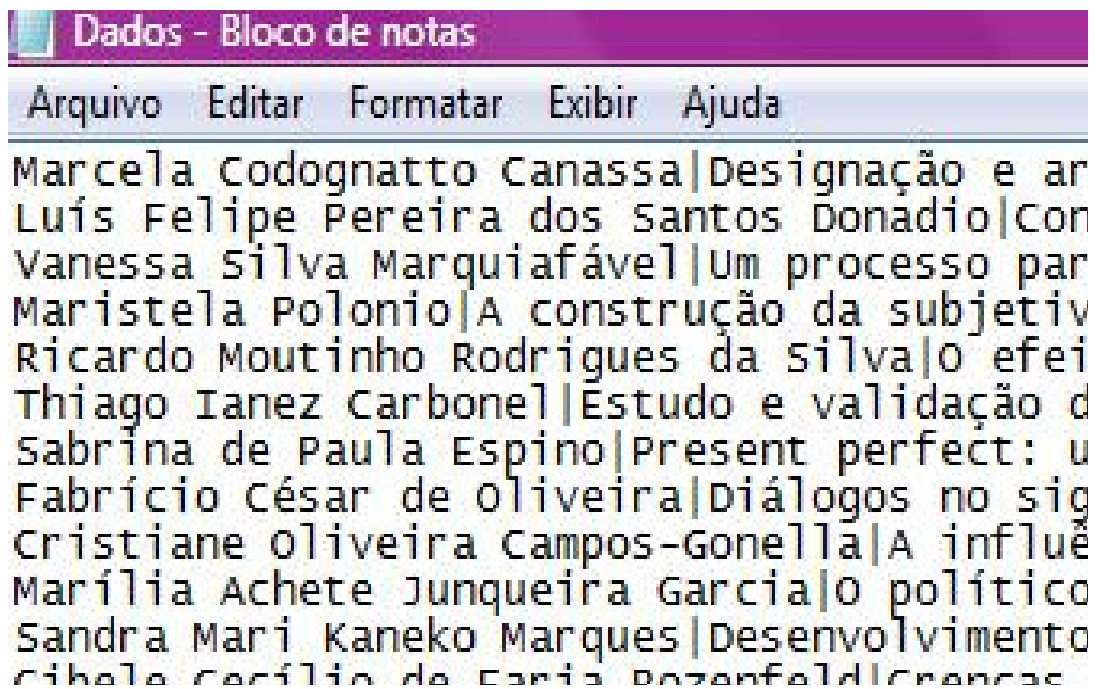


Figura 5 – Tratamento dos dados estabelecendo requisitos para aplicação do MX.

No Passo 4 o aplicativo MX foi executado para a geração de uma base de dados ISIS a partir do arquivo em formato .txt com os dados referenciais dos trabalhos científicos, tratados no Passo 3 (Figura 6).

```

D:\Linguistica>mx BaseLin iso=BaseLing.iso now
mf n= 1
1 «Marcela Codognatto Canassa»
2 «Designação e argumentação em uma ação popular da cidade de São Carlos»
3 «2007»
4 «Semântica; Designação; Lógica; Texto jurídico»
5 «1350»
mf n= 2
1 «Luís Felipe Pereira dos Santos Donadio»
2 «Content-based Instruction para professores pré-serviço: uma análise da pr
blemática do processo de avaliação»
3 «2007»
4 «Professores - formação; Língua inglesa - ensino; Content-based instructio
Avaliação de aprendizagem»
5 «1363»
mf n= 3
1 «Vanessa Silva Marquias»
2 «Um processo para a geração de recursos lingüísticos aplicáveis em ferrame
tas de auxílio à escrita científica»
3 «2007»
4 «Lingüística - processamento de dados; Lingüística de corpus; Língua ingle
- ensino; Ferramenta de apoio à escrita científica; Gênero textual»
5 «1393»
mf n= 4
1 «Maristela Polonio»

```

Figura 6 - Geração da base de dados com o aplicativo MX da família ISIS.

Após a construção de uma interface de gestão com o aplicativo WWINISIS, foi realizada a indexação dos dados, ilustrada na Figura 7:

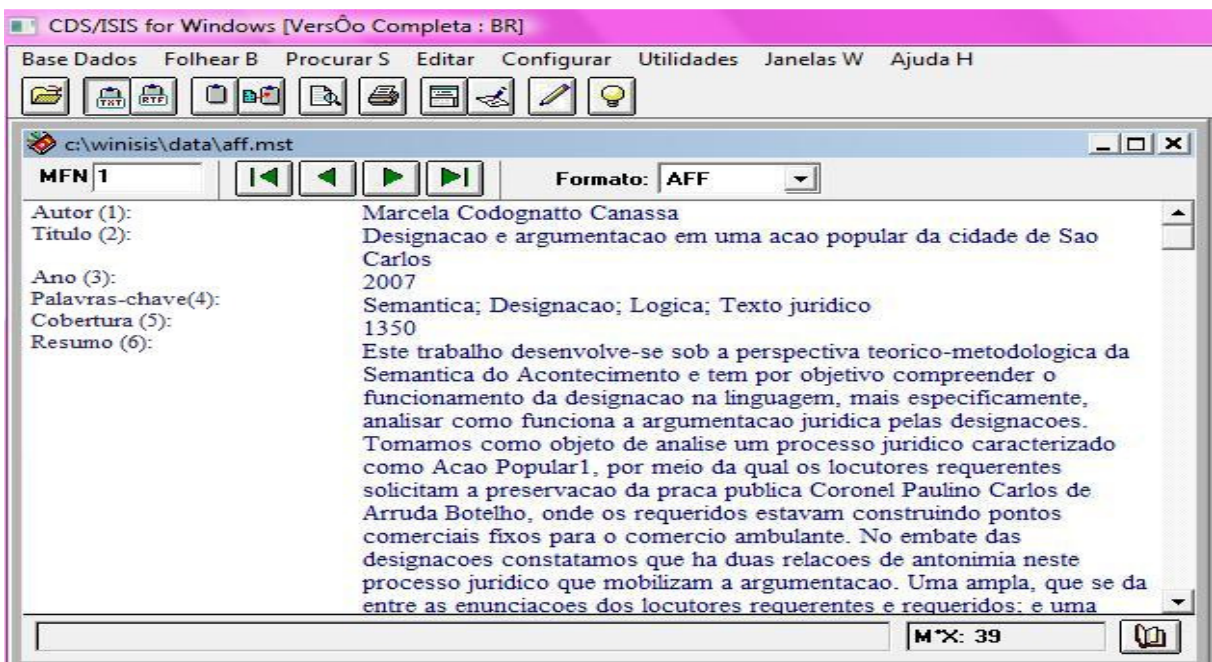


Figura 7 - Indexação da base de dados no WINISIS.

Após a indexação, foram geradas as interfaces gráficas da solução com o software Genisis da Família ISIS (Figura 8). Foram construídas as interfaces

de busca, lista de resultados e resultado detalhado. Optou-se por realizar as buscas por: Todos os campos, Autor, Título, Palavras-chave, Ano, Cobertura e Resumo. A Figura 8 ilustra o processo de confecção das interfaces.

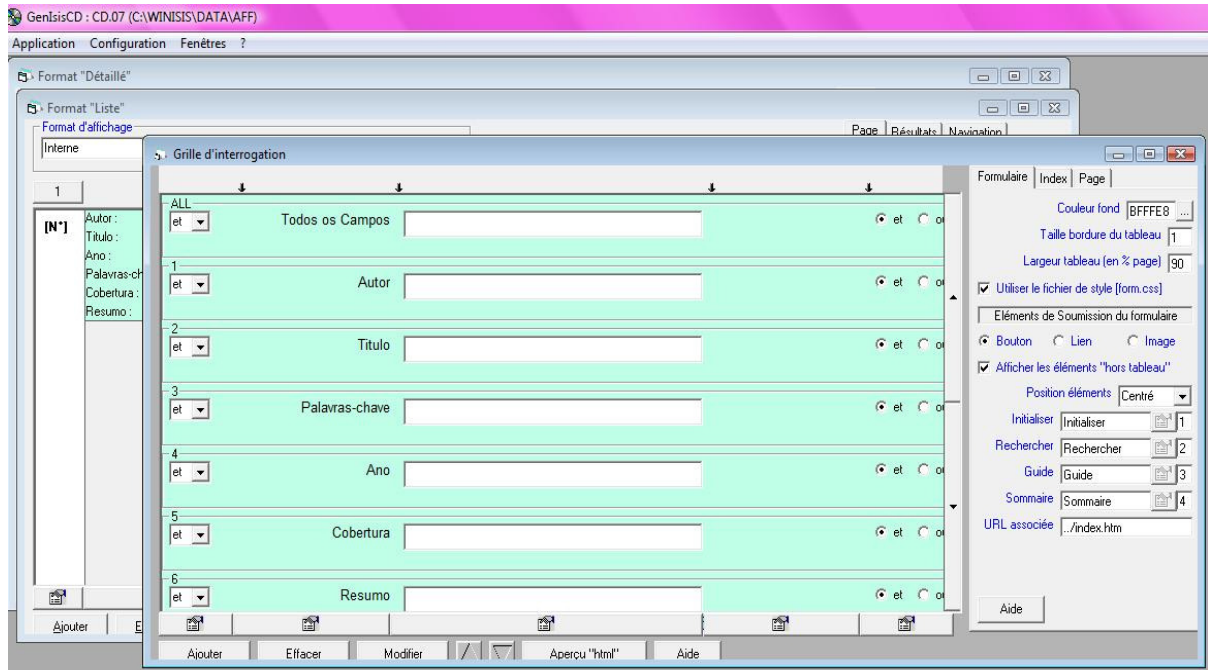


Figura 8 - Aplicação do software GENESIS para desenvolver as interfaces.

No Passo 6, foram realizadas melhorias nas interfaces gráficas, como a inserção do logo UFSCar e Bco, do link de acesso para o texto completo do trabalho em PDF e demais ajustes. As melhorias foram realizadas com o editor de páginas HTML Macromedia Dreamweaver, como ilustra a Figura 9.

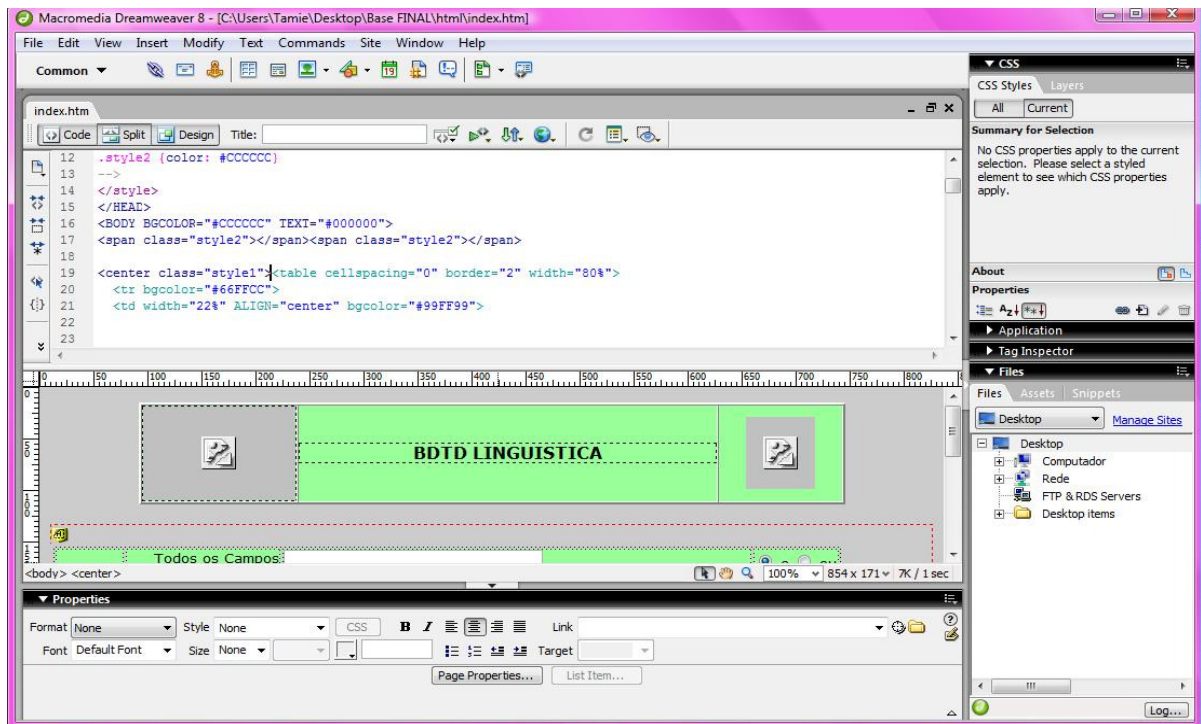


Figura 9 - Implementação de melhorias nas interfaces através do Macromedia Dreamweaver.

Passo 7 gravação do DVD – Após a personalização da interfaces gráficas (Passo 6), foram gravados em um DVD: as interfaces gráficas; a base de dados referenciais dos trabalhos científicos do PPG-L e os documentos digitais dos trabalhos em formato PDF.

Até o presente momento a solução foi apresentada aos funcionários da organização estudada e a sua avaliação (Passo 8) está em andamento pelos mesmos.

4. RESULTADOS

4.1 Formatação e Aplicação de uma Estrutura de Referência para Apoiar o Processo de Diagnóstico do *Status* da Obsolescência Tecnológica

Combinando a análise da literatura sobre preservação digital com a experiência dos funcionários da organização e da pesquisadora, adquirida na atuação com a BDTD/UFSCar como estagiária em um período de dois anos (2008-2010), pode-se diagnosticar a obsolescência tecnológica da BDTD/UFSCar, seu resultado é apresentado no Quadro 3:

Abordagem	Diagnóstico da conformidade da BDTD/UFSCar aos requisitos necessários à preservação digital de uma biblioteca digital aberta e interoperável		Descrição	
		SIM NÃO		
Social	Estabilidade social da BDTD/UFSCar		X	Devido ao alto índice de “ <i>turn over</i> ” de estagiários, responsáveis pela inserção de dados e pelo fato de que a maioria dos componentes da equipe de trabalho não possui dedicação exclusiva à atuação com a BDTD/UFSCar.
	Gestores com preparação		X	Ausência de treinamentos formais na gestão de repositórios digitais.
	Profissionais preparados	X*		Parte da equipe recebeu treinamentos básicos sobre a operacionalização do sistema TEDE, ainda se faz necessários treinamentos que compreendam: linguagens de programação; editoração eletrônica entre outros.
	Equipe multidisciplinar	X*		01 Analista de sistema: dedicação parcial; 01 Bibliotecária: dedicação parcial; 01 Auxiliar administrativo: dedicação exclusiva; 02 Estagiários do Curso de Biblioteconomia e Ciência da Informação da UFSCar: dedicação exclusiva. Dada a demanda pelo serviço há necessidade de ampliação da equipe de trabalho.
	Planejamento de preservação digital		X	Incipiente: esta pesquisa é a primeira iniciativa da organização no sentido de preservação do conteúdo digital das teses e dissertações.
	Definição de objetivos (O que deve ser preservado?)		X	

	Estudo de iniciativas e experiências nacionais e internacionais		X	
	Aspectos legais	X		Há entendimento dos aspectos legais.
	Direitos autorais	X		Há entendimento dos direitos autorais.
	Possui atividades de acompanhamento da obsolescência tecnológica		X	Esta pesquisa é a 1ª iniciativa neste sentido.
	Financiamento sustentável	X		Atualmente os recursos são providos pela UFSCar. Existem iniciativas no sentido de encontrar novas fontes de fomento, como por exemplo, participação em editais.
	Responsabilidade pela acessibilidade		X	Atualmente não é possível acessar o conteúdo completo da BDTD/UFSCar pelo sistema BDTD/IBICT.
	Mecanismos de certificação	X		Por utilizar o sistema TEDE a BDTD/UFSCar é certificada pelo IBICT como biblioteca digital aberta e interoperável.
Estrutural	Softwares apropriados		X	Existe a demanda por aplicativos na editoração eletrônica dos documentos digitais.
	Softwares livres	X*		Sistema operacional Linux, Sistema de Gestão do repositório TEDE, Sistema servidor web Apache e Sistema de banco de dados: Livre. Aplicativos para a editoração de documentos: código proprietário.
	Hardware apropriado	X*		Equipamentos satisfatórios quanto ao número e desempenho. Dada a crescente demanda de serviço há necessidade de novos equipamentos.
	Modelo OAI	X		Conforme especificações BDTD/IBICT.

	Adoção de padrões de metadados (MTD-BR)	X		Conforme especificações BDTD/IBICT.
	Registros de metadados em XML/MTD-BR	X		Conforme especificações BDTD/IBICT.
	Adoção do protocolo OIA-PMH	X		Conforme especificações BDTD/IBICT.
	Adoção do padrão ETD-MS da NDLTD	X		Conforme especificações BDTD/IBICT.
	Padrões na criação dos objetos digitais	X		A geração dos documentos obedece à norma ABNT 14701. Existem anomalias quanto à aplicação da norma, porém nada que comprometa a qualidade dos documentos.
	Padrões no formato dos objetos digitais	X		Formato PDF.
	Garantia de integridade do objeto digital.	X*		Não há possibilidade de edição dos documentos disponibilizados na BDTD/UFSCar. Porém há o problema da falta de controle no acesso a pessoas externas e internas no Departamento responsável pela editoração e armazenamentos dos documentos.
	Garantia de acessibilidade plena		X	No momento não é possível acessar o conteúdo completo dos documentos pela BDTD/IBICT.
	Disponibilidade de conexão, velocidade e segurança	X*		As condições são satisfatórias. Há ocorrências de falhas de acesso temporárias e intermitentes.
Operacional	Cópias de segurança (backup) – digital	X*		Somente no servidor da UFSCar.
	Cópias de segurança (backup) – analógica	X*		Os documentos estão disponíveis em dois exemplares em papel. Isto ocasiona um alto consumo de espaço físico da BCo. Outro agravante é a falta de

				mecanismos de controle da entrega das teses e dissertações impressas.
	Espaço físico	X		A BDTD/UFSCar esta localizada no Departamento de Processamento técnico da BCo/UFSCar. Não há mecanismos de controle de acesso físico de pessoas internas e externas ao departamento. Situação crítica, porque proporciona falta de segurança aos documentos sobre direitos autorais e de integridade dos documentos disponibilizados na BDTD/UFSCar, pois é possível ter acesso aos equipamentos e aplicativos utilizados na editoração eletrônica dos mesmos.

*Sim, porém com restrições.

Quadro 3 - Da conformidade da BDTD/UFSCar aos requisitos necessários à preservação digital.

Fonte: Adaptado de Conway (1994), Sant'Anna (2001), Arellano (2004, 2005), Sayão (2007), Ferreira (2007), BDTD (2009), Saramago (2004), Tammaro e Salarelli (2008), IBICT (2009), Moreno e Arellano (2005), Gracio e Fadel (2008).

4.2 Solução computacional para a preservação digital

A solução proposta para a preservação digital dos trabalhos científicos dos PPG UFSCar (Quadro 2) , identificada na literatura como *Refrescamento*, foi desenvolvida com software livre, e se implementada na organização estudada proporcionará maior acessibilidade ao conteúdo digital, necessária devido aos períodos de indisponibilidade identificados no diagnóstico. Acrescentará ainda uma rotina de *backup* mais segura e diminuirá a ocupação de espaço físico ocupado pelas teses e dissertações em formato impresso, que poderão contar apenas com um exemplar em papel na BCo, já que a solução servirá também como fonte de consulta e acesso ao texto completo em formato PDF.

Como projeto piloto aplicou-se os passos do Quadro 2 à 39 documentos e suas devidas referências do PPG em Lingüística, escolhido aleatoriamente. A base foi então denominada '*BDTD Lingüística*'. A busca pode ser realizada por *Autor, Título, Ano, Palavras-chave, Cobertura e Todos os Campos*. Outra sugestão de classificação dos DVDs-ROM é o armazenamento anual, contendo os documentos de todos os PPG do ano vigente. Nesse caso o campo *Ano* seria substituído pelo campo *PPG*. As interfaces, resultado do processo de desenvolvimento já descrito e ilustrado, podem ser observadas nas Figuras 10, 11, 12, 13 e 14 a seguir. Todas as funcionalidades da solução desenvolvida podem ser analisadas através do DVD-ROM que acompanha este trabalho. A Figura 10 ilustra o processo de instalação da solução em DVD-ROM no computador do usuário. É necessário instalar somente no primeiro acesso. Nos acessos seguintes basta inserir o DVD-ROM e acessar a opção "*Acessar a BDTD Lingüística*".

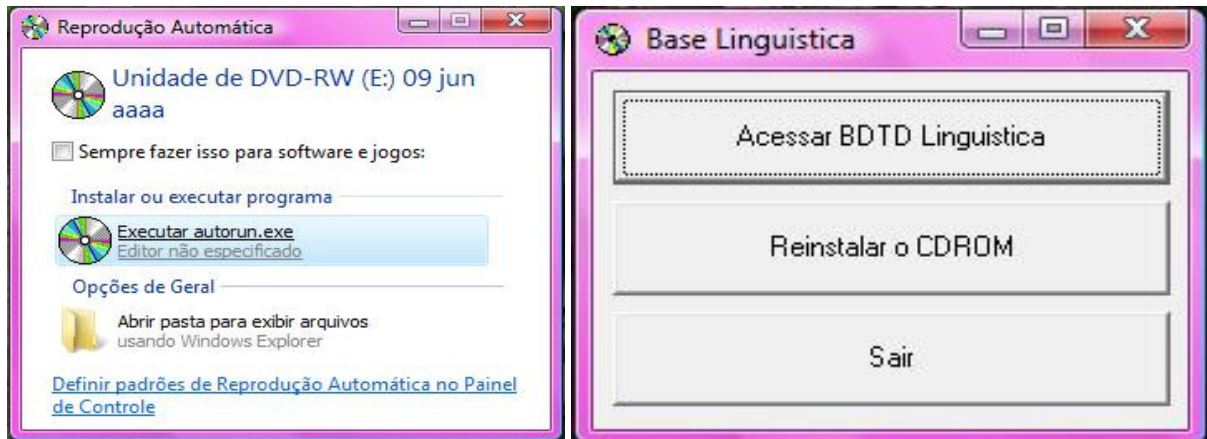


Figura 10 – Processo de instalação do DVD-ROM da ‘BDTD Lingüística’.

A Figura 11 ilustra a interface de busca da solução, com os campos de busca, as opções de truncagem (entre as palavras digitadas no campo e entre os campos) e a opção de escolher o número de registros apresentados por página nos resultados.

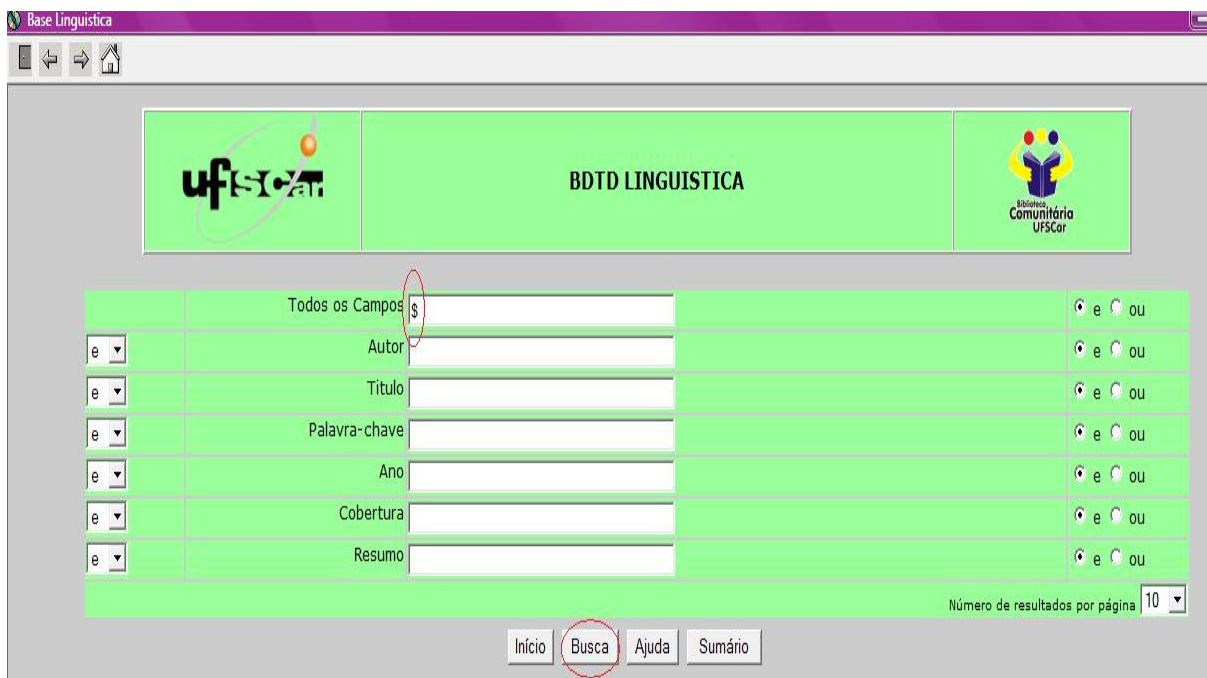


Figura 11 – Interface de busca da ‘BDTD Lingüística’.

Após realizada a busca, executada na ilustração com o termo '\$', que recupera todos os registros indexados, tem-se a interface dos resultados. Optou-se, na lista de resultados, apresentar somente o título do documento, mantendo assim uma interface simples e clara, conforme é visualizado na Figura 12.

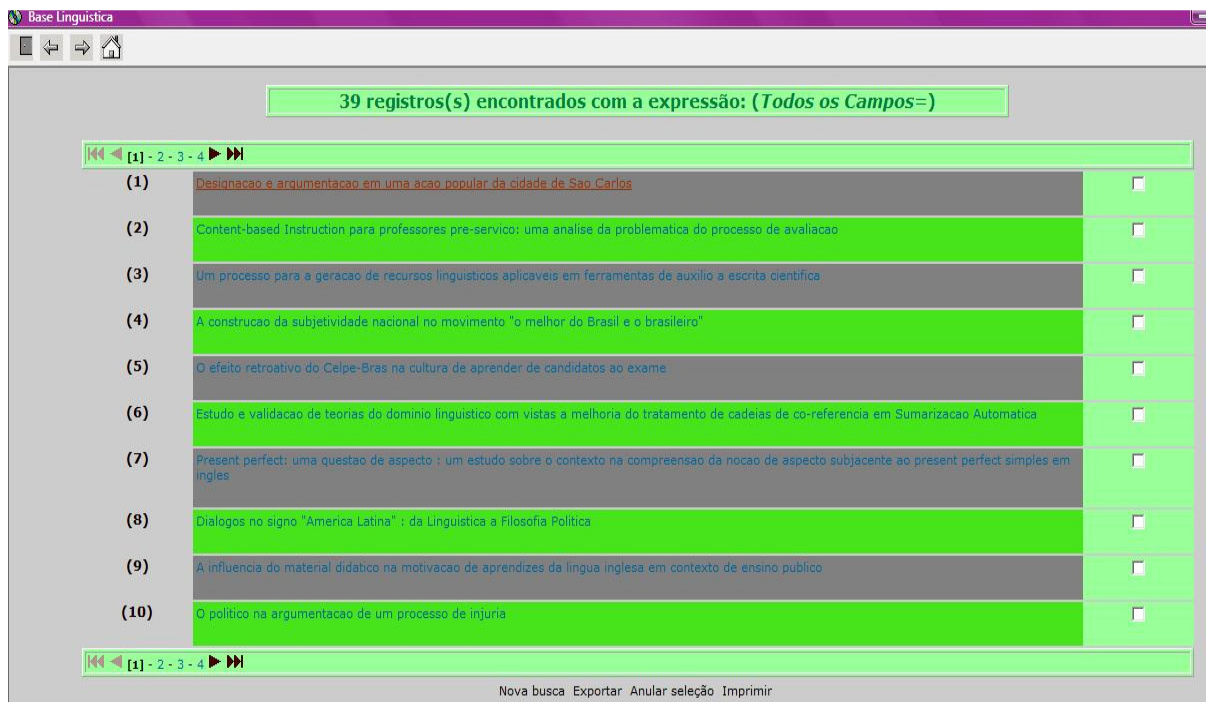


Figura 12 – Interface da lista de resultados na 'BDTD Lingüística'.

Com um clique no título desejado, tem-se acesso, através de um hiperlink, à interface dos resultados detalhados, com a exibição de todos os campos indexados, ilustrado a na Figura 13.

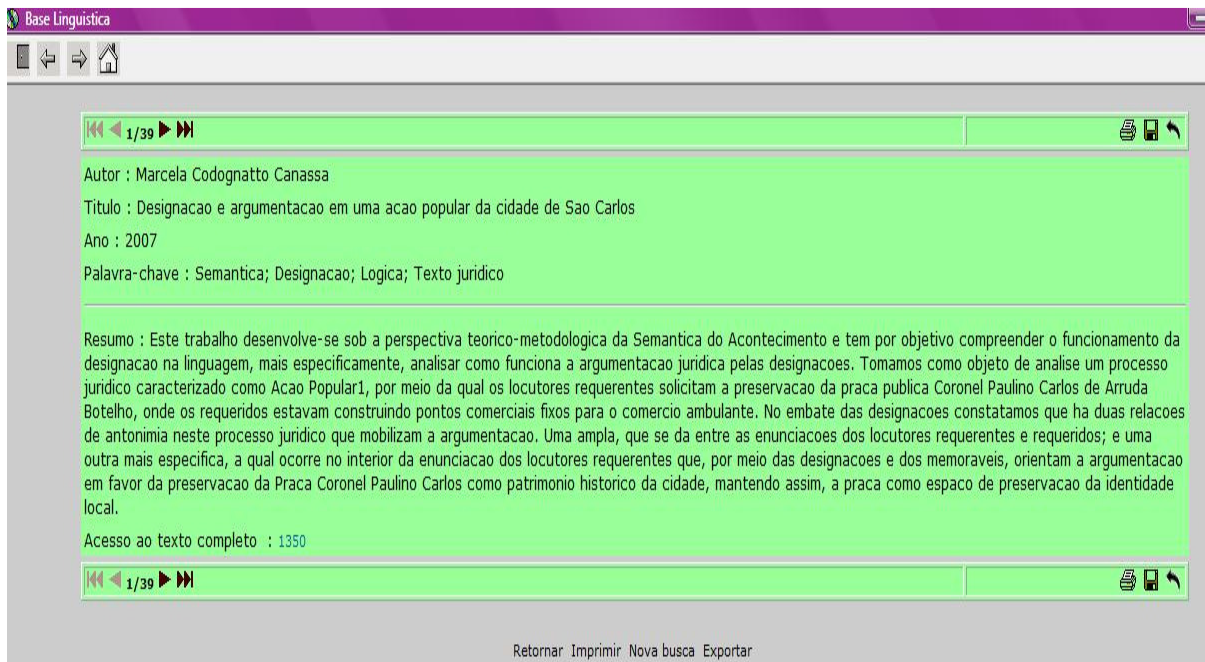


Figura 13 – Interface do resultado detalhado na ‘BDTD Lingüística’.

Clicando no link disponível em “Acesso ao texto completo” (representado pelo número de cobertura - campo indexado), tem acesso ao PDF do documento, como ilustra a Figura 14.

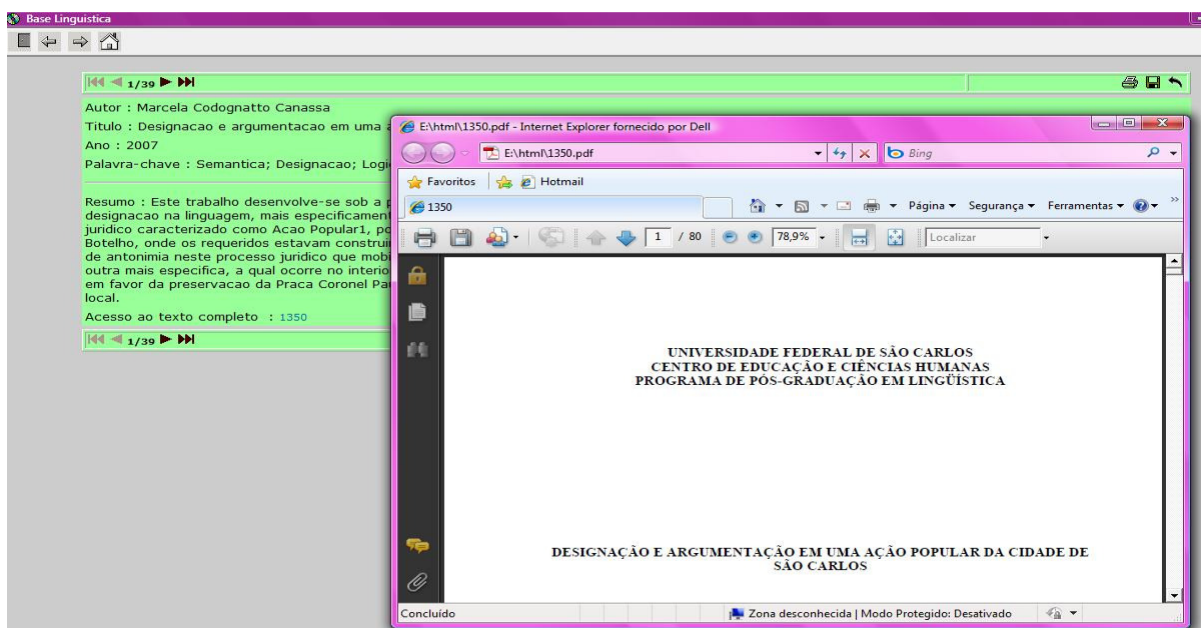


Figura 14 – Acesso ao texto completo em PDF na ‘BDTD Lingüística’.

5. CONSIDERAÇÕES FINAIS

A preservação digital é um dos grandes desafios do século XXI e as iniciativas se multiplicam na busca de soluções, principalmente para as informações relacionadas à ciência e tecnologia. A estrutura de referência desenvolvida nesta pesquisa contribui para as iniciativas de preservação digital nas bibliotecas universitárias e demais centros de informação, ao disponibilizar modelo de referência para diagnóstico da obsolescência tecnológica envolvendo requisitos sociais, estruturais e operacionais.

Os profissionais da informação precisam se envolver na questão da preservação digital por serem um dos principais atores nesta área. A falta de conhecimento quando o assunto é preservação digital, além da inexistência de pessoal com preparação nas instituições é observada por diversas pesquisas da área: o tema é um assunto novo na maioria das instituições que tratam a informação científica e tecnológica. A informação digital é uma realidade e o volume de informações disponibilizadas nesse formato cresce a cada dia, tornando inevitável o desenvolvimento de políticas e planejamentos de preservação digital.

Conclui-se que o objetivo do trabalho foi alcançado, ao desenvolver, além do modelo de referência, uma proposta de solução computacional envolvendo a estratégia de preservação denominada *Refreshamento* para apoiar a preservação digital da BDTD/UFSCar. Os resultados auxiliam para que a contribuição da comunidade científica da UFSCar, divulgada pelos documentos digitais gerados nos PPG da instituição, possa ser ininterruptamente acessada, em termos de conteúdo e funcionalidades através dos recursos tecnológicos disponíveis na época, tendo garantida sua integridade, autenticidade e confiabilidade. O modelo de referência para o diagnóstico da obsolescência tecnológica permite o acompanhamento da evolução da obsolescência tecnológica do repositório digital. A aplicação da solução computacional promove a liberação do espaço físico da biblioteca, o acesso nos períodos de instabilidade da rede e do servidor e uma rotina de *backup* mais segura aos documentos digitais. Ambos são de fácil aplicação e não há necessidade de

trabalhos terceirizados. Com a solução computacional fica exemplificado que o desenvolvimento de estratégias de preservação digital não significa dispendiosos recursos e trabalho: com ferramentas simples, softwares livres, baixo custo de material e fácil desenvolvimento pode-se promover estratégias de preservação digital de maneira eficiente.

Novas pesquisas podem dar continuação ao tema abordado, como o desenvolvimento de políticas de preservação nas instituições de ensino e pesquisa, apoiadas nas iniciativas nacionais e internacionais já realizadas, como a Biblioteca Nacional da Austrália com a criação do PADI – “*Preserving Access To Digital Information*”, que possui um site com artigos, livros, eventos, tutoriais e guias relacionados às políticas de preservação digital adotadas em várias instituições no mundo. Iniciativas como a “Carta para a Preservação do Patrimônio Digital” publicada pela UNESCO (*United Nations Educational, Scientific and Cultural Organization*), em 2003, que aponta os princípios para o desenvolvimento de políticas, ou a “Carta para a Preservação do Patrimônio Arquivístico Digital: Preservar para garantir o acesso”, aprovada pelo Conselho Nacional de Arquivos (CONARQ), podem nortear as formulações e implementações das políticas de preservação digital. Outra proposta é analisar as equipes nas unidades de informação a fim de orientar esforços para planejar, captar, desenvolver e avaliar as competências necessárias à implementação das políticas bem como desenvolvimento e aplicação dos planejamentos de preservação digital.

REFERÊNCIAS

ALENCAR, E. M. L. S. O estímulo à criatividade em programas de pós-graduação segundo seus estudantes. *Psicol. Reflex. Crit.* [online], vol.15, n.1, 2002. Disponível em: <<http://www.scielo.br/pdf/prc/v15n1/a08v15n1.pdf>>. Acesso em: 10 nov. 2009.

AMARAL, R. M. *Impacto da obsolescência de hardware e software no gerenciamento de acervo digital*. Texto não publicado cedido pelo autor, 2009.

ARELLANO, M. A. Preservação de documentos digitais. *Ci. Inf.* [online], v.33, n.2, 2004. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewFile/305/271>>. Acesso em: 20 out. 2009.

ARELLANO, M. A. Preservação digital. In: 1º Seminário de Gestão da Informação Jurídica em Espaços Digitais, 2007, Brasília. *Anais...* Brasília: Supremo Tribunal Federal, 2007. Disponível em: <<http://www.stf.jus.br/arquivo/sijed/10.pdf>>. Acesso em: 10 out 2009.

ARELLANO, M. A.; ANDRADE, R. S. Preservação digital e os profissionais da informação. *Ci. Inf.* [online], v.7, n.5, out. 2006. Disponível em: <http://www.dgz.org.br/out06/Art_05.htm>. Acesso em: 18 nov. 2009.

BAX, M. P. Agentes de Interface para Bibliotecas Digitais: a Arquitetura SABiO. In: VI Seminário sobre Automação em Bibliotecas e Centros de Documentação, n.6, 1997, Águas de Lindóia. *Anais...* Águas de Lindóia: [s.n.], 1997.

BDTD – Biblioteca Digital Brasileira de Teses e Dissertações. Disponível em: <<http://bdttd.ibict.br/>>. Acesso em: 30 dez. 2009.

BIREME. *MX(c)Bireme*: algumas aplicações úteis. Tradução e adaptação de Elysio Mira Soares de Oliveira. Disponível em: < <http://www.elysio.com.br/site/mx.htm>>. Acesso em: 20 fev 2010.

BOERES, S. A.; ARELLANO, M. A. M. Políticas e estratégias de preservação de documentos digitais. In: VI CIFORM – Encontro de Ciência da Informação, 2005, Salvador. Anais... Salvador: UFBA, 2005.

BOTOMÉ, S. P.; KUBO, O. M. Responsabilidade social dos programas de Pós-graduação e formação de novos cientistas e professores de nível superior. *Inter. em Psico.*, v.6, n.1, 2002.

CDS/ISIS database software. In: UNESCO. Disponível em: <http://portal.unesco.org/ci/en/ev.php-URL_ID=2071&URL_DO=DO_TOPIC&URL_SECTION=201.html>. Acesso em: 12 fev 2010.

COELHO, P. *Onze minutos*. Rio de Janeiro: Rocco, 2003.

CONWAY, P. *Preservação no universo digital*. Rio de Janeiro: Arquivo Nacional, 1997.

CUNHA, M. B. Biblioteca digital: aspectos econômicos. Brasília: [s.d.]. Apresentação em PowerPoint. Disponível em: <<http://www.ime.usp.br/~cesar/simposio99/murilo/sld001.htm>>. Acesso em: 20 out. 2009.

CUNHA, M. B. Biblioteca digital: bibliografia internacional anotada. *Ci. Inf.* [online], v.26, n.2, 1997. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewFile/403/362>>. Acesso em: 18 out. 2009.

CUNHA, M. B. Construindo o futuro: a biblioteca universitária brasileira em 2010. *Ci. Inf.* [online], v.29, n.1, 2000. Disponível em: <<http://www.scielo.br/pdf/ci/v26n2/v26n2-12.pdf>>. Acesso em: 06 nov. 2009

CUNHA, M. B. Desafios na construção da biblioteca digital. *Ci. Inf.* [online], v.28, n.3, 1999. Disponível em: <<http://www.scielo.br/pdf/ci/v28n3/v28n3a3.pdf>>. Acesso em: 17 out. 2009.

DRABENSTOTT, K. M., BURMAN, C. M., MACEDO, N. D. Revisão analítica da biblioteca do futuro. *Ci. Inf.* [online], v.26, n.2, 1997. Disponível em: <<http://www.scielo.br/pdf/ci/v26n2/v26n2-11.pdf>>. Acesso em: 10 out. 2009.

DSpace - Repositórios Digitais. Disponível em: <<http://dspace.ibict.br/>>. Acesso em 12 nov. 2009.

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

FUJINO, A.; STAL, E.; PLONSKI, A. A proteção do conhecimento na universidade. *Rev. de Adm.*, v.34, n.4, 1999. Disponível em: <http://www.propesp.ufpa.br/spi/arquivos/prot_conhec_universid.pdf>. Acesso em: 08 out. 2009.

FUJITA, M. S. L. A biblioteca digital no contexto da gestão de bibliotecas universitárias: análise de aspectos conceituais e evolutivos para a organização da informação. In: encontro nacional de ciência da informação, v.6, 2005, Salvador. *Anais...* Salvador: UFBA, 2005. Disponível em: <<http://dici.ibict.br/archive/00000506/>>. Acesso em: 05 out. 2009.

GOMES, A. A. *Pesquisa-ação: conceitos e aplicação - Uma alternativa metodológica para a pesquisa em educação*. Texto didático da disciplina Tópicos Especiais de Pesquisa em Educação. Presidente Prudente: UNESP, 2007.

GRACIO, J. C. A.; FADEL, B. A Política de preservação digital nas instituições de ensino. In: IX Encontro de Pesquisadores - Uni-FACEF, 2008. *Anais...* Franca: Centro Universitário de Franca, 2008. Disponível em: <<http://www.facef.br/novo/publicacoes/IIforum/Textos%20EP/Jose%20Carlos%20e%20Barbara%20Fadel.pdf>>. Acesso em 03 maio 2010.

IMMIANOVSKY, E. *Ferramenta de indexação e recuperação de documentos eletrônicos utilizando Oracle intermídia*. 2007. 55p. Monografia (Graduação em Ciência da Computação) - Universidade Regional de Blumenau, Blumenau. Disponível em: <<http://www.inf.furb.br/tcc/index.php?cd=10&tcc=1008>>. Acesso em: 15 junho 2009.

IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia. Disponível em: <<http://www.ibict.br/>>. Acesso em: 16 março 2010.

INNARELI, H. C. Preservação de acervos digitais. In: 1º Seminário de Gestão da Informação Jurídica em Espaços Digitais, 200, Brasília. *Seminário...* Brasília: Supremo Tribunal Federal, 2007. Disponível em: <<http://www.stf.jus.br/arquivo/sijed/10.pdf>>. Acesso em: 9 junho 2010.

KATO, R. *Biblioteca da USP importa robô único no Brasil para digitalizar livros*. Disponível em: <<http://www.abril.com.br/noticias/diversao/biblioteca-usp-importa-robo-unico-brasil-digitalizar-livros-415743.shtml>>. Acesso em 10 nov. 2009.

LASTRES, H. M. M. Informação e conhecimento na nova ordem mundial. *Ci. Inf.* [online], v.28, n.1, 1999. Disponível em: <<http://www.scielo.br/pdf/ci/v28n1/28n1a09.pdf>>. Acesso em: 11 out. 2009.

LOPES, O. U. Pesquisa básica *versus* pesquisa aplicada. *Estud. Av.* [online], vol.5, n.13, 1991. Disponível em: <<http://www.scielo.br/pdf/ea/v5n13/v5n13a15.pdf>>. Acesso em: 03 dez. 2009.

MARCONDES, C. H.; SAYÃO, L. F. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. *Ci. Inf.* [online], v.30, n.3, 2001. Disponível em: <<http://dici.ibict.br/archive/00000211/>>. Acesso em: 15 nov. 2009.

MORENO, F. P.; ARELLANO, M. A. M. Publicação científica em arquivos de acesso aberto. *Arquivística.net*, Rio de Janeiro, v.1, n.1, 2005. Disponível em: <http://www.periodicos.ufrgs.br/admin/sobrelinks/arquivos/Publicacao_acesso_aberto.pdf>. Acesso em: 05 maio 2010.

MORIGI, J. V.; PAVAN, C. Tecnologias de informação e comunicação: novas sociabilidades nas bibliotecas universitárias. *Ci. Inf.* [online], v.33, n.1, 2004. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/72/69>>. Acesso em: 20 out. 2009.

PACHECO, R. C. S.; KERN, V. M. Transparência e gestão do conhecimento por meio de um banco de teses e dissertações: a experiência do PPGE/UFSC. *Ci. Inf.* [online], v.30, n.3, 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-9652001000300009>. Acesso em: 20 out. 2009.

Portal do Sistema Eletrônico de Editoração de Revistas. Disponível em: <<http://seer.ibict.br/>>. Acesso em 12 nov. 2009.

RCAAP – Repositório Científico de Acesso Aberto de Portugal. Disponível em: <http://projecto.rcaap.pt/index.php?option=com_content&view=article&id=24&Itemid=36&lang=pt>. Acesso em 12 nov. 2009.

SANT'ANNA, M. L. Os desafios da preservação de documentos públicos digitais. *Revista IP*, v.3, n.2, 2001. Disponível em: <http://www.ip.pbh.gov.br/ANO3_N2_PDF/ip0302santanna.pdf>. Acesso em: 18 nov. 2009.

SARAMAGO, M. L. Metadados para preservação digital e aplicação do modelo OAIS, [s.d], Guimarães - Portugal. *Anais...* Guimarães: APBDA: [s.d.]. Disponível em: <<http://badinfo.apbad.pt/congresso8/comm2.pdf>>. Acesso em: 18 nov. 2009.

SAYÃO, L. F. Padrões para bibliotecas digitais abertas e interoperáveis. *Pesq. Bras. em Ci. Info. e Biblio*, v.2, n.2, 2007. Disponível em: <<http://revista.ibict.br/pscib/index.php/pscib/article/view/794>>. Acesso em: 22 out. 2009.

SENA, N. K. Open archives: caminho alternativo para a comunicação científica. *Ci. Inf.* [online], v.29, n.3, 2000. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewFile/227/202>>. Acesso em: 18 nov. 2009.

TEDE - Sistema de publicação eletrônica de teses e dissertações. Disponível em: <http://tedesite.ibict.br/> Acesso em: 12 nov. 2009.

SOUZA, C. M. Aviso aos navegantes ou onde fica a biblioteca? *Transifo.*, v.9, n.2, 1997. Disponível em: <<http://www.csouza952.com.br/aviso.pdf>>. Acesso em: 02 out. 2009.

TALAGALA, D. Web interfaces for CDS/ISIS. GENISISweb version 3.0.0. *Manual...* Colombo: Sri Lanka Library Association, 2003. Disponível em: <<http://www.unesco.org/isis/files/winisis/genisis/web/genisisman.pdf>>. Acesso em: 25 março 2010.

TAMMARO, A. M.; SALARELLI, A. *A biblioteca digital*. Brasília: Briquet de Lemos, 2008.

TERENCE, A. C. F.; ESCRIVÃO FILHO, E. Abordagem quantitativa, qualitativa e a utilização da pesquisa-ação nos estudos organizacionais. In: XXVI ENEGEP - Fortaleza, 2006. *Anais...* Fortaleza: ABRPRO, 2006. Disponível em: <http://www.abepro.org.br/biblioteca/ENEGEP2006_TR540368_8017.pdf>. Acesso em: 5 dez 2009.

TORRES, E. F.; MAZZONI, A. A. Conteúdos digitais multimídia: o foco na usabilidade e acessibilidade. *Ci. Inf.* [online], v.23, n.2, 2004. Disponível em <<http://www.scielo.br/pdf/ci/v33n2/a16v33n2.pdf>>. Acesso em: 18 junho 2009.

TRIPP, D. Pesquisa-ação: uma introdução metodológica. *Educ. e Pesq.*, v.31, n.3, 2005. Disponível em: <<http://www.scielo.br/pdf/ep/v31n3/a09v31n3.pdf>>. Acesso em: 20 junho 2009.

VALLE, Eduardo. *Sistemas de informações multimídia na preservação de acervos permanentes*. 2003. 128p. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Minas Gerais, Belo Horizonte. Disponível em <<http://www.eduardovalle.com/permalink.htm?doc=doc00001>>. Acesso em: 10 nov. 2009.

VINCENTINI, L. A. A homepage e a Internet como instrumentos de disponibilização dos serviços bibliotecários. In: Seminário sobre Automação em Bibliotecas e Centros de Documentação, n.6, 1997. *Anais...* Águas de Lindóia: [s.d.], 1997.