

**UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE COMUNICAÇÃO E BIBLIOTECONOMIA
CURSO DE BIBLIOTECONOMIA**

DENISE EUGÊNIA MOREIRA

**A IMPORTÂNCIA DA INDEXAÇÃO EM BIBLIOTECAS DIGITAIS
COMO APOIO AO ENSINO A DISTÂNCIA: PROPOSTA DE POLÍTICA DE
INDEXAÇÃO PARA O CENTRO INTEGRADO DE APRENDIZAGEM EM REDE DA
UNIVERSIDADE FEDERAL DE GOIÁS.**

**Goiânia
2008**

DENISE EUGÊNIA MOREIRA

**A IMPORTÂNCIA DA INDEXAÇÃO EM BIBLIOTECAS DIGITAIS
COMO APOIO AO ENSINO A DISTÂNCIA: PROPOSTA DE POLÍTICA DE
INDEXAÇÃO PARA O CENTRO INTEGRADO DE APRENDIZAGEM EM REDE DA
UNIVERSIDADE FEDERAL DE GOIÁS.**

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia, da Faculdade de Comunicação e Biblioteconomia, da Universidade Federal de Goiás, para obtenção do título de Bacharel em Biblioteconomia.
Orientação: Prof^a. Tatiane Ferreira
Co-orientação: Prof^a. Leda Moreira Nunes Mendonça.

**Goiânia
2008**

M838i

Moreira, Denise Eugênia.

A importância da indexação em bibliotecas digitais como apoio ao ensino a distância: proposta de política de indexação para o Centro Integrado de Aprendizagem em Rede da Universidade Federal de Goiás. / Denise Eugênia Moreira. Goiânia, 2008. 59p.

Monografia, apresentada à Universidade Federal de Goiás – Faculdade de Comunicação e Biblioteconomia. Orientadora: Tatiane Ferreira.

Co-orientadora: Leda Moreira Nunes Mendonça.

1. Biblioteca digital.
2. Ensino a distância.
3. Políticas de Indexação.
4. Recuperação da Informação.

002:004

DENISE EUGÊNIA MOREIRA

**A IMPORTÂNCIA DA INDEXAÇÃO EM BIBLIOTECAS DIGITAIS COMO APOIO
AO ENSINO A DISTÂNCIA: PROPOSTA DE POLÍTICA DE INDEXAÇÃO PARA O
CENTRO INTEGRADO DE APRENDIZAGEM EM REDE DA UNIVERSIDADE
FEDERAL DE GOIÁS.**

Trabalho de Conclusão de Curso defendido no Curso de Biblioteconomia da
Universidade Federal de Goiás, para obtenção de grau de Bacharel,
provado em ____ de _____ de _____, pela
Banca Examinadora constituída pelos seguintes professores:

Prof.^aLeda Moreira Nunes Mendonça

Prof.^a Tatiane Ferreira

Prof.^a Dr.^a Suely Henrique de Aquino Gomes

Goiânia
2008

Aos meus pais Esly e Benedito, ao
meu companheiro José dos Reis e
aos meus filhos Diovanny, Mychelle e
Leandro.

AGRADECIMENTOS

À Deus pelo dom da vida que me conferiu força, consolo e abrigo ao longo da minha jornada.

Aos meus pais pelo amor abnegado; aos meus irmãos pelo zelo; ao meu companheiro por ser meu "porto-seguro"; aos meus filhos por me ensinarem o que é amar; e às minhas sobrinhas por simplesmente existirem.

Às minhas amadas amigas Inês, Kellen, Jaqueline Rocha e Taísse, pela ajuda incondicional e compartilhamento de momentos bons e ruins.

À Professora Leda pela orientação e incentivo os quais foram indispensáveis para a elaboração deste trabalho.

À Professora Tatiane pela orientação e disponibilidade.

À Professora e Coordenadora Suely pelo apoio, compreensão e benevolência os quais jamais serão esquecidos.

Aos professores do Departamento de Biblioteconomia da Universidade Federal de Goiás por terem me conferido a honra de ser sua aluna e o amor à Biblioteconomia

Aos queridos Cirlene, Thomaz, Moisés e Ideltônio funcionários da secretária da Facomb, pela disponibilidade em atender às necessidades dos alunos.

À Universidade Federal de Goiás pela realização do sonho da aprovação em seu concurso vestibular.

Á todos que perto ou não, acreditaram em mim e não me permitiram desistir.

Finalmente: Aos que desacreditaram, pois sem vocês não teria tanta graça em concluir um projeto!

“Grandes realizações são possíveis quando se dá importância aos pequenos começos.”

Lao Tzu

RESUMO

Ensino a distância é o processo de ensino-aprendizagem, mediado por tecnologias, onde professores e alunos estão separados espacial e/ou temporalmente. No EAD a biblioteca digital é de extrema importância, visto que o acesso à informação neste modelo de ensino se dá preferencialmente através deste tipo de biblioteca. Para se obter eficiência neste processo a Biblioteca Digital necessita ter uma política de indexação que deve ser constituída de estratégias que permitam o alcance dos objetivos de recuperação do sistema de informação. O objetivo deste trabalho é sugerir um modelo de política de indexação ao Centro de Aprendizagem em Rede da Universidade Federal de Goiás com base em pesquisa documental com fins comparativos de propostas, revisão da literatura sobre os processos de indexação e análise da Universidade de São Paulo instituição de nível superior que já trabalha com modelo de Biblioteca digital. Os resultados obtidos demonstraram que a metodologia pode e foi utilizada para definição de uma política de indexação para que se tenha eficiência no acesso aos documentos inseridos no sistema utilizado pelo CIAR. Conclui que, a indexação é um importante processo que pode nortear a área de representação e recuperação da informação.

SUMÁRIO

	RESUMO.....	8
1	INTRODUÇÃO.....	12
2	METODOLOGIA.....	14
3	O ENSINO A DISTÂNCIA.....	15
4	O CIAR.....	16
4.1	ESTRUTURA ORGANIZACIONAL DO CIAR.....	16
4.2	PÓLOS DE ENSINO A DISTÂNCIA.....	17
5	BIBLIOTECAS INFORMATIZADAS.....	18
6	A RECUPERAÇÃO DA INFORMAÇÃO.....	19
7	HISTÓRICO DA INDEXAÇÃO.....	21
7.1	A INDEXAÇÃO.....	23
7.2	INDEXAÇÃO ALFABÉTICA DE ASSUNTO.....	23
7.3	INDEXAÇÃO COORDENADA.....	24
7.4	TESAURO.....	25
7.5	INDEXAÇÃO AUTOMÁTICA.....	26
7.6	LEITURA DOCUMENTÁRIA.....	26
7.7	TEMATICIDADE.....	27
7.8	ASPECTOS LINGÜÍSTICOS.....	27
7.9	ASPECTOS LÓGICOS.....	28
7.10	ASPECTOS COGNITIVOS.....	28
7.11	PROCEDIMENTOS DE INDEXAÇÃO.....	29
7.12	AMBIGÜIDADE.....	30
7.13	METADADOS.....	31
8	RESULTADOS E DISCUSSÃO DA PESQUISA.....	32
8.1	A UNIVERSIDADE DE SÃO PAULO (USP).....	32
8.2	SISTEMA DE BIBLIOTECAS.....	33
9	ANÁLISE DOS DADOS DA BIBLIOTECA DIGITAL DE TESES E DISSERTAÇÕES DA UNIVERSIDADE DE SÃO PAULO.....	33
9.1	OBJETIVOS DA BDTD DA USP.....	35
9.2	COOPERAÇÃO.....	35
9.3	SOFTWARE UTILIZADO.....	35
9.4	ACERVO EXISTENTE.....	36
9.5	ACESSO AO ACERVO.....	36

9.6	BUSCA.....	37
9.7	AUTORIZAÇÃO DE PUBLICAÇÃO.....	40
9.8	FORMATO DOS DOCUMENTOS.....	41
9.9	RESTRIÇÃO DE USO E PUBLICAÇÃO.....	41
9.10	ESTRUTURA DE APRESENTAÇÃO DO DOCUMENTO.....	43
9.11	FORMATOS DE ARQUIVOS.....	44
9.12	INDEXAÇÃO ADOTADA.....	45
9.13	COMUNICAÇÃO ATRAVÉS DE CORREIO ELETRÔNICO.....	46
9.14	AVALIAÇÃO DO SISTEMA UTILIZADO.....	47
10	PROPOSTA DE MODELO DE POLÍTICA DE INDEXAÇÃO PARA O CIAR.....	47
10.1	ETAPAS PARA CRIAÇÃO DE POLÍTICA DE INDEXAÇÃO.....	49
11	CONSIDERAÇÕES FINAIS.....	50
	REFERÊNCIAS.....	52

ANEXOS.....	56
ANEXO A – QUESTIONÁRIO.....	57
ANEXO B – ROTEIRO PARA ANÁLISE da BDTD da USP.....	59

1 INTRODUÇÃO

No Ensino a Distância (EAD) uma das preocupações mais freqüentes é como recuperar as informações contidas em documentos disponibilizados via internet nas bibliotecas digitais. Para recuperá-los se faz necessário indexá-los, ou seja, representá-los de forma que possam ser acessados independentes do ambiente em que estejam.

O EAD surgiu no final do século XVIII, com os cursos feitos por correspondência, tendo o total apoio dos serviços dos Correios. Ao longo do tempo e com o avanço tecnológico, à medida que se tornavam disponíveis foram incorporadas tecnologias de comunicação que permitiam a interação com os demais alunos, professores e tutores. O EAD tem o mesmo valor que os cursos presenciais, sendo reconhecidos pelo Ministério da Educação (MEC) caso as exigências sejam cumpridas. Dentre as exigências do MEC, está o acesso à informação e o acesso a bibliografia básica de cada curso disponibilizado a distância.

O conceito para “Ensino a Distância”, na definição de Trompieri Filho (apud SANTOS, 2004, p. 2) é:

ensino à distância é o termo utilizado para designar o método de instrução na qual a interação entre aluno e professor ocorre de forma intermediada por ferramentas tecnológicas, sejam convencionais (como correio, rádio, televisão) sejam mais modernas como a Internet.

O EAD vem se tornando uma alternativa do mundo globalizado como forma de se obter e aperfeiçoar o conhecimento acadêmico. Muitas pessoas, às vezes, deixam de estudar devido a falta de tempo e necessidade de trabalhar. E como auxílio a este problema, a informática ligada à educação está fazendo chegar até o conforto do lar ou escritório cursos nos quais irão ajudar e muito o desenvolvimento profissional e intelectual. Educação a distância possibilita maior liberdade para que o aluno satisfaça suas necessidades informacionais, seja por modelos tradicionais, não tradicionais ou pela mistura de ambos. Nesse cenário, a biblioteca desempenha importante papel na recuperação, transferência e disseminação de informações para estudantes e demais interessados. Sendo que, o processo de recuperação da informação compreende basicamente três etapas: indexar, armazenar e recuperar.

Santos (2004) pensando nesta nova realidade cita como é importante que saibamos administrar bem a biblioteca, seja ela convencional, virtual, digital ou

híbrida, para oferecer condições favoráveis ao ensino a distância, ou seja, dispor de acervo atualizado, amplo e representativo de livros e periódicos, acervo de imagens, áudio, vídeos, sites na internet à disposição de alunos e professores. Adotar procedimentos que garantam atendimento a cada aluno ou professor, independentemente do local onde ele esteja.

Uma política de indexação deve ser oferecida por uma biblioteca digital para otimizar a mediação da informação e os seus usuários finais, ampliando as possibilidades de acessos aos documentos armazenados, garantindo assim a recuperação eficiente de todas as versões dos materiais produzidos no âmbito dos cursos de EAD oferecidos pela instituição.

Na Universidade Federal de Goiás (UFG) temos o Centro Integrado de Aprendizagem em Rede (CIAR) que desde o ano de 2007 oferece cursos no modelo EAD. Após análise constatou-se que na Biblioteca Digital do CIAR não existe um modelo de política de indexação adotado que possa oferecer uma variedade de serviços que permitirão aos usuários tirar melhor proveito dos dados representados, organizados e disponibilizados em mídia eletrônica, pois a política de indexação seguida pela biblioteca digital é que irá limitar quais os níveis de representação da informação que serão adotados num determinado acervo.

Obviamente tais decisões terão consideráveis implicações na representação do conteúdo a ser utilizado quer no tamanho do arquivo e nos mecanismos para a sua recuperação.

Esses indicadores reforçam a necessidade de se pensar a infra-estrutura informacional de apoio aos alunos *off campus*. Por isso este estudo tem como objetivo geral pesquisar entre instituições de ensino superior que já trabalham com Biblioteca Digital modelos de políticas de indexação para bibliotecas digitais. E como objetivos específicos avaliar dentre as bibliotecas digitais escolhidas para análise o modelo que melhor atende as necessidades do CIAR e após esta análise propor ações que permitam a criação de uma política de indexação para representação dos documentos com fins de recuperação, transferência e disseminação de informações e também identificar qual a linguagem de indexação mais apropriada para a Biblioteca Digital do CIAR.

2 METODOLOGIA

São inúmeros os conceitos destinados à “pesquisa” e ainda não se conseguiu chegar a um consenso sobre o assunto. Como fundamento para este trabalho foi escolhido o conceito de Lakatos e Marconi (1992, p. 155) que define pesquisa como “um procedimento formal com método de pensamento reflexivo, que requer um tratamento científico e se constitui no caminho para conhecer a realidade ou para descobrir verdades parciais”.

Como foi constatado que no CIAR ainda não existe uma política de indexação para os documentos que são e serão disponibilizados para consulta pelo site de sua Biblioteca Digital para prosseguir com a proposta de pesquisa e sugerir um modelo de política de indexação para o CIAR foi definido que os dados básicos para este estudo seriam reunidos por meio de pesquisa documental com fins comparativos de propostas, revisão da literatura sobre os processos de indexação que serviram de suporte para elaboração de questionário e roteiro de análise para fontes de informações digitais e coleta de dados por amostragem intencional, onde o pesquisador usa critério profissional na seleção dos entrevistados. Rea e Parker (2002, p. 150) afirmam que esta é uma amostragem não probabilística que consiste em selecionar um subgrupo de população que, com base nas informações disponíveis, possa ser considerado representativo, de toda a população.

No caso deste trabalho foram escolhidas instituições de ensino superior que já trabalham com o modelo de Biblioteca Digital para, a partir dos dados obtidos, ser possível traçar um perfil do papel da biblioteca informatizada no EAD, ou seja, identificar qual a linguagem de indexação mais apropriada e sugerir propostas e ações que ajudem a definir a política de indexação que será adotada pelo CIAR

Para obtenção dos dados pertinentes a este trabalho foram enviados questionários para as bibliotecas digitais das seguintes instituições:

- a) Universidade de São Paulo (USP);
- b) Universidade Norte do Paraná (UNOPAR);
- c) Universidade do Estado do Rio de Janeiro (UERJ);
- d) Universidade Federal da Bahia (UFBA);
- e) Universidade de Campinas (UNICAMP);
- f) Pontifícia Universidade Católica do Rio (PUC-RIO);

- g) Pontifícia Universidade Católica de São Paulo (PUC-SP);
- h) Universidade Federal de Santa Catarina (UFSC).

Foi estipulado o prazo de uma semana para espera de respostas aos questionários enviados, com contato posterior reiterando a solicitação de colaboração em responder e novamente mais uma semana de prazo foi fixada para retorno.

Quando as instituições não responderam ao questionário enviado, os dados da pesquisa foram colhidos estritamente através do roteiro para análise de fontes de informações digitais. A escolha de um modelo para análise recaiu sobre a Universidade de São Paulo (USP) que é a instituição que, pela sua experiência, está destacada na literatura como referência no assunto Biblioteca Digital com trabalhos e seminários apresentados em congressos.

3 O ENSINO A DISTÂNCIA

O ensino a distância tem sido intensamente discutido no cenário acadêmico, nacional e internacional, em função da necessidade da democratização do ensino. Segundo Belloni (1999) o conceito de ensino a distância diz respeito tanto às estratégias desenvolvidas, quanto a organização do trabalho acadêmico e a produção de materiais pedagógicos visando os direitos de acesso irrestrito à educação e à cultura de todo ser humano, da formação continuada, e da superação dos impedimentos sociais e pessoais (distância geográfica, disponibilidade de horário, compromissos familiares e profissionais, etc.). A incorporação nas universidades das novas tecnologias e recursos informacionais tem proporcionado a implantação e intensificado a implementação dos programas de EAD.

No Brasil, observa-se que o EAD tem sido amplamente adotado por várias instituições de ensino superior inclusive a UFG. As iniciativas de implantação de cursos a distância buscam consolidação para os mesmos junto aos órgãos competentes que com base no decreto (que regulamenta este tipo de ensino) nº. 2.494 de 10 de fevereiro de 1998, art. 80 da LDB (Lei nº. 934/96), prevê a existência da biblioteca informatizada, como item obrigatório para a aprovação destes cursos, conforme destacamos abaixo:

Portaria nº. 301, de 7 de abril de 1998, dispõe sobre a necessidade de normatizar os procedimentos de credenciamento de instituições para a oferta de cursos de graduação e educação profissional tecnológica a distância, resolve:

Art. 3º A solicitação para o credenciamento do curso de que trata o § 1º deverá ser acompanhada de projeto, contendo pelo menos, as seguintes informações:

IV – descrição da infra-estrutura, em função do projeto a ser desenvolvido: instalações físicas, destacando salas para atendimento aos alunos; laboratórios; biblioteca atualizada e informatizada, com acervo de periódicos e livros, bem como fitas de áudio e vídeo. (BRASIL, 1998, p. 2)

4 O CIAR

As informações apresentadas a seguir têm como base a página do CIAR da UFG (acesso em 2008). O Centro Integrado de Aprendizagem em Rede (CIAR) da Universidade Federal de Goiás é um órgão suplementar vinculado a Reitoria que tem como competência a implementação e apoio de atividades acadêmicas de graduação, pós-graduação, extensão e pesquisa, integradas pelas tecnologias da informação e comunicação, desenvolvidas na UFG e tem por objetivos acompanhar projetos interinstitucionais e intra-institucionais que envolvem aprendizagem em redes com apropriação de tecnologias da informação e comunicação; promover atividades de formação continuada para profissionais envolvidos nos projetos de aprendizagem em rede desenvolvidos pela UFG; promover articulação pedagógica entre os projetos de EAD realizados pela UFG; orientar e coordenar a produção de material didático apropriado aos projetos de EAD; contribuir e incentivar a criação da cultura de EAD junto à comunidade da UFG. Visa ainda constituir grupos de pesquisa com integrantes das equipes do CIAR e outros pesquisadores com interesse em avaliar e validar os instrumentos didáticos, tecnologias, concepções, metodologias, sistema de gestão e ambientes virtuais de aprendizagem e outros temas específicos aos projetos realizados na UFG; incentivar o investimento de pesquisas e investigações científicas e técnicas nas diversas áreas do conhecimento científico; realizar atividades que contribuam para a inclusão digital e o desenvolvimento de competências e habilidades para o avanço da comunidade do Estado de Goiás.

4.1 ESTRUTURA ORGANIZACIONAL DO CIAR

Na estrutura organizacional da UFG o CIAR é um órgão suplementar, pois dentro do organograma da universidade há a reitoria, a vice – reitoria, as pró – reitorias, as comissões especiais, os órgãos complementares, órgãos suplementares, órgãos administrativos e as unidades acadêmicas e que se dividem

em três instâncias, da seguinte forma: Conselho Gestor, Diretoria e Coordenadoria de Equipe.

O CIAR oferece os seguintes cursos de graduação: Bacharelado de Administração; Licenciatura em Artes cênicas, Artes Visuais; Física, Biologia e Educação Física. Cursos de pós – graduação em: Especialização em Metodologia do Ensino Fundamental e Especialização em gestão escolar.

Oferece também as atividades de extensão desenvolvidas pelo CIAR, que são realizadas pelas unidades acadêmicas da UFG ou até mesmo pelo próprio CIAR em convênio com outras instituições superiores e instâncias federais, estaduais e municipais. Os cursos oferecidos são: Geo-processamento, Moodle para professores, Mídias na Educação Gestão de Pólos em EAD, Rede de Conhecimento: produção de material em EAD, Formação, Pedagógica em EAD, Formação de Orientadores e Tutores.

E ainda existe um Grupo de Estudo Novas Tecnologias e Educação (GENTE) que surgiu no ano de 2005 e desde então vem expandindo suas ações em diversas instituições, com o objeto de investigação temáticas envolvendo as tecnologias educacionais.

4.2 PÓLOS DE ENSINO A DISTÂNCIA

Estão sendo beneficiados municípios do interior goiano que constituem pólos de ensino da Universidade Aberta do Brasil.

Os pólos de ensino a distância do CIAR se encontram nas seguintes cidades: Alexânia, Alto Paraíso, Anápolis, Aparecida de Goiânia, Catalão, Ceres, Cezarina, Cidade de Goiás, Firminópolis, Formosa, Goianésia, Goiânia, Iporá, Jataí, Luziânia, Morrinhos, Porangatú, Quirinópolis São Simão e Uruana.

Para Blattmann e Belli (2000) a Internet pode fortalecer o ensino a distância e as bibliotecas, mas somente o trabalho conjunto entre educadores e bibliotecários possibilitará a integração entre as novas tecnologias e a explosão informacional. As barreiras econômicas e geográficas para acessar a Internet estão acabando e assim se torna necessário observar os impactos da Internet no treinamento profissional,

educacional e social e a necessidade de desenvolver novas ferramentas e interfaces para atender as demandas informacionais dos usuários.

Teixeira (1997) ressalta que a importância do comprometimento da instituição no fornecimento de serviços de suporte aos alunos *off campus* devem ser equivalentes aos disponíveis para alunos *on campus*, incluindo acesso à base de dados *on line* e recursos informacionais e que o nível dos serviços a serem desenvolvidos dependerá diretamente de informações como: formato dos cursos, número de alunos envolvidos, bibliografia básica de cada curso, como serão fornecidos os materiais, o nível das instruções e a forma de comunicação.

5 BIBLIOTECAS INFORMATIZADAS

Para Cunha (1999, p. 258), a biblioteca informatizada é definida por dezenas de autores por encontrar uma ou várias das características seguintes:

- a) acesso remoto pelo usuário, por meio de um computador conectado a uma rede;
- b) utilização simultânea do mesmo documento por duas ou mais pessoas;
- c) inclusão de produtos e serviços de uma biblioteca ou centro de informação;
- d) existência de coleções de documentos correntes onde se pode acessar não somente a referência bibliográfica, mas também o seu texto completo;
- e) provisão de acesso em linha e outras fontes externas de informação (bibliotecas, museus, bancos de dados, instituições públicas e privadas);
- f) utilização de maneira que a biblioteca local não necessite ser proprietária do documento solicitado pelo usuário;
- g) utilização de diversos suportes de registro da informação, tais como: texto, som, imagem e números;
- h) existência de unidade de gerenciamento do conhecimento, que inclua sistema inteligente ou especialista para ajudar na recuperação de informação mais relevante.

A Biblioteca informatizada engloba os conceitos de biblioteca virtual, eletrônica, digital e híbrida, que serão especificados a seguir:

Segundo LaCruz (apud FUJITTA, 2005) biblioteca virtual é uma denominação que faz referência a uma coleção de documentos eletrônicos, alojados na WEB e sem local físico, organizado e postos à disposição de usuários que vão acessá-los *on-line*, à distância, pela rede de tele processamento em seus endereços eletrônicos.

Para Arnold, et al. (apud BLATTMANN; BELLI, 1993) o conceito de biblioteca eletrônica é tratado como um ambiente de ensino, aprendizagem e estudos para

educação superior onde a informação é considerada primeiramente na forma eletrônica, sendo não restrita ao local físico, permitindo acesso de qualquer lugar, devendo possuir acervo de periódicos e livros, bem como fitas de áudio e vídeo. Necessita, portanto, estar conectada (linkada) com a indústria de publicação e livrarias.

Para Pereira e Rutina (1999) a biblioteca digital é aquela que tem, além de seu catálogo, os textos dos documentos de seu acervo armazenados de forma digital, permitindo sua leitura na tela do monitor ou sua importação (*download*) para o disco rígido do computador, ou outra forma de armazenamento.

De acordo com Rusch-Feja (apud GARCEZ; RADOS, 1999), biblioteca híbrida é o conceito que deve integrar o acesso a diferentes tecnologias para o mundo da biblioteca digital e através de diferentes mídias, refletindo o estado transacional da biblioteca, objetivando o enfoque da importância do papel das bibliotecas híbridas no atendimento aos vários tipos de usuários dos cursos a distância, por meio da flexibilização dos bens e serviços oferecidos e da integração dos suportes impresso e digital, nesse momento de transição por que passam as bibliotecas convencionais.

Devido a similaridade conceitual dos tipos de biblioteca informatizada, será usada neste trabalho a denominação biblioteca informatizada para designar os modelos de biblioteca virtual, eletrônica, digital e híbrida.

6 A RECUPERAÇÃO DA INFORMAÇÃO

Teixeira (1997) explica que a explosão documentária aumentou significativamente a dificuldade de recuperação das informações em sistemas manuais. Com o advento da informática, indexar, armazenar e recuperar as informações tornaram-se tarefas mais simples e eficientes, por haver recursos que permitem maior rapidez nestes três processos.

Cunha (1994) afirma que inicialmente, o computador foi usado nas bibliotecas para armazenar informações bibliográficas e gerar índices impressos que poderiam ser consultados da mesma maneira que os índices manuais. Hoje, os sistemas informatizados oferecem uma infinidade de recursos de busca, como por exemplo, as consultas com combinações complexas de termos, que não podemos comparar aos dos sistemas manuais. Nas novas tecnologias de recuperação da informação, a

Internet surge para reduzir drasticamente as distâncias por intermédio das redes de transmissão de dados e a multiplicação, não só da potência e velocidade de processamento, mas também dos recursos lógicos por meio de redes locais e remotas, que se integram nos chamados sistemas abertos e equipamentos informáticos de tipos mais variados.

Para Teixeira (1997) o avanço tecnológico, que possibilita a disseminação da informação por computadores, CD-ROM e outros meios de armazenamento, torna-nos cada dia mais exigente, buscando meios mais eficientes de armazenar e de principalmente recuperar a informação. As bibliotecas e os demais serviços de informação construíram suas bases de dados bem como os mecanismos necessários para que essas bases sejam consultadas remotamente por meio das redes de telecomunicações em nível nacional e internacional.

Os trabalhos mais recentes na área da recuperação da informação baseiam-se na premissa de que ferramentas de busca, ao fazerem uso da linguagem natural, necessitam de conhecimento sobre o significado das expressões que são tratadas e das relações que se estabelecem entre elas. Essas ferramentas devem, ainda, ser capazes de tratar determinados fenômenos lingüísticos que afetam a qualidade da recuperação, como o da ambigüidade, a qual será tratada no âmbito deste trabalho.

Segundo Cunha (2000), as bibliotecas informatizadas têm como desafio prover a recuperação eficiente de documentos aos seus usuários, porque na medida em que a informação digital se expande necessita-se utilizar arquivos de textos completos de periódicos, imagens digitais, dados numéricos e multimídia. Esta recuperação eficiente só é possível com a preocupação especial na indexação, organização e armazenagem do acervo, pois com o advento da Internet outros tipos de documentos tais como: páginas iniciais (home-pages) e periódicos eletrônicos merecem a especial atenção dos profissionais designados para o processamento técnico das informações, pois são verdadeiras obras de referência, cuja importância cresce a cada dia e estão provocando a criação de novos padrões para a eficiente descrição dos formatos que melhoram os requisitos para seus acessos e usos.

No Código de Catalogação Anglo-Americano (AACR-II) e no formato MARC as normas contidas se mostram insuficientes para atender as novas necessidades técnicas, os catalogadores, além de conhecer seus instrumentos de trabalho, necessitam dominar outros, tais como metadados e marcação de textos, e também estar aptos a lidar com as características multimídia dos novos documentos. O setor de processamento técnico é desafiado a prover novos meios de descrever o registro e o conteúdo de itens com estruturas informacionais e manipulação bem diferentes daqueles tradicionalmente arrolados pelo controle bibliográfico. Esta preocupação se deve à natureza efêmera de muitos recursos da Internet, e por esse motivo muitas bibliotecas somente catalogarão aquelas fontes consideradas de qualidade ou que possuam certa segurança de acesso e confiabilidade.

7 HISTÓRICO DA INDEXAÇÃO

A indexação surgiu com a atividade de elaboração de índices, como instrumento de armazenagem e recuperação da informação. O modelo mais antigo de indexação que se tem conhecimento são as tábuas de argila que foram produzidas na Mesopotâmia no século II a.C. Nestas tábuas foram grafadas uma espécie de resumo de livros, constatando o conteúdo informacional dos mesmos.

Na biblioteca de Alexandria, organizada pela classificação de Calímaco, o catálogo era organizado em ordem alfabética de autores e subordinados a assuntos mais gerais, onde as obras como histórias e peças dos dramaturgos da época eram condensadas. Vale salientar que o primeiro tipo de indexação existente era baseado na memória, pois textos célebres como grandes epopéias eram transmitidas oralmente e sua indexação era feita pela primeira sentença de cada parágrafo.

Muitas foram as formas utilizadas de indexação até chegarmos à padronização que se conhece hoje. Podemos citar no século II, o trabalho de Cláudio Galeno intitulado *Libris Propriis Liber* que determinou o aparecimento dos primeiros guias para obras isoladas: cabeçalhos de capítulos, tábuas de matéria ou sumários, cabeçalhos nas margens dos parágrafos, ou descritos no alto das páginas.

Segundo Silva e Fujita, (2004, p. 139):

era comum, no século XIV, a elaboração de inventários ou catálogo dos livros existentes nos mosteiros, nessa época, eles tinham a elaboração de listas de completas como um objetivo definido. Esses monges copistas, tentando esclarecer e indicar os pontos principais dos assuntos tratados em trechos ou

parágrafos mais longos dos livros, escreviam às margens algumas palavras ou sentenças que indicassem o conteúdo.

Somente no século XVII é que se iniciou a grande época do índice facilitada pela Reforma Protestante que possibilitou a tradução da Bíblia tornando-a acessível ao público em geral. Com a disseminação dos procedimentos de indexação, surge na Alemanha a contribuição da idéia de palavra-chave na representação de um item com o sistema de escolha da *schlagwort* (palavra-chave), que representou um avanço para a busca da informação. Até a invenção da imprensa, os índices eram a única forma de acesso aos livros utilizados nas bibliotecas dos mosteiros, a partir do registro dos títulos dos livros. A partir daí, houve um crescente aumento da literatura que estimulou o surgimento de várias listas com diversas finalidades.

A indexação é conhecida com uma ação de descrever, identificar e recuperar um determinado documento de acordo com sua área do conhecimento. Silva e Fujita (2004) nos permite afirmar que os índices antigos não tinham uma indexação, ou “indexação crítica”. A literatura da área alega que isso se justifica pelo fato de que as pessoas da época desconheciam o processo analítico que nos dias atuais consiste na indexação e por esse motivo, os índices apresentam uma entrada limitada e simples como nomes próprios ou entradas por acontecimentos. A literatura da área registra também, que a história da indexação tem seu início juntamente com a história da bibliografia e que a indexação teve maior delicadeza e atenção a partir do aparecimento das publicações periódicas. Esse fato impulsionou o surgimento da necessidade de elaboração de técnicas para a organização por assunto do conteúdo que esse tipo de publicação tratava.

A partir do lançamento de *Reader's Guide to Periodical Literature* por Wilson (apud SILVA; FUJITA, 2004, p. 140) surgiram muitos índices bons, tanto específicos como gerais. O aperfeiçoamento de técnicas da indexação acompanhava o crescimento da publicação de livros. Silva e Fujita (2004) citam como exemplo da melhoria da técnica de indexar livros, o índice da *Encyclopaedia Britânica* que representa o aprimoramento das técnicas da indexação no século XIX, dada a clareza e apresentação gráfica de seus índices.

O século XIX representou a grande fase da indexação, devido ao grande crescimento da massa documental. A indexação evoluiu dos índices das obras isoladas para os índices de vários volumes chegando aos índices cooperativos e em nível internacional.

7.1 A INDEXAÇÃO

Segundo Silva e Fujita (2004), a partir da revisão da literatura fundamental verificou-se uma explosão dos princípios teóricos e metodológicos da indexação e a partir daí, obteve-se subsídios teóricos para a discussão de dúvidas/afirmações presentes na evolução da indexação.

Considerando-se a indexação como tema principal de nosso estudo, a revisão de literatura fundamental dividiu-se em: indexação alfabética de assunto; indexação coordenada, sobretudo a pré-coordenação dos cabeçalhos de assunto; indexação automática e a relação entre o indexador humano e o automático; leitura em indexação os aspectos ligados a ela como os aspectos lógicos, lingüísticos e cognitivos e a tematicidade, que influenciam na identificação dos conceitos que irão representar o conteúdo do documento; e os procedimentos de indexação, com base nos princípios do Sistema Internacional de Informação para a Ciência e Tecnologia (UNISIST), Associação Brasileira de Normas Técnicas (ABNT) e na abordagem sistemática de identificação de conceitos do sistema de indexação (PRECIS).

7.2 INDEXAÇÃO ALFABÉTICA DE ASSUNTO

A indexação alfabética de assuntos e a catalogação de assuntos são equivalentes porque são resultados de um mesmo processo: a análise de assunto. Lancaster (1993) afirma que a indexação de assunto é idêntica à catalogação de assunto. A utilização do termo “catalogação de assunto” é influência dos Estados Unidos da América reforçado pelo uso das listas de cabeçalho de assuntos por diversas instituições. A origem do termo catalogação de assuntos está ligada à construção dos catálogos de bibliotecas, essencialmente, do catálogo de assuntos que é organizado mediante determinação de cabeçalhos de assuntos que funcionam como enunciados de assuntos formados a partir da composição ordenada de palavras.

A primeira tentativa de sistematização para a representação alfabética de assuntos, foram os cabeçalhos de assunto. Seu surgimento deu-se aos seguintes fatores: os títulos das obras não representavam de forma adequada o assunto nelas

tratado; problemas relativos às subdivisões de assuntos; existiam obras com mais de um assunto: a interdisciplinaridade dos assuntos em uma mesma obra; obras que relacionavam os assuntos a lugares e épocas diferentes.

Os princípios básicos que norteiam Cutter a estabelecer um cabeçalho de assunto estão pautados em princípios fundamentais que, segundo Cesarino e Pinto, (apud SILVA; FUJITA, 2004, p. 143), podem ser resumidos em:

- a) princípio específico - os assuntos são representados pela entrada do termo mais específico, desconsiderando a classe de assunto a qual está subordinado;
- b) princípio de uso - prevê a necessidade do usuário;
- c) princípio sindético - desenvolvimento nas listas de cabeçalhos de assunto de estruturas sindéticas, entendidas como uma rede de remissivas cruzadas para superar o problema de entrada alfabética dos cabeçalhos de assunto.

Segundo Silva e Fujita (2004) a obra de Cutter representa o primeiro estudo dos problemas relacionados à indexação alfabética de assunto, as implicações de entradas específicas de assuntos, Cutter também estudou as dificuldades da utilização dos termos simples, compostos e geográficos.

A notação de índice, ao longo do tempo, se encontrou muito ligada ao processo de indexação. Outrora, os índices existentes em sistemas de recuperação da informação foram pensados dentro de um panorama classificatório, devido aos grandes cabeçalhos de assuntos serem compostos por influência da terminologia classificatória e não do texto ou de seu conteúdo.

A análise de facetas representa o grande elemento transformador dentro do sistema de indexação alfabética. Essa análise foi proposta por Kaiser e Ranganathan, eles determinaram uma maior especificidade e uniformidade com o uso dos conceitos elementares da indexação: espaço, tempo, processo, concreto, coisa, ação, etc.

7.3 INDEXAÇÃO COORDENADA

Silva e Fujita (2004) caracterizam esse tipo de indexação por sua composição de assunto usada para representar o conteúdo informacional de um determinado

documento. A diferença básica que caracteriza a indexação pré e pós-coordenada é que na indexação pré-coordenada os termos são agrupados no ato da elaboração do índice e já na indexação pós-coordenada os termos são agrupados no ato da busca para recuperação da informação.

7.4 TESAURO

Para Jesus (2002, p. 4) as linguagens documentárias em uso nas bases terminológicas, como instrumento de representação e recuperação permitem uma eficiente comunicação entre a informação disponibilizada e o usuário. No entanto, este registro necessita de uma inter-relação com os conceitos indexados e a conseqüente alteração no significado dos conceitos existentes na área do conhecimento e de suas relações com as outras áreas. No intuito de solucionar este problema, percebem-se esforços científicos, cada vez maiores, no sentido de padronizar a terminologia usada pelas várias ciências.

Jesus (2002, p. 4) afirma que:

só o conjunto de termos estruturados numa linguagem particular pós-coordenada em sua forma e conteúdo, é que permitirá gerar novos termos e/ou alterar significados já existentes, que garanta ao pesquisador no processo de busca e seleção da informação e do documento, encontrar a resposta de que necessita numa linguagem padronizada e de qualidade.

Nos Sistemas de Recuperação da Informação, vários são os instrumentos utilizados para representar o conhecimento de uma ciência, e um deles é denominado Tesouro que é um instrumento dinâmico que apresenta terminologias estruturadas, que permitem atualização periódica e a identificação de conceitos relacionada à sistematização e representação de uma determinada área do conhecimento e definido por Cavalcante (apud JESUS, 2002, p. 3) como:

uma lista estruturada de termos associada empregada por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entrada, e para permitir aos pesquisadores a recuperação da informação que procura

Ou seja, Tesouro é uma linguagem documentária, que apresenta algumas peculiaridades, pois sua hierarquia de assuntos possui uma relação associativa e

sua estrutura não se baseia em conceito de palavras simplesmente, mas sim de termos conceituados e relacionados aos assuntos disponíveis para consulta.

7.5 INDEXAÇÃO AUTOMÁTICA

A indexação automatizada é representada pelo resultado do trabalho intelectual de um profissional para checagem do valor dos termos atribuídos a determinado documento através de um programa de computador.

Na década de 60 surge um novo método de indexação, o Key-Word-in Context (KWIC) – indicação pela palavra-chave no contexto. O KWIC representa a primeira indexação automática de documentos técnicos, tendo com base as palavras significativas dos títulos. O índice KWIC representa a linguagem natural pois não há um controle de termos e os sinônimos não são idênticos.

Já o índice Keywuord-out-of-Context (KWOC) - palavra fora do contexto, para Lancaster (1993, p. 48) é:

aquele em que a palavra-chave usada como ponto de entrada não se repete no título, mas é substituído por asterisco (*) ou outro símbolo. Nesse tipo de índice, as palavras significativas do título são retiradas e organizadas em ordem alfabética e não ocorre uma permutação do título como no KWIC; o título aparece na mesma ordem de sua apresentação e a “palavra significativa” em linguagem natural fica na ordem alfabética como cabeçalho.

Os índices automáticos possuem um grande problema, não representam os assuntos dos documentos da mesma forma que um homem o faz. Isso se explica pelo fato de que ainda se desconhece o processo mental de análise de assunto envolvido durante o processo da indexação. O computador, por sua vez, representa maior agilidade em tarefas respectivas mais simples, em que a análise conceitual não se faz necessária.

7.6 LEITURA DOCUMENTÁRIA

Segundo Silva e Fujita (2004) a leitura documentária faz-se presente no momento em que o indexador realiza a análise do assunto em um documento a fim de representá-lo em termos de indexação. Essa leitura está condicionada a determinados objetivos e ao contexto do sistema de informação. Torna-se, pois, difícil, dissociar esses aspectos da atividade de leitura.

A leitura de indexação difere da leitura tratada de modo geral por possuir finalidades profissionais e pragmáticas. No entanto os conhecimentos basilares necessários para uma boa compreensão de um texto são comuns a ambas.

A leitura documentária caracteriza-se, de acordo com Ginez de Lara (apud SILVA; FUJITA, 2004, p. 147), pela presença de operações seletivas para o processo de identificação e extração de informações. E é realizada sob variáveis que lhe são específicas, dentre elas destacam-se:

- a) instruções textuais;
- b) condições de produção;
- c) circunstâncias de enunciação;
- d) momento e lugar de enunciação.

Silva e Fujita (2004) explicam que essa leitura, realizada pelo leitor-indexador na fase da análise, corresponde à primeira fase de abordagem entre o leitor-indexador e o texto a ser analisado. O indexador, na atividade de leitura para fins documentários, tem por objetivo identificar o tema ou assunto do documento. Deve ter consciência do porque estar lendo um texto, essa consciência torna a leitura meta cognitiva. O leitor indexador durante a atividade de leitura documentária, deve buscar a compreensão, pois somente assim terá condição de concretizar sua atividade de indexação.

7.7 TEMATICIDADE

Para Silva e Fujita (2004) durante a leitura documentária, o indexador tem por objetivo identificar conceitos que compõem o tema do documento, que é a informação relevante abordada no texto, mas é preciso ressaltar que a seleção do tema sofre a influência da política do sistema de indexação, portanto a tematicidade sempre estará relacionada aos interesses informacionais dos usuários do sistema.

7.8 ASPECTOS LINGÜÍSTICOS

Os aspectos lingüísticos na indexação são evidenciados no momento da leitura, pois a “análise do conteúdo é um método de apreensão da informação transmitida por uma língua” (CHAUMIER apud SILVA; FUJITA, 2004, p. 151), sobre tudo no momento da tradução dos termos que representam os conceitos identificados no documento em linguagem natural para a linguagem de indexação.

Aspectos lingüísticos estão presentes na leitura para indexação, por meio da lingüística textual que, segundo Favero e Kock (apud SILVA; FUJITA, 2004, p. 152) é a área responsável por “determinar o que faz com que um texto seja um texto e diferenciar as várias espécies de textos.”

Para Silva e Fujita (2004, p. 152):

Os resultados obtidos no estudo sobre leitura documentária comprovam que se os indexadores conhecerem e explorarem a estrutura textual tendem a encontrar, mais facilmente, a informação relevante em determinado documento e para o que, salientamos a necessidade do conhecimento de estrutura textual, por parte do indexador, como um fator a ser considerado nas propostas de novas metodologias para a análise de conteúdo.

7.9 ASPECTOS LÓGICOS

No momento da leitura do documento para indexação, o indexador realiza esforços mentais classificados dentro de uma das divisões da “Ciência Normativa Lógica”, e por sua vez pertence a uma ciência maior a Filosofia. A divisão da lógica, conforme Santanella (apud SILVA; FUJITA, 2004, p 152), pode ser entendida como:

- a) lógica Utens que é direcionada para o raciocínio comum dos seres humanos, visando formar opiniões, mas que, porém não é suficiente por si mesmo;
- b) lógica Docens que exige do ser humano, estudos do processo de raciocínio e a investigação de métodos que dêem bases mais confiáveis ao pensar, e apressem o avanço do conhecimento para os resultados desejados.

7.10 ASPECTOS COGNITIVOS

O indexador, além dos processos lingüísticos, utiliza, também, processos cognitivos que interagem na leitura. Os processos cognitivos envolvidos na lei documental estabelecidos por Monday (apud SILVA; FUJITA, 2004, p. 153), são :

- a) representação do conhecimento;
- b) esquemas;
- c) unidades cognitivas;
- d) organização do conhecimento;
- e) compreensão e estruturas do texto;
- f) estruturas semânticas e esquemáticas do texto.

Naves (apud SILVA; FUJITA, 2004, p. 153) identifica o processo de inferência como um dos muitos processos cognitivos, caracterizando em:

- a) inferência lógica (usada para estabelecer causas, motivações);
- b) inferência integrativa (executada no momento da compreensão, e baseada nos conceitos e propriedades da organização hierárquica);
- c) inferência construtiva (baseada no conhecimento do indexador).

Bräsher (2002) afirma que os avanços tecnológicos influenciam a área de informação e conduzem ao surgimento de novas técnicas de representação e recuperação de conteúdo.

Os sistemas de recuperação exigem, para isso, modelos de representação do conhecimento que possibilitem contextualizar os significados expressos nos textos armazenados. As técnicas de indexação podem ser usadas de um lado para organizar os conceitos em instrumentos de recuperação da informação, e também por analogia para analisar e organizar as perguntas em conceitos representados como descritores ou combinações de descritores, símbolos de classificação, etc.

O tipo de documento a ser indexado pode também afetar o produto. As duas características de um índice provavelmente *mais* afetadas por estes parâmetros são: exaustividade e especificidade.

Entende-se por exaustividade o ato de identificar todos os conceitos de um documento que teriam um valor potencial para os usuários de um sistema de informação. Para a exaustividade na indexação, o indexador deverá procurar todos os conceitos de um documento que possam ter um valor potencial para os usuários de um sistema de informação.

Por especificidade entende-se que quando em certos casos o indexador identifica que o nível da indexação pode ser afetado pela importância dada a um conceito pelo autor. Se o indexador considera que uma idéia não foi completamente desenvolvida, ou foi tratada de forma superficial pelo autor, justifica-se a indexação num nível mais geral.

7.11 PROCEDIMENTOS DE INDEXAÇÃO

O controle de qualidade será alcançado mais eficientemente se os indexadores tiverem contato com os usuários, eles podem juntos, por exemplo, determinar se alguns descritores produzem falsas recuperações ou criam problemas

no estágio de saída. A qualidade da indexação depende também de certas propriedades dos métodos e procedimentos de indexação. A qualidade de indexação pode também ser testada pela análise dos resultados de recuperação, por exemplo, pelo cálculo das taxas de revocação e precisão.

Do ponto de vista da NBR 12676 (1992, p. 2):

A indexação corresponde ao ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos. O processo de indexação, segundo a norma, possui três estágios:

- a) exame do documento e estabelecimento do assunto de seu conteúdo;
- b) identificação de conceitos presentes no assunto;
- c) tradução desses conceitos em termos de uma dada linguagem de indexação.

No exame a norma indica a verificação das seguintes partes do documento:

- a) título;
 - b) resumo;
 - c) lista de conteúdos;
 - d) introdução,
 - e) conclusão,
 - f) ilustrações;
 - g) diagramas;
 - h) tabelas;
- palavras em destaque.

7.12 AMBIGÜIDADE

Bräsher (2002) cita a ambigüidade como um importante fator a ser considerado para se ter uma eficiente indexação. A ambigüidade é entendida como uma expressão da língua (palavra ou frase) que possui vários significados distintos, podendo, conseqüentemente, ser compreendida de diferentes maneiras por um receptor e causa ruído na recuperação da informação, pois, sob um mesmo termo, o usuário encontrará informação relevante e irrelevante. A ambigüidade pode ser ocasionada por diversos fatores tais como: polissemia, homografia, policategorização, relação contextual e estrutura sintática das frases. E dependendo do fator que a ocasiona pode ser classificada pelos diferentes tipos:

- a) ambigüidade morfológica;
- b) ambigüidade lexical;

- c) ambigüidade sintática;
- d) ambigüidade semântica
- e) ambigüidade pragmática .

Para Bräsher (2002) desambiguação é o processo pelo qual uma ambigüidade é solucionada. Este processo exige diferentes níveis de conhecimentos lingüísticos e extralingüísticos e como os avanços tecnológicos influenciam a área de informação e conduzem ao surgimento de novas técnicas de representação e recuperação de conteúdo.

No contexto tecnológico atual, há tendência para o desenvolvimento de sistemas inteligentes de recuperação de informação com base em processamento de linguagem natural, em função da disponibilidade de textos completos em máquina e da necessidade de interfaces voltadas para o usuário final. Os sistemas de recuperação exigem, para isso, modelos de representação do conhecimento que possibilitem contextualizar os significados expressos nos textos armazenados.

7.13 METADADOS

Alves (apud CASTRO; SANTOS, 2005, p. 15) afirma que os metadados são elementos que compõem estruturas para dar tratamento às informações, tendo em vista o gerenciamento informacional.

Um dos objetivos do uso de metadados no contexto da Web e das bibliotecas digitais é permitir não só descrever documentos eletrônicos e informações em geral, possibilitando sua avaliação de relevância por usuários humanos, mas também permitir agenciar computadores e programas especiais, para que eles compreendam os metadados associados a documentos e possam então recuperá-los, avaliar sua relevância e manipulá-los com mais eficiência.

Diante disso Castro e Santos (2007, p. 16) afirmam que sem o uso de metadados, não haverá o estabelecimento de uma biblioteca digital de acordo com um conjunto de padrões e normas que garantam o acesso, a recuperação, o uso e a reutilização dos recursos informacionais e que atendam às necessidades dos usuários apresentadas nas solicitações de busca destacando como ponto primordial proporcionar meios mais adequados de representar e organizar os conteúdos digitais, tomando como base os aspectos semânticos inerentes a tais conteúdos, no que diz respeito à sua recuperação pelo usuário final.

8 RESULTADOS E DISCUSSÃO DA PESQUISA

Na metodologia utilizada para este trabalho procurou-se estudar o sistema de informação da Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo, para assim evidenciar suas políticas de indexação e observar em que medida elas influenciam na recuperação da informação. Pretendendo assim conseguir subsídios metodológicos para a elaboração de uma proposta de modelo de política de indexação destinada à Biblioteca Digital do CIAR que seja utilizada na uniformização dos procedimentos de indexação com a intenção de se obter excelência na recuperação da informação.

8.1 A UNIVERSIDADE DE SÃO PAULO (USP)

As informações que serão apresentadas sobre a USP têm como base sua página disponível na internet (acesso em 2008). A fundação da USP tem como data o dia: 25 de janeiro de 1934 é uma instituição pública caracterizada como autarquia, sendo mantida pelo governo do estado de São Paulo.

Atualmente a USP que tem como lema a frase em latim *Scientia Vincet*, que em português se traduz: *Vencerás pela ciência*, é formada por 36 unidades de pesquisa e ensino, 24 das quais se localizam em São Paulo, cidade que abriga também a reitoria, um centro de práticas esportivas (CEPEUSP), 9 museus (Estação Ciência, Museu de Anatomia, Museu de Arqueologia e Etnologia, Museu de Arte Contemporânea, Museu de Ciências, Museu de Geociências, Museu Paulista, Museu Oceanográfico e Museu de Zoologia), 2 hospitais Hospital Universitário (HU) e Hospital de Reabilitação de Anomalias Crânio-Faciais (HRACF), o Centro Universitário "Maria Antônia" e diversos órgãos especializados da universidade.

Além disso, se vinculam ou a ela se subordinam para fins de ensino, pesquisa e extensão, diversos outros órgãos públicos do estado. Recentemente, a USP expandiu-se para um novo local em São Paulo, dando origem à EACH (a Cidade Universitária localiza-se na Zona Oeste do município), que iniciou suas atividades de graduação e extensão em 2002.

Universidade de São Paulo é formada por escolas, unidades e institutos, que abrangem todas as áreas do conhecimento divididas em sete campi e uma base avançada:

- a) campus de São Paulo
- b) campus de Bauru

- c) campus de Lorena
- d) campus de Piracicaba
- e) campus de Pirassununga
- f) campus de Ribeirão Preto
- g) campus de São Carlos
- h) centro de biologia marinha (base avançada)

8.2 SISTEMA DE BIBLIOTECAS

A Universidade de São Paulo possui 42 bibliotecas instaladas nas unidades de ensino dos diversos campi. Essas bibliotecas são geridas pelo Sistema Integrado de Bibliotecas (SIBi), que também é responsável pelo catálogo on-line das bibliotecas da USP (Dedalus).

O Dedalus é composto por um catálogo global, que possibilita a consulta simultânea em todas as bibliotecas da Universidade ou pelo catálogo específico de cada biblioteca.

Assim, o usuário é capaz de localizar a referência bibliográfica de qualquer item do acervo da USP. Alguns registros do Dedalus já apontam para Programas de Revistas Eletrônicas, com possibilidade de acesso on-line ao texto completo dos artigos.

O usuário pode pesquisar o conteúdo desejado a partir de palavras-chave, além de consultar a relação dos materiais disponíveis por tipo (tese de doutorado, dissertação de mestrado ou tese de livre-docência), área do conhecimento (humanas, exatas e biológicas) ou pela unidade responsável pelo trabalho.

A Biblioteca Digital está associada a Networked Digital Library of Theses and Dissertations (NDLTD), iniciativa global reconhecida pela Unesco, o que lhe garante confiabilidade e abrangência.

9 ANÁLISE DOS DADOS DA BIBLIOTECA DIGITAL DE TESES E DISSERTAÇÕES DA UNIVERSIDADE DE SÃO PAULO

A Biblioteca Digital de Teses e Dissertações (BDTD) foi criada em 2001 para disponibilizar as consultas ou *downloads* dos trabalhos acadêmicos defendidos na USP e para que as comunidades brasileira e mundial possam ter uma versão digital completa dos documentos, teses e dissertações que lhes interessarem de forma gratuita. O compromisso da BDTD da USP como agente para a construção e

mediação desses conjuntos de conteúdos digitais, criando condições para uma fácil recuperação das informações pelos usuários no novo contexto que se estabelece é notados nas imagens da página de acesso a seguir:

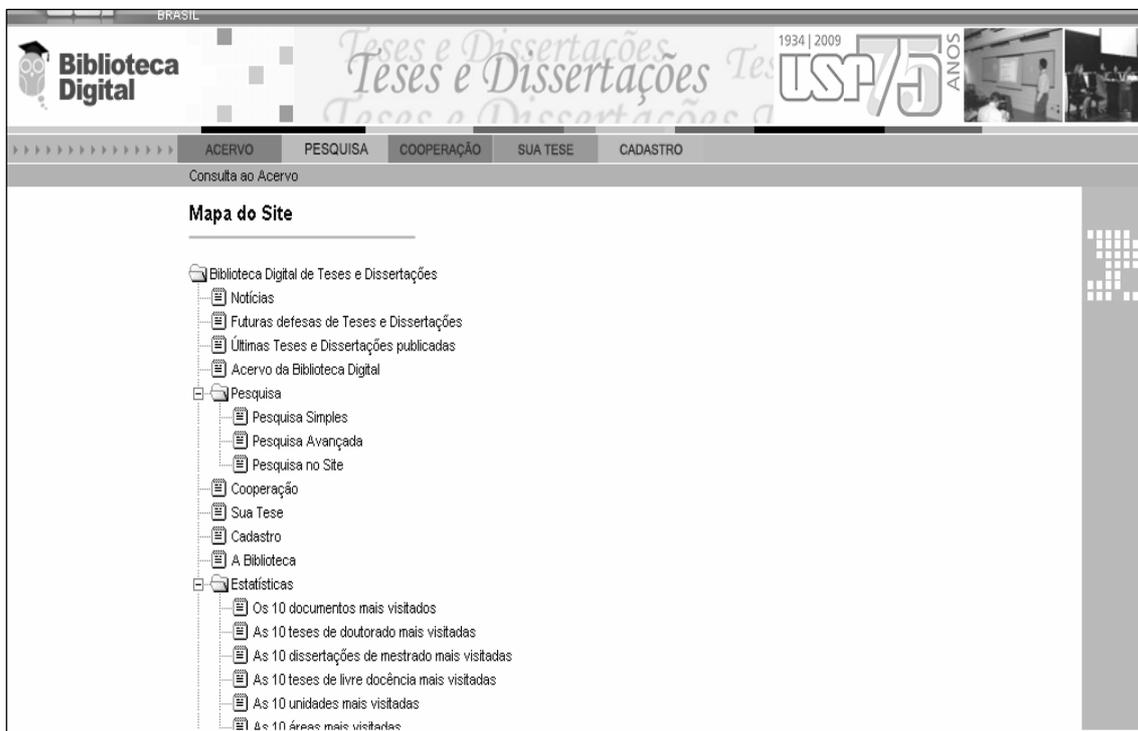


Ilustração 1-Estrutura da página da Biblioteca Digital da USP

Fonte: Fonte: Universidade de São Paulo (acesso em 2008)

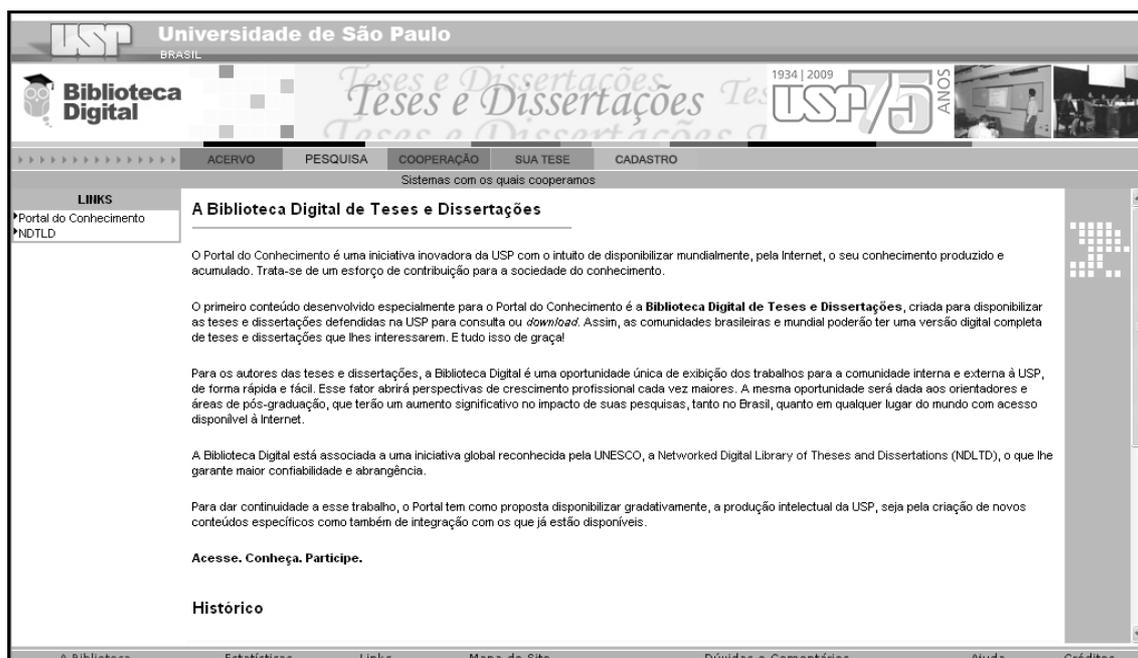


Ilustração 2 - Página de apresentação da Biblioteca Digital da USP

Fonte: Universidade de São Paulo (acesso em 2008)

9.1 OBJETIVOS DA BDTD DA USP

Além de favorecer o acesso a teses e dissertações de forma gratuita para o público em geral, a Biblioteca Digital é uma oportunidade única de exibição dos trabalhos para a comunidade interna e externa à USP, de forma rápida e fácil.

Esta oportunidade favorece não só aos autores, mas aos orientadores e áreas de pós-graduação, que terão um aumento significativo no impacto de suas pesquisas, tanto no Brasil, quanto em qualquer lugar do mundo com acesso disponível à Internet.

A Biblioteca Digital está associada a uma iniciativa global e reconhecida pela UNESCO, a Rede Mundial de Biblioteca Digital de Teses e Dissertações ou *Networked Digital Library of Theses and Dissertations* (NDLTD), o que lhe garante maior confiabilidade e abrangência. O conteúdo da NDLTD pode ser acessado na página: <http://www.ndltd.org>.

O acesso a Biblioteca Digital se dá através do Portal do conhecimento, que tem como proposta disponibilizar gradativamente, a produção intelectual da USP, seja pela criação de novos conteúdos específicos como também de integração com os que já estão disponíveis.

9.2 COOPERAÇÃO

A Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo contou com uma equipe multidisciplinar para o seu desenvolvimento e implantação e assim foram pesquisadas iniciativas semelhantes e referências relacionadas para um embasamento teórico através da Digital Library of Theses and Dissertations (NDLTD), que é o consórcio entre seis (06) instituições que juntas fazem a interface com o usuário, a pesquisa e o desenvolvimento da proposta de Biblioteca Digital de Teses e Dissertações, oferecendo todo suporte necessário para o desenvolvimento e a implantação de uma biblioteca digital.

9.3 SOFTWARE UTILIZADO

Como membro do NDLTD, a USP teve acesso e o direito de utilizar o pacote desenvolvido pela Virginia Tech (Virginia Polytechnic Institute and State University) para o gerenciamento e armazenamento de Teses e Dissertações Eletrônicas (TDE). Este pacote utiliza softwares de domínio público.

O pacote foi adaptado para o contexto da Universidade, integrando-se com o sistema administrativo da Pós-Graduação e com o Banco de Dados Bibliográficos da USP (DEDALUS), mantido pelo Sistema Integrado de Bibliotecas, SIBi). Esta adaptação também envolve uma atualização tecnológica do banco de dados e da linguagem utilizados.

Desde 2003, a USP faz parte da Biblioteca Brasileira Digital de Teses e Dissertações, mantida pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), do Ministério da Ciência e Tecnologia, exportando as informações de suas teses e dissertações.

9.4 ACERVO EXISTENTE

Segundo a página da BDTD/USP, a BD conta com 14656 acervos distribuídos em 8910 dissertações, 5660 teses de doutorado e 86 de livre docência.

As unidades que disponibilizam os acervos são: unidades de ensino e pesquisa; centros e institutos especializados; hospitais e serviços anexos; museus; órgãos externos vinculados e programas conjuntos. As áreas que fazem parte do acervo são: Biológicas; Exatas e Humanas.

9.5 ACESSO AO ACERVO

De acordo com Tammaro e Salarelli (2008, p. 248) o acesso aos documentos compartilhados na Internet poderá ter a obrigatoriedade de receber autorização dos detentores de direitos autorais. Na BDTD da USP o detentor de direito previamente autoriza através do preenchimento de um documento formal o acesso ao seu trabalho.

O acesso ao acervo da Biblioteca Digital de Teses e Dissertações da USP pode ser feito das seguintes formas:

O acesso direto através de menus e sub-menus do *site*, compreendidos como Área de Concentração, Autores e Unidades.

A **Área de Concentração** corresponde a um campo específico do conhecimento e são classificadas em Biológicas, Exatas e Humanas, sendo que as áreas multidisciplinares do conhecimento aparecem em mais de uma área.

No que se refere a **Autores**, estes são ordenados alfabeticamente pelo seu sobrenome.

E, as **Unidades** correspondem às unidades de ensino e de pesquisa da Universidade formadores de mestres e doutores. Estas unidades são classificadas em:

- a) Unidades de Ensino e Pesquisa;
- b) Institutos Especializados;
- c) Institutos Associados;
- d) Órgãos Complementares;
- e) Programas Inter-unidades.

O acesso indireto é feito através de buscas no acervo de **Teses e Dissertações** que permite encontrar documentos que contenham determinadas palavras-chaves que são utilizadas nos trabalhos disponibilizados para consulta, percebendo assim que o uso de vocabulário controlado das áreas atendidas pela BDTD da USP se faz presente no momento da indexação dos assuntos.

9.6 BUSCA

Tammaro e Salarelli (2008, p. 263) explicam que o *online public access catalogue* (OPAC) quando destinado ao acesso remoto do usuário deve observar os seguintes requisitos para ser eficiente:

- a) buscar registros catalográficos relativos aos itens presentes no acervo;
- b) rolagem de listas (autores, títulos, áreas, etc.);
- c) aplicar filtros e outros limites na busca.

A página da BDTD da USP permite vários tipos de busca. No menu superior a busca pode ser realizada clicando o *link* Acervo e Pesquisa.

Na coluna esquerda a página permite a busca através das opções Pesquisa simples e Pesquisa avançada.

Na parte do meio existem várias opções de busca: por **Autores**, clicando na letra do alfabeto referente ao nome; por **Documentos**, clicando nos tipos de documentos que deseja (Dissertação de mestrado, tese de doutorado e tese de livre docência); por **Áreas** (Biológicas, Exatas e Humanas); e por **Unidades** que disponibilizaram o acervo. Como as imagens a seguir exemplificam:

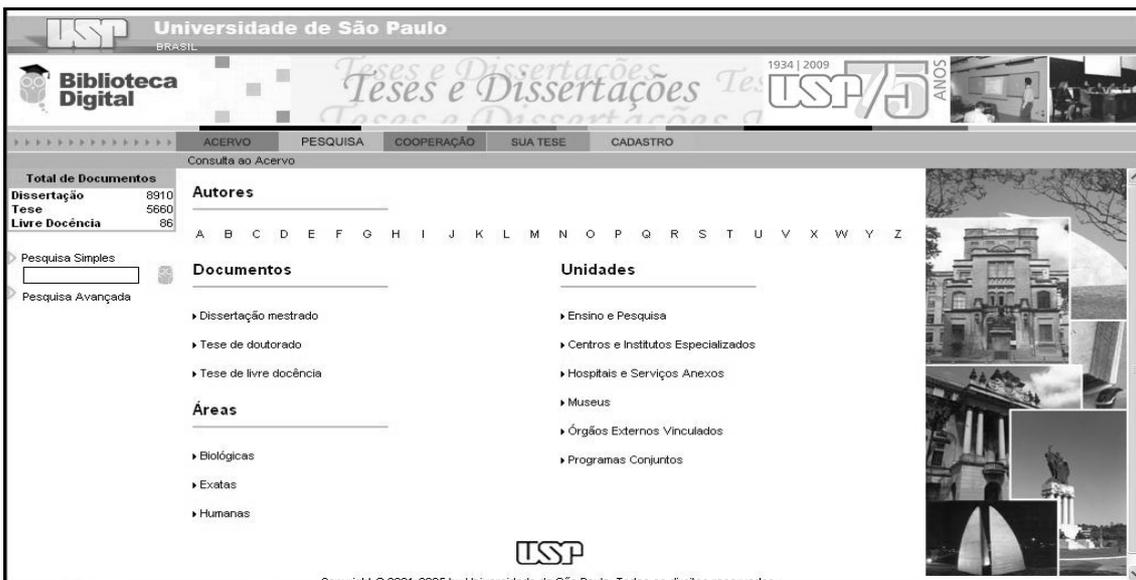


Ilustração 3 - Página de Consulta ao acervo da biblioteca Digital da USP

Fonte: Universidade de São Paulo (acesso em 2008)



Ilustração 4 - Página de Consulta por autor ao acervo da biblioteca Digital da USP

Fonte: Universidade de São Paulo (acesso em 2008)



Ilustração 5 - Página de consulta simples ao acervo da Biblioteca Digital da USP

Fonte: Universidade de São Paulo (acesso em 2008)

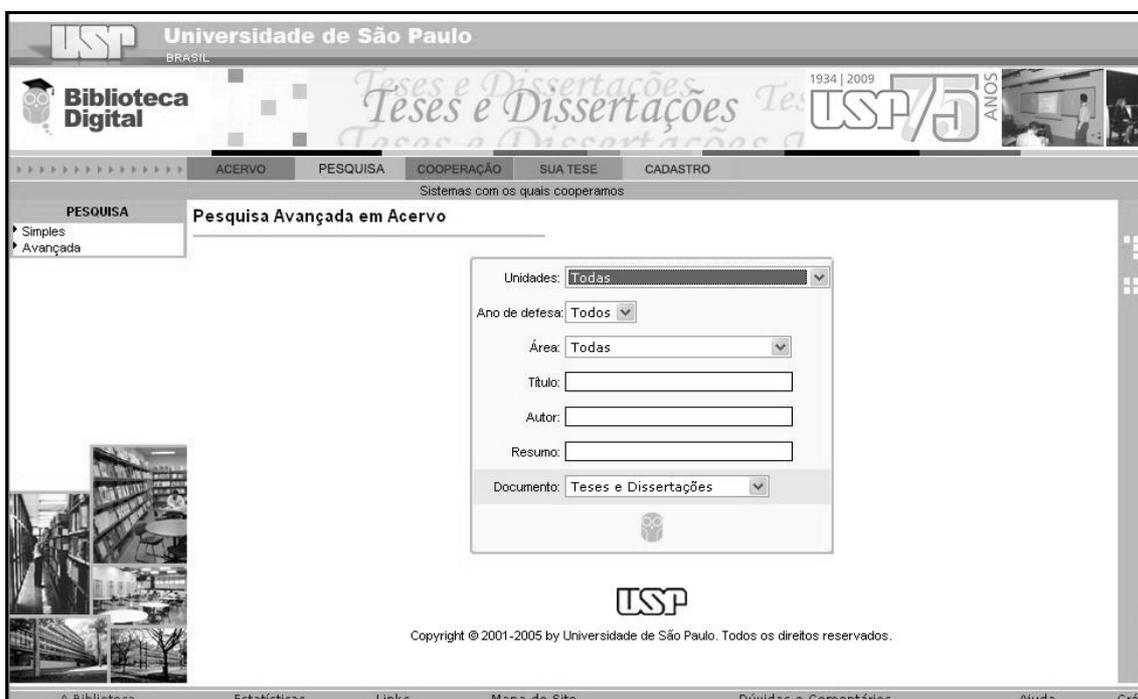


Ilustração 6 - Página de consulta avançada ao acervo da Biblioteca Digital da USP

Fonte: Universidade de São Paulo (acesso em 2008)

Cardoso e Oliveira (2000, p. 3) consideram que um dos objetivos principais de uma biblioteca digital seja satisfazer o usuário com a informação que ele necessita,

sua interface deve possibilitar opções de busca necessárias e adequadas ao contexto de biblioteca digital de maneira a oferecer ao usuário incentivo ao descobrimento de novas formas de se realizar a consulta, recuperação e visualização do documento.

Uma vez que as opções de busca são extremamente necessárias para localizar itens relevantes com maior rapidez e precisão, e que temos pessoas com diferentes perfis inseridos neste universo é preciso trabalhar a questão da interface de usuário com o intuito de suprir tais dificuldades. A BDTD da USP mostra estar ciente desta realidade quando disponibiliza várias formas de pesquisa para seus usuários como demonstra os exemplos nas imagens apresentadas anteriores.

9.7 AUTORIZAÇÃO DE PUBLICAÇÃO

Tamaro e Salarelli (2008, p.193) explicam que no momento da negociação sobre os direitos de propriedade intelectual é preciso chegar a um acordo quanto as licenças respectivas e ou garantias de preservação dos recursos.

Como citado anteriormente a BD da USP disponibiliza as publicações das teses e dissertações através de um termo de autorização, que é preenchida pelo autor. O preenchimento pode ser feito de forma on-line, inserindo dados nos locais determinados do termo. Após o preenchimento o termo deverá ser impresso e encaminhado para a direção da BD. Assim, a BDTD respeita de forma integra a propriedade intelectual dos autores.

The image shows a screenshot of a Mozilla Firefox browser window displaying a PDF document titled "autorizacao.pdf (objeto application/pdf)". The browser's address bar shows the URL "http://www.teses.usp.br/info/autorizacao.pdf". The PDF content is a form titled "AUTORIZAÇÃO" with a sub-header "Como preencher". The form contains the following text and fields:

1 Eu, _____ NOME COMPLETO _____ ?

2 nacionalidade _____, _____ profissão, _____ ?

3 residente e domiciliado em _____ cidade _____ estado _____, na _____ ?

4 endereço completo _____ ?

5 portador do documento de identidade _____ seleçione documento _____ n.º _____ ?

6 órgão emissor _____, inscrito no Cadastro de Pessoas Físicas do Ministério da Fazenda _____ ?

7 sob o n.º _____, na qualidade de titular dos direitos morais e patrimoniais de _____ ?

8 autor que recaem sobre minha _____ seleçione a obra _____, intitulada _____ ?

9 _____ título da obra _____ ?

10 com fundamento nas disposições da Lei n.º 9.610, de 19 de fevereiro de 1998, autorizo a _____ ?

11 Universidade de São Paulo a: _____ ?

Ilustração 7 – Autorização de publicação de teses e dissertações da Digital da USP
Fonte: Universidade de São Paulo (acesso em 2008)

9.8 FORMATO DOS DOCUMENTOS

O formato utilizado na Biblioteca Digital é o mesmo adotado pela Network Digital Library of Theses and Dissertations (NDLTD) para textos das teses e dissertações: o formato digital Portable Document Format (PDF), padrão para difusão em formato eletrônico que preserva as características originais do documento (TAMMARO; SALARELLI, 2008, p. 50). Esse tipo de padrão já é aceito mundialmente para distribuição de documentos.

O site <http://pandora.cisc.usp.br/tde/>, pode ser utilizado para a conversão do documento, conforme pode ser observado abaixo.

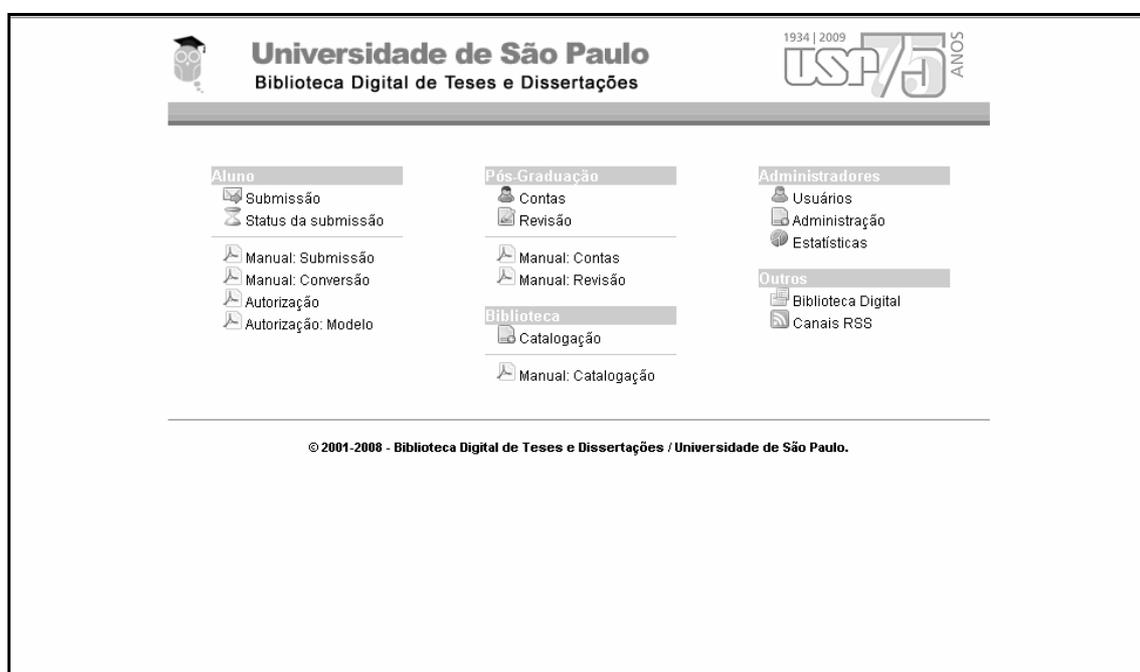


Ilustração 8 – Página do site Pandora na Biblioteca Digital da USP para conversão de documentos
Fonte: Universidade de São Paulo (acesso em 2008)

9.9 RESTRIÇÃO DE USO E PUBLICAÇÃO

A restrição do documento é observada ao acessá-lo. O usuário se depara com a seguinte mensagem:

O autor é o titular dos direitos autorais da tese ou dissertação que você está preste a baixar em seu computador. A tese ou dissertação destina-se para uso pessoal ou científico. Está proibida a comercialização de qualquer espécie sem autorização prévia do autor. Eu li e concordo com as condições acima.

Após a mensagem, o usuário tem a opção de marcar se concorda ou não com as condições. Se não concorda não tem direito ao acesso da publicação, se clica na opção concordo o usuário terá acesso ao documento desejado. Ao clicar a opção concordando com as condições, o documento poderá ser acessado através de um *download*, podendo salvar o arquivo em formato PDF e o conteúdo do documento não poderá ser copiado em outro formato.

As imagens a seguir demonstram como funciona a restrição e o acesso aos documentos:



Ilustração 9 – Página de acesso para cópias de teses e dissertações da Biblioteca Digital da USP
Fonte: Universidade de São Paulo (acesso em 2008)

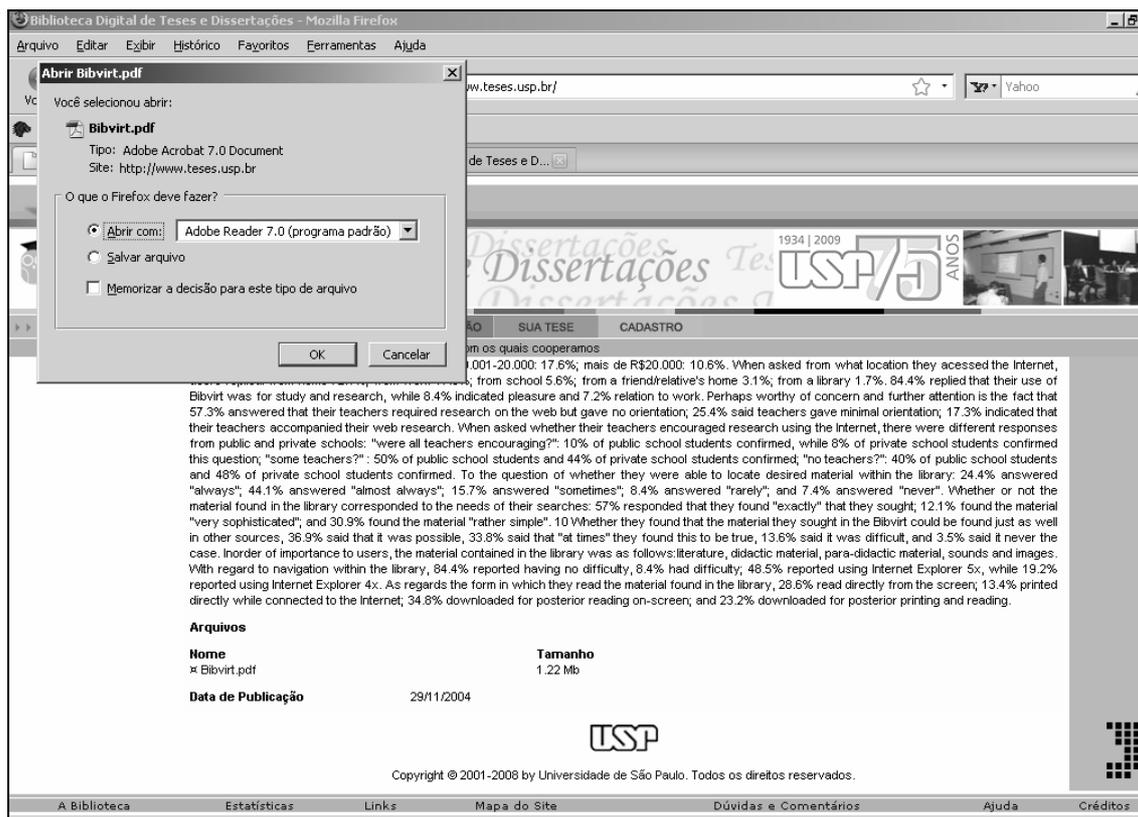


Ilustração 10 – Programa PDF para acesso às teses e dissertações da Biblioteca Digital da USP
 Fonte: Universidade de São Paulo (acesso em 2008)

9.10 ESTRUTURA DE APRESENTAÇÃO DO DOCUMENTO

Para que o trabalho possa fazer parte do acervo digital da BDTD da USP é necessário no momento da indexação o preenchimento da ficha de metadados que são as informações necessárias para administrar a informação em um ambiente de rede (TAMMARO; SALARELLI, 2008, p. 78) como exemplifica o quadro a seguir:

Título original	"A biblioteca virtual do estudante brasileiro da Escola do Futuro da Universidade de São Paulo: um estudo da sua estrutura e dos seus usuários"
Autor	Salgado, Luciana Maria Allan
E-mail	
Unidade	Escola de Comunicações e Artes (ECA)
Área de concentração	Cinema, Rádio e Televisão
Orientador	Litto, Fredric Michael
Banca Examinadora	Litto, Fredric Michael Chaves, Eduardo Oscar de Campos Passarelli, Brasilina
Data da Defesa	15/04/2002
Palavras-chave	▫ bibliotecas digitais ▫ bibliotecas virtuais ▫ comunicação ▫ informação

	☒ Internet
Resumo Original	
Título em Inglês	The Brazilian Student's Virtual Library of The School of the Future of The University of São Paulo: a study of its structure and its users.
Palavras-chave em Inglês	☒ communication ☒ digital librarys ☒ information ☒ Internet ☒ virtual librarys
Resumo em Inglês	
Nome	
Tamanho	1.22 Mb
Data de Publicação	

Quadro1 – Estrutura dos metadados dos documentos digitais da Biblioteca Digital da USP
Fonte: Universidade de São Paulo (acesso em 2008)

9.11 FORMATOS DE ARQUIVOS

Segundo Tamaro e Salarelli (2008, p. 37) o termo arquivo (*file*) pode ser entendido como uma reunião de dados armazenados em discos ou fitas e as publicações eletrônicas/objetos no ambiente digital são identificadas por um conjunto de elementos que os caracterizam e localizam. A normalização desses conjuntos, chamados formatos propiciam as condições necessárias a sua recuperação e intercâmbio via rede. Na Biblioteca Digital de Teses e Dissertações da USP são usados os formatos do quadro a seguir:

Mídia	Formato (extensão)				
Texto	PDF (.pdf)				
Imagem	GIF (.gif)	JPEG (.jpg ou .jpeg)	PNG (.png)	TIFF (.tif ou .tiff)	CGM (.cgm)
	PDF (.pdf)	PhotoCD			
Vídeo	MPEG (.mpg ou .mpeg)	AVI (.avi)	Quicktime (.mov)		
Áudio	MPEG (.mpg ou .mpeg)	WAVE (.wav)	AIF (.aif)	SOUND (.snd)	MIDI (.mid ou .midi)

Quadro 2 – Formatos de arquivos aceitos pela Biblioteca Digital da USP
Fonte: Universidade de São Paulo (acesso em 2008)

9.12 INDEXAÇÃO ADOTADA

Para proporcionar um serviço útil aos usuários de biblioteca digital é necessário fazer uma intermediação para que no momento da busca seja possível interagir com êxito com o sistema informatizado (TAMMARO; SALARELLI, 2008, p. 78). O uso de um vocabulário estruturado permite ao pesquisador recuperar a informação com o termo exato utilizado para descrever o conteúdo daquele documento científico, provenientes de termos consistentes, que permite ao usuário selecionar a informação que necessita.

Na Biblioteca Digital de Teses e Dissertações da USP essa intermediação ocorre com a utilização do Vocabulário Controlado da USP disponível pelo endereço eletrônico: <http://143.107.73.99/Vocab/Sibix652.dll>.

VOCABULÁRIO CONTROLADO DO SIBi/USP

Apresentação Histórico e Metodologia Equipe de Implantação Equipe Atual

MACROESTRUTURA

LISTA DE ASSUNTOS

- Geral Alfabética
- Geral Hierárquica
- Locais Geográficos e Históricos
- Gênero e Forma
- Profissões e Ocupações

TABELA DE QUALIFICADORES

[Pesquisar no DEDALUS](#) [Relatório de Atualização](#) [Contadores](#) [e-m3](#)

© 2006-2008 SIBi/USP - Departamento Técnico - SIBi/DT

Ilustração 11 – Página de Vocabulário Controlado da Biblioteca Digital da USP
Fonte: Universidade de São Paulo (acesso em 2008)

Para que seja feito o processamento dessas informações, é necessário o uso de senhas pessoais dos colaboradores.

9.13 COMUNICAÇÃO ATRAVÉS DE CORREIO ELETRÔNICO

Fujitta (2005, p.12) explica que:

Com a Biblioteca Digital, a Universidade torna acessível, de forma gratuita, sua produção científica e posteriormente algumas coleções do seu acervo, como: obras raras, mapas, coleções especiais, dentre outros. Isto significa otimizar e agilizar a divulgação da pesquisa por ela desenvolvida com acesso em tempo real, extrapolando as barreiras inerentes ao formato em papel: consulta monousuário, se um documento é emprestado, todos os demais interessados ficam impossibilitados da consulta, problema que a disponibilização on-line resolve: acesso multiusuário, simultâneo, não limitado pelo espaço físico, esteja o pesquisador onde estiver, desde que conectado à Internet.

Na Internet, a comunicação pode ser de modo assíncrono ou interativa. Para Fujitta (2005, p. 11) a comunicação assíncrona (correio eletrônico) significa que um usuário pode digitar uma mensagem e enviar sem que haja necessidade do destinatário estar utilizando a rede no momento. É um serviço básico de comunicação em redes de computadores e o processo de troca de mensagens eletrônicas ocorre de forma rápida e fácil, necessitando apenas de um programa de correio eletrônico e do endereço eletrônico dos envolvidos.

A comunicação interativa significa que o usuário pode estabelecer uma conversação, em tempo real, por computador com outro usuário.

Através do correio eletrônico também é possível utilizar outros serviços de rede, tais como: listas de discussão de assuntos pertinentes aos documentos disponíveis em formato digital ou através Really Simple Syndication (RSS), um aplicativo que através de feeds (fontes) envia resumos e atualizações de sites e páginas da internet para os usuários fazendo assim a disseminação seletiva da informação.

Nesta análise, foi detectado na BDTD da USP um link para contato usuário/BD, para acesso a Disseminação Seletiva da Informação (DSI) e RSS, como a imagem a seguir exemplifica.



Ilustração 12 – Cadastro para DSI ou RSS da Biblioteca Digital da USP
Fonte: Universidade de São Paulo (acesso em 2008)

9.14 AVALIAÇÃO DO SISTEMA UTILIZADO

Sampaio et al. (2004, p. 143) explica que:

Os modelos atuais de qualidade estão focados em descobrir quais são as expectativas dos clientes, pois quanto mais se aproximam dessas expectativas, maior qualidade esses serviços contêm. Incorporar qualidade é o primeiro passo para reter o usuário no atual ambiente competitivo.

Foi a partir dessa perspectiva que, em 2002, o Sistema Integrado de Bibliotecas da Universidade de São Paulo (SIBi/USP), visando à estruturação de um programa efetivo de avaliação, realizou um estudo de experiências no uso de modelos para avaliação em bibliotecas e serviços de informação nacionais e internacionais, que serviu de apoio para o delineamento de uma proposta que contemplasse as especificidades da BDTD da USP. Esta proposta veio na forma de um formulário on-line que pode ser preenchido por docentes, alunos e funcionários da USP e recebeu o nome de PAQ – Programa de Avaliação da Qualidade dos Produtos e Serviços.

10 PROPOSTA DE MODELO DE POLÍTICA DE INDEXAÇÃO PARA O CIAR

A Biblioteca Digital do CIAR ainda não trabalha com um modelo de política de indexação. Portanto, como explicado na metodologia, este trabalho tem como

objetivo fazer a sugestão de um modelo de política de indexação no intuito de disponibilizar material informacional de forma eficiente para os usuários.

Para Cunha (1999, p. 262) o armazenamento digital amplia as possibilidades de acesso a um determinado documento e dois pontos distintos devem ser salientados e identificados quando abordamos a biblioteca no contexto do ensino a distância:

- a) os serviços/produtos que já são oferecidos para o usuário presencial devem se adequar para o usuário não presencial;
- b) na recuperação da informação um dos fatores determinantes que irá delinear quais os níveis de representação da informação que serão adotados no acervo será a política de indexação adotada.

Lancaster (1993) acredita que o propósito principal da política de indexação é a elaboração de índices e resumos para constituírem representações de documentos publicados numa forma que se preste a sua inclusão em algum tipo de base de dados.

Strehl (apud SANTOS; AMARAL, p. 74) comenta que dentre os elementos que compõem uma política de indexação podem-se destacar os seguintes:

- a) cobertura de assuntos;
- b) processo de indexação;
- c) estratégia de busca;
- d) tempo de resposta do sistema;
- e) forma de saída;
- f) avaliação do sistema.

Rowley (1994) explica que a partir destas diretrizes estabelecidas pela política de indexação, os sistemas de informação possuem condições de desenvolver as atividades de representação temática dos documentos de forma racional e consistente.

Os termos de indexação a serem atribuídos serão extraídos de uma lista-padrão, baseada na ocorrência de palavras num registro. Os computadores também podem ser convocados para pôr em ordem termos de indexação atribuídos por seres humanos.

10.1 ETAPAS PARA CRIAÇÃO DE POLÍTICA DE INDEXAÇÃO

Para a definição de uma política de indexação para a Biblioteca Digital do CIAR com base na análise da BDTD da USP e no referencial teórico deve-se primeiro observar as seguintes etapas:

- a) criação de um comitê para criação do manual de indexação que deve ser didático e conter o propósito do sistema;
- b) implantação de um banco de dados institucional contendo itens documentais e respectivas informações sobre classificação, leitura documentária, cobertura de assuntos, vocabulário controlado e indexação que serão adotadas, para assim ser possível fazer atualizações tecnológicas quando necessário;
- c) criação de um Tesouro que apresente terminologias estruturadas dos cursos oferecidos pelo CIAR;
- d) definição de modos de busca - autor, palavra-chave, título, assunto, área do conhecimento, editora, ISBN / ISSN, série e DGM (designação geral do material);
- e) descrição dos formatos de arquivo para download disponíveis;
- f) criação de uma ficha para preenchimento de metadados que são as informações necessárias para administrar a informação.

A bibliografia utilizada não apresenta sugestões com relação ao número necessário de descritores que representarão exhaustivamente um documento, pois este número dependerá do sistema de informação e do tipo de material por ele adquirido.

Vale ressaltar, no entanto, que, na indexação dois aspectos são imprescindíveis: a exaustividade e a especificidade, e na recuperação, outros dois fatores devem ser observados, a precisão e a revocação.

No caso da exaustividade, o indexador, no momento da leitura, deverá estar ciente do número de descritores que poderá extrair de cada documento, selecionando aqueles que atendam ao critério da exaustividade.

No caso da especificidade, o indexador no momento da leitura do documento, deve ter em mente o nível de especificidade exigida e o público-alvo.

A qualidade de indexação pode também ser testada pela análise dos resultados de recuperação, por exemplo, pelo cálculo das taxas de revocação e precisão que permitem saber se o usuário está mesmo recebendo as respostas que procura no sistema.

Conclui-se, portanto que os fatores relacionados à indexação (exaustividade e especificidade) estão diretamente ligados à política de indexação adotada pelo sistema e à leitura documentária e que influenciarão no desempenho da indexação realizada pelo sistema de informação, o que refletirá na recuperação da informação.

11 CONSIDERAÇÕES FINAIS

As mudanças tecnológicas também atingem as bibliotecas que colaboram com o ensino a distância, facilitando o acesso às diferentes fontes de informação já que as necessidades do homem e das organizações se modificam constantemente.

Segundo Cunha (1999, p. 257) Os profissionais que atuam em bibliotecas acadêmicas confrontam-se com novas perspectivas de atendimento às necessidades de seus usuários, geradas com o advento das novas tecnologias da informação que foram, incorporadas às atividades das bibliotecas, gerando mudanças internas e transformando os seus produtos e serviços. Nos últimos anos, as mudanças tecnológicas têm aumentado num espaço de tempo cada vez menor. Esse novo acontecimento colocou a biblioteca num período de transição passando de material impresso para forma digital.

Observou-se neste trabalho que a política de indexação em bibliotecas digitais é um dos fatores determinantes para a eficiência dos resultados na recuperação da informação e o indexador é ponto de partida para a geração do conhecimento organizacional sobre política de indexação dentro dos sistemas de informação onde atua.

Torna-se, portanto necessário que o CIAR crie sua política de indexação, pois de acordo com Carneiro (apud RUBI; FUJITTA, p. 68), uma política de indexação:

deve servir como um guia para tomada de decisões, deve levar em conta os seguintes fatores: características e objetivos da organização, determinantes do tipo de serviço a ser oferecido; identificação dos usuários, para atendimento de suas necessidades de informação e recursos humanos, materiais e financeiros, que delimitam o funcionamento de um sistema de recuperação de informações.

A política de indexação e suas ferramentas devem, ainda, ser capazes de tratar determinados fenômenos lingüísticos que afetam a qualidade da recuperação, como a tematicidade e a ambigüidade os quais foram tratados no decorrer deste trabalho.

O projeto não se finaliza aqui, representa apenas uma etapa de um estudo mais amplo, onde a metodologia apresentada poderá ser estendida e aplicada na confecção de um Tesouro que apresente terminologias estruturadas dos cursos oferecidos pelo CIAR, pois dentre as linguagens controladas, o Tesouro constitui um dos meios mais utilizados para indexação e recuperação de documentos e/ou informações.

Para Jesus (2002) a identificação de conceitos relacionada ao processo de representação de uma determinada área do conhecimento na linguagem do Tesouro, garante uma terminologia atualizada para os Sistemas de Recuperação da Informação, o que facilita ao pesquisador no processo de busca e seleção da informação e do documento, uma linguagem padronizada e por ordem de relevância.

Com a atual diversidade de termos técnicos e científicos, bem como, pela demanda por maior rapidez e facilidade na recuperação da informação, se faz sentir uma necessidade crescente no desenvolvimento de sistemas mais avançados e efetivos de organização e gestão de informações, onde o processamento de dados em controle da terminologia é feito por área do conhecimento.

Enfim, entre todas as formas de recuperação da informação o que mais importa é que o usuário identifique os documentos que vão ao encontro de suas necessidades de informação, pois melhorar a qualidade dos serviços tem sido uma preocupação freqüente de bibliotecários e outros profissionais que trabalham, administram ou simplesmente usam esses serviços.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: Métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992. 4 p.

BELLONI, Maria Luiza. Os paradigmas econômicos: contribuição para a educação a distância. In: _____. **Educação a distância**. São Paulo: Autores Associados, 1999.

BLATTMANN, Ursula; BELLI, Mauro José. As bibliotecas na educação a distância: revisão de literatura. **Rev. Online Bibl. Prof. Joel Martins**, Campinas, v. 2, n. 1, p. 23-31, out. 2000.

Disponível em: <<http://www.ced.ufsc.br/~ursula/papers/ciberead.html>>

Acesso: 20 maio 2008.

BRÄSHER, Marisa. A Ambigüidade na Recuperação da Informação. **Ciência da Informação**, Brasília, DF - v.3 n.1 fev/02. 286 p.

Disponível em: <<http://www.scielo.br>>

Acesso: 20 maio 2008.

BRASIL. Ministério da Educação e do Desporto. Portaria nº. 301, de 7 de abril de 1998. A necessidade de normatizar os procedimentos de credenciamento de instituições para a oferta de cursos de graduação e educação profissional tecnológica a distância. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 9 abr. 1998.

CARDOSO, Jiani Cordeiro; OLIVEIRA, João Batista. Problemáticas em interfaces de busca de bibliotecas digitais Atas do V Congresso Ibero-americano de Informática Educativa. 2000, **Anais...** Viña del Mar – Chile.

Disponível em: <<http://ism.dei.uc.pt/ribie/pt/textos/doc.asp?txtid=34#54>>

Acesso: 15 de out. 2008

CASTRO, Fabiano Ferreira de; SANTOS, Plácida Leopoldina V. A. da Costa. Os metadados como instrumentos tecnológicos na padronização e potencialização dos recursos informacionais no âmbito das bibliotecas digitais na era da web semântica. **Informação & Sociedade: Estudos**, João Pessoa v.17, n.2, p.13-19, maio/ago. 2007.

Disponível em: <<http://revista.ibict.br/pbcib/index.php/pbcib/article/view/764#>>

Acesso: 15 de out.2008

CUNHA, Murilo Bastos da. As tecnologias de informação e a integração das bibliotecas brasileiras. **Ciência da Informação**, Brasília, DF, v. 23, nº. 2, p. 182-189, maio/ago. 1994.

Disponível em: <<http://www.scielo.br>>

Acesso: 20 maio 2008.

_____. Desafios na construção de uma biblioteca digital. **Ciência da Informação**, Brasília, DF v. 28, nº. 3, pp. 257-268, set./dez. 1999.

Disponível em: <<http://www.scielo.br>>

Acesso em: 20 maio 2008.

_____. Construindo o futuro: a biblioteca universitária brasileira em 2010. **Ciência da Informação**, Brasília, DF v. 29, n. 1, p. 71-89, jan./abr. 2000.

Disponível em: <<http://www.scielo.br>>

Acesso: 20 maio 2008.

FUJITTA Mariângela S.L. A biblioteca digital no contexto da gestão de bibliotecas universitárias: análise de aspectos conceituais e evolutivos para a organização da informação. **ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO**, 2005 – UFBA.

Disponível em: <<http://dici.ibict.br/archive/00000506/01/MariangelaFujita.pdf>>

Acesso: 20 maio 2008.

GARCEZ, E. M. S.; RADOS, G.J.V. Biblioteca híbrida: Um novo enfoque no suporte a educação a distancia. **REVISTA ACB**. Florianópolis, SC: Associação Catarinense de Bibliotecários, 1996. 6 v.: il. ISBN 1414-0594. v. 6, n. 1, p. 71-96, 2001

Disponível em: <<http://www.scielo.br>>

Acesso: 20 maio 2008.

JESUS, Jerocir Botelho Marques de. Tesouro: um instrumento de representação do conhecimento em sistemas de recuperação da informação. In: XII Seminário Nacional de Bibliotecas Universitárias. 2002. **Anais...** Recife. 18 p.

Disponível em: http://www.ndc.uff.br/textos/jerocir_tesauros.pdf

Acesso em: 05 de nov. de 2008

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Metodologia do trabalho científico**. São Paulo: Atlas, 1992. 270 p.

LANCASTER, F. Wilfrid. Indexação e resumos: teoria e prática. Brasília, DF: Briquet de Lemos, 1993. 347 p.

PEREIRA, Edmeire Cristina; RUTINA, Raquel. O século XXI e o sonho da biblioteca universal: quase seis mil anos de evolução na produção, registro e socialização do conhecimento. **Perspectivas Ciência da Informação**, v. 4, n. 1, jan./jun. 1999.

Disponível em:<<http://www.eci.ufmg.br/pcionline/index.php/pci/article/view/590/359>>

Acesso: 20 maio 2008.

REA, Louis M.; PARKER, Richard A. **Metodologia de pesquisa: do planejamento a execução**. São Paulo. Pioneira Thomson Learning, 2002. 262 p.

ROWLEY, Jennifer E. **Informática para bibliotecas**. Brasília DF.: Briquet de Lemos, 1994. 307p.

RUBI, Milena Polsinelli.; FUJITA Mariângela Spotti Lopes. Elementos de política de indexação em manuais de indexação de sistemas de informação especializados.

Perspect. cienc. inf., Belo Horizonte, v. 8, n. 1, p.66-77, jan./jun. 2003

Disponível :<<http://www.eci.ufmg.br/pcionline/index.php/pci/article/viewFile/375/193>>

Acesso: 20 maio 2008.

SAMPAIO, Maria Imaculada Cardoso; et al. PAQ – Programa de avaliação da qualidade de produtos e serviços de informação: uma experiência no SIBi/USP. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 142-148, jan./abril 2004.
Disponível em: <<http://www.usp.br/>>.
Acesso: 28 de out, de 2008.

SANTOS, Andréa Pereira dos; GOMES, Suely. Ensino a distância: como as bibliotecas do futuro podem contribuir para o aprendizado do Aluno *off campus*. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 13; 2004, **Anais...** Natal, Rio Grande do Norte: BCZM, 2004.
Disponível em: <http://www.bczm.ufrn.br/snbu2004/segundas/trab_livre_oral.html>
Acesso: 20 maio 2008.

SANTOS, Gildenir Carolino; AMARAL, Sérgio Ferreira do Amaral. Rede de conhecimento digital (Bednet): metodologia para a construção da rede de bibliotecas escolares digitais. **Revista Brasileira de Biblioteconomia e Documentação**, Nova Série, São Paulo, v. 2, n. 1, p. 57-82, jan/jun. 2006.
Disponível em: <<http://revista.ibict.br/pbcib/index.php/pbcib/issue/view/2>>
Acesso em: 15 de out. 2008

SILVA, Maria dos Remédios da; FUJITA, Mariângela Spoti. A prática da indexação: análise e evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, v. 16, n. 2, p.133-161, maio/ago.2004.
Disponível em:<<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=65>>
Acesso: 20 maio 2008.

TAMMARO, Anna Maria; Salarelli, Alberto. **A biblioteca digital**. Brasília/DF: Briquet de Lemos, 2008. 378 p.

TEIXEIRA, Cenidalva Miranda de Sousa; SCHIEL, Ulrich. A internet e seu impacto nos processos de recuperação da informação. **Ciência da Informação**, v. 26, n. 1, jan./abr. 1997. 16p.
Disponível em: <<http://www.scielo.br/>>
Acesso: 20 maio 2008.

UNIVERSIDADE FEDERAL DE GOIÁS - UFG. Centro Integrado de Aprendizagem em Rede. [Página do CIAR].
Disponível em: <<http://www.ciar.ufg.br/>>
Acesso: 20 maio 2008.

UNIVERSIDADE FEDERAL DE GOIÁS - UFG. Centro Integrado de Aprendizagem em Rede. **Projeto político Pedagógico**. 2007. 38 p.

UNIVERSIDADE DE SÃO PAULO – USP. **Portal da universidade**.
Disponível em: <<http://www.usp.br/>>.
Acesso: 28 de out, de 2008.

_____. **Biblioteca digital de teses e dissertações**.
Disponível em: <<http://www.teses.usp.br/>>.
Acesso em: 28 de out. de 2008.

____. **Autorização de publicação.**

Disponível em: <<http://www.teses.usp.br/info/autorizacao.pdf>>

Acesso: 28 de out. de 2008.

____. **Autorização de publicação.**

Disponível em: <<http://pandora.cisc.usp.br/tde2/manuais/conversao.pdf>>

Acesso: 28 de out. de 2008.

ANEXOS

ANEXO A – QUESTIONÁRIO:

- 1. Qual sua área de formação? Participou de treinamento específico para fazer a indexação dos materiais disponíveis na biblioteca digital?**
- 2. Como funcionam os processos de recuperação das informações contidas na biblioteca digital? Utiliza senha para consulta somente para alunos e professores ou aberta a todos pelo site da biblioteca?**
- 3. No caso de EAD o professor passa o nome do texto para se fazer a pesquisa ou existem opções de outros textos similares sobre o mesmo assunto?**
- 4. Quem define quais textos, links, etc, deve estar disponível para os usuários da biblioteca digital? (professores, monitores, o currículo do curso que foi definido previamente). Em sua opinião o que deve ser mudado neste processo de definição e por quê?**
- 5. Em caso de perguntas e dúvidas dos usuários da biblioteca digital existe um canal de comunicação ou serviço de referência pelo site para que possam ser respondidas e esclarecidas?**
- 6. Já existem ou está em andamento uma política de indexação para os materiais disponíveis na biblioteca digital? Se sim faça um breve resumo**
- 7. Qual o software utilizado para indexação?**
- 8. Ao escolherem o atual formato se espelharam em outras experiências? Qual (is) e por quê?**
- 9. Existe avaliação do sistema utilizado? Por quem? Como?**
- 10. O acesso aos documentos digitais tem algum tipo de restrição como senha ou obrigatoriedade do usuário ser cadastrado? Os documentos/serviços disponíveis na biblioteca digital correspondem com os objetivos da instituição?**
- 11. Utilizam algum serviço on-line Really Simple Syndication (RSS), um aplicativo que através de feeds (fontes) envia resumos e atualizações de sites e páginas da internet para os usuários para disseminação de informações?**
- 12. Qual o tipo de indexação adotado?**
 - a. () Indexação alfabética de assunto
 - b. () Indexação coordenada
 - c. () Indexação automática

13. Em sua opinião quais dos princípios teóricos e metodológicos da indexação citados abaixo é mais adequado para ser utilizado por uma biblioteca digital?

- a. () Tematicidade
- b. () Aspectos lógicos
- c. () Aspectos cognitivos

14. Qual serviço de referência citado abaixo é utilizado com mais frequência na biblioteca digital?

- a. () Levantamento bibliográfico
- b. () Envio de cópias on-line
- c. () Sumários correntes
- d. () Fichas catalográficas
- e. () Normalização de documentos
- f. () Livros, dicionários, enciclopédias e periódicos on-line
- g. () Lista de Periódicos eletrônicos.

15. Qual o critério definido na leitura documentária realizado na biblioteca digital?

- a. () Instruções textuais
- b. () Condições de produção
- c. () Circunstâncias de enunciação
- d. () Momento e lugar de enunciação

ANEXO B – ROTEIRO PARA ANÁLISE da BDTD da USP:

1. Histórico da USP.
2. Objetivos da BDTD da USP.
3. Existe colaboração com outras instituições?
4. Qual o software utilizado para indexação?
5. Quais áreas do conhecimento fazem parte do acervo?
6. A BDTD da USP pode ser usada somente por alunos e professores ou é aberta a todos pelo site da biblioteca?
7. Como funcionam os processos de busca e recuperação das informações contidas na Biblioteca Digital da USP?
8. Utiliza senha para consulta?
9. Qual política de indexação para os materiais disponíveis na biblioteca digital da USP?
10. Espelharam-se em outras experiências para escolher o atual formato?
11. Como é tratado os direitos de propriedade intelectual?
12. Em caso de perguntas e dúvidas dos usuários da biblioteca digital da USP existe um canal de comunicação ou serviço de referência pelo site para que possam ser respondidas e esclarecidas?
13. Existe avaliação do sistema utilizado? Por quem? Como?