



UNIVERSIDADE
ESTADUAL DE LONDRINA

LIGIA PATRICIA TORINO

**ORGANIZAÇÃO DA PRODUÇÃO CIENTÍFICA
EM REPOSITÓRIOS INSTITUCIONAIS:
UM PARÂMETRO PARA A UTFPR**

LONDRINA
2010

LIGIA PATRICIA TORINO

**ORGANIZAÇÃO DA PRODUÇÃO CIENTÍFICA
EM REPOSITÓRIOS INSTITUCIONAIS:
UM PARÂMETRO PARA A UTFPR**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão da Informação, Mestrado Profissional, da Universidade Estadual de Londrina como requisito parcial à obtenção do título de Mestre.

Orientação: Profa. Dra. Brígida Maria Nogueira Cervantes

LONDRINA
2010

Dados Internacionais de Catalogação na Publicação

T683 Torino, Lígia Patrícia
Organização da produção científica em repositórios institucionais : um parâmetro para a UTFPR/ Lígia Patrícia Torino. – Londrina, 2010.
150 f. : il. ; 30 cm.

Orientação: Brígida Maria Nogueira Cervantes.
Dissertação (Mestrado em Gestão da Informação) – Universidade Estadual de Londrina, Programa de Pós-Graduação em Gestão da Informação (Profissional), 2010.
Bibliografia : f. 135-144

1. Organização da informação. 2. Repositórios institucionais. 3. Produção científica. 4. Gestão da informação. 5. Universidade Tecnológica Federal do Paraná. I. Autor. II. Cervantes, Brígida Maria Nogueira. III. Universidade Estadual de Londrina. Programa de Pós-Graduação em Gestão da Informação. IV. Título.

CDD 025.4

LIGIA PATRICIA TORINO

**ORGANIZAÇÃO DA PRODUÇÃO CIENTÍFICA
EM REPOSITÓRIOS INSTITUCIONAIS:
UM PARÂMETRO PARA A UTFPR**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão da Informação, Mestrado Profissional, da Universidade Estadual de Londrina como requisito parcial à obtenção do título de Mestre.

COMISSÃO EXAMINADORA

Brígida Maria Nogueira Cervantes
Universidade Estadual de Londrina

Luiz Ernesto Merkle
Universidade Tecnológica Federal do Paraná

Terezinha Elisabeth da Silva
Universidade Estadual de Londrina

Londrina, ____ de _____ de 2010.

À minha família

AGRADECIMENTOS

Ao concluir este trabalho gostaria de agradecer a algumas pessoas bem próximas, outras mais distantes, que fizeram toda a diferença e me tornaram melhor a cada dia. Meu muito obrigado a todos os que acreditaram que este sonho seria possível. Especialmente:

A Deus, por ter me abençoado com a possibilidade de alçar mais um grau de conhecimento, me presenteando com saúde e força, pois somente Ele conhece as minhas limitações, meus anseios e angústias.

Ao meu esposo Márcio, por ter me apoiado, acompanhado e ajudado ao longo da minha trajetória profissional e, sobretudo, por ter proporcionado que a minha trajetória pessoal tomasse um rumo diferente, mais leve, mais centrada e acima de tudo, por ter sido um pai de excelência e um companheiro para todas as horas.

Ao meu filho Cauã, pelo colorido que trouxe à minha vida, pelo carinho e alegria que me proporciona diariamente e pela paciência de dividir a mãe com os livros, as pesquisas e as viagens. É por você meu filho que tudo faz sentido!

Aos meus pais Valdecir e Tereza, meus irmãos Leandro e Janyne e minha cunhada Liliane, por compartilharem comigo os cuidados com o Cauã e também a alegria e satisfação de mais esta realização.

Aos meus avós Joaquim e Antonia e a minha tia Jeane, pelo carinho, pelas orações e pelo incentivo que ofereceram em todas as etapas de minha vida acadêmica.

À família de Londrina: Ivanilde, Márcio, Décio, Elda, Vinícius, Gabriel, Fernando, Nani, Letícia, Tereza e Sebastião (*in memoriam*), Valdir, Bia e Elisa pela torcida e apoio, pelas orações, pelas palavras certas nos momentos certos e por estarem sempre presentes nos momentos de alegria e de dificuldade.

Especialmente à Bel, Cida, Toninho e Ana Carolina por terem me acolhido de maneira afetuosa, alegre e carinhosa, oferecendo cama cheirosa e pão quentinho. Vocês são exemplos de dedicação e gratuidade. A vocês, minha eterna gratidão!

À minha orientadora Brígida, pela paciência, disponibilidade, sabedoria, profissionalismo, apoio irrestrito e principalmente pelo relacionamento honesto e verdadeiro que conseguimos construir.

Aos membros da banca Terezinha, Luiz Ernesto e Miguel Luiz Contani pela leitura do texto e por todas as contribuições oferecidas, e a docente suplente Silvana Drumond Monteiro.

Às gestoras dos repositórios em estudo.

À comissão coordenadora do MPGI pelo apoio.

Aos docentes do Programa de Pós-Graduação do Mestrado Profissional pelo compartilhamento de informação e conhecimento.

Aos *teachers*: Denize, Paulo, Leonilde, Gisele e Adriana pela disposição, pelas estratégias adotadas e pelas boas risadas durante as aulas que antecederam a etapa de proficiência em língua estrangeira.

Aos colegas de trabalho: Solange, Andréia, Rafael, Osmar, Vítor, Tiago, Paula, Radames e Marcos Sehaber pelo apoio irrestrito nestes anos de estudo.

Aos bibliotecários do SIBI-UTFPR, especialmente à Anna Caruso, pelo incentivo.

Aos amigos bibliotecários que sempre me prestaram apoio: Bel, Márcia, Di Maria, Lau, Adriano e Mariana.

Aos amigos de viagem Sívio e Ednei.

À Carla Beatriz pela amizade, carinho e capricho nos cuidados com o Cauã;

À minha irmã e colega de turma Emanuelle – Deus nos deu a oportunidade de fazer parte da mesma família, mesma profissão, mesma Instituição e mesmo programa de mestrado... Irmã: quem sabe no futuro sentaremos novamente no banco da UEL e começaremos a mesma conversa... “*Há alguns anos atrás*”....Obrigada pela força!

TORINO, Ligia Patrícia. *Organização da informação em repositórios institucionais: um parâmetro para a UTFPR*. 2010. 150 f. Dissertação (Mestrado Profissional em Gestão da Informação) - Universidade Estadual de Londrina, Londrina, 2010.

RESUMO

Analisa os elementos de organização da informação presentes em repositórios de instituições acadêmicas, com o intuito de subsidiar a implantação desta ferramenta na Universidade Tecnológica Federal do Paraná (UTFPR). Apresenta como objetivo geral propor uma estrutura de organização da informação em repositórios institucionais como um parâmetro para a UTFPR e como objetivos específicos: identificar nos diretórios *OpenDoar* e *Roar* os repositórios de instituições de ensino superior; analisar as estruturas de organização da informação existentes; compará-las e selecionar as estruturas mais adequadas à UTFPR. A metodologia da pesquisa caracteriza-se como um estudo de caso, descritivo, com enfoque na análise qualitativa. A etapa de análise e apresentação dos resultados enfatiza os elementos necessários para auxiliar a elaboração de uma estrutura de organização da informação para repositórios institucionais, contemplando as necessidades que abrangem a elaboração da política de informação, a definição de padrões de metadados e a interoperabilidade humana e semântica. Os repositórios analisados cumprem com os objetivos a que se destinam e apresentam uma estrutura de organização sistematizada levando em consideração tanto a instituição que representam quanto a facilidade de acesso às informações depositadas. Concluiu-se que implementar um repositório vai além da instalação de um *software* e seu “povoamento”, pressupõe compartilhamento e integração das equipes de trabalho e deve ser responsabilidade de toda a instituição, não apenas de um segmento. A possibilidade de aliar questões teóricas à realidade profissional neste estudo foi um fator positivo, sobretudo pela possibilidade de transformar o meio onde a pesquisadora encontra-se inserida, colaborando com a gestão da informação institucional. Partindo dos resultados obtidos, foi possível propor uma estrutura de organização da informação para o repositório da UTFPR.

Palavras-chave: Organização da informação. Repositórios institucionais. Gestão da informação.

TORINO, Ligia Patricia. The Organization of Information at Institutional Repositories: a recommendation for the UTFPR. 2010. 150 f. Dissertation (Master's Degree on Information Resources Management) - State University of Londrina. Londrina, 2010.

ABSTRACT

In this work I analyze some common elements of information organization in educational institutions repositories. With this analysis I intend to promote the organization of a repository at the Federal Technological University of Paraná (UTFPR). My main objective is to propose a structure of information organization in institutional repositories as a parameter for the UTFPR. I have particularly analyzed and compared the respective structure of educational institutions repositories present at the OpenDoar and Roar directories in order to select those that could be more adequate to the future repository of UTFPR. The research methodology is characterized as a descriptive case study, with emphasis on the qualitative analysis. The stage of analysis and presentation of results emphasizes the requirements to help develop a structure of information organization for institutional repositories considering the needs that include the development of the information policy, the definition of meta-data standards, and semantic and human interoperability. The structure of the analyzed repositories comply with their objectives and present a systematized organization structure taking into consideration both the institution they represent and the ease of access to the archived information. It was concluded that to implement a repository goes further than its software installation or the success of capturing the institution's research output; it presupposes the staff's integration and the sharing of information and it has to be the institution's responsibility as a whole. The fact that I work at UTFPR, in particular at the library system, enabled me to positively link theoretical and professional issues. Based on the results, it was possible to propose a structure of information organization for the repository at the UTFPR.

Key words: Information organization. Institutional repositories. Information management.

LISTA DE QUADROS

Quadro 1 - Organização da informação em ambientes tradicionais e digitais	53
Quadro 2 - Elementos de organização da informação segundo Rosenfeld e Morville	55
Quadro 3 - Elementos de metadados MARC	63
Quadro 4 - Elementos de metadados <i>Dublin Core</i>	64
Quadro 5 - Informações gerais sobre os repositórios Dspace UFPR, Lume e RIUnB	97
Quadro 6 - Síntese das estruturas de organização da informação dos repositórios	110
Quadro 7 - Ações para implementação de repositórios institucionais	126

LISTA DE ILUSTRAÇÕES

Figura 1 - O ciclo documentário - visão sistêmica	59
Figura 2 - Atividades do ciclo documentário	60
Figura 3 - Estrutura de organização da informação DSpace UFPR	99
Figura 4 - Uso do mesmo termo na comunidade e coleção DSpace UFPR	100
Figura 5 - Customização da página inicial do DSpace UFPR.....	101
Figura 6 - Estrutura de organização da informação do Lume Repositório Digital	103
Figura 7 - Sub-comunidades do Lume Repositório Digital	103
Figura 8 - Organização das coleções do Lume Repositório Digital	104
Figura 9 - Uso do mesmo termo na comunidade e coleção Lume	105
Figura 10 - Estrutura de organização da informação do RIUnB	108
Figura 11 - Indicação do formato do arquivo disponível no RIUnB.....	118
Figura 12 - Estrutura de organização da informação do repositório da UTFPR	128
Figura 13 - Comunidades e coleções do Repositório da UTFPR	129

LISTA DE SIGLAS E ABREVIATURAS

AACR2 - *Anglo-American Cataloguing Rules*

BDTD - Biblioteca Digital de Teses e Dissertações

BOAI - *Budapest Open Access Initiative*

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CDD - Classificação Decimal de Dewey

CDU - Classificação Decimal Universal

CEFET - Centro Federal de Educação Tecnológica do Paraná

CPD - Centro de Processamento de Dados

IBICT - Instituto Brasileiro de Informação Ciência e Tecnologia

LC - Linguagem controlada

LN - Linguagem natural

MARC - *Machine-Readable Cataloging*

MIT - *Massachusetts Institute of Technology*

NDLTD - *Networked Digital Library of thesis and dissertations*

OAI-PMH - *Open Archive Initiative-Protocol for Metadata Harvesting*

OPEN-DOAR - *Directory of Open Access Repositories*

PROGRAD - Pró-Reitoria de Graduação e Educação Profissional e Técnica

PROPPG - Pró-Reitoria de Pesquisa e Pós-Graduação

ROAR - *Registry of Open Access Repositories*

SABi - Sistema de Automação de Bibliotecas

SEER - Sistema Eletrônico de Editoração de Revistas

SIBI - Sistema de Bibliotecas

UFPR - Universidade Federal do Paraná

UFRGS - Universidade Federal do Rio Grande do Sul

UnB - Universidade de Brasília

UTFPR - Universidade Tecnológica Federal do Paraná

SUMÁRIO

1 INTRODUÇÃO	12
2 UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ: OS FATORES QUE MOTIVARAM A IMPLANTAÇÃO DE UM REPOSITÓRIO INSTITUCIONAL ...	18
3 O MOVIMENTO DE ACESSO ABERTO À INFORMAÇÃO: O CONTEXTO DOS REPOSITÓRIOS INSTITUCIONAIS.....	23
4 REPOSITÓRIOS INSTITUCIONAIS: OS OBJETIVOS, OS CRITÉRIOS DE ORGANIZAÇÃO E A DEFINIÇÃO DE POLÍTICAS PARA SUA IMPLEMENTAÇÃO	30
5 ORGANIZAÇÃO DA INFORMAÇÃO EM REPOSITÓRIOS INSTITUCIONAIS	51
6 PROCEDIMENTOS METODOLÓGICOS.....	80
6.1 INSTRUMENTOS DE COLETA DE DADOS	82
6.2 UNIVERSO DA PESQUISA.....	83
6.3 POPULAÇÃO ALVO.....	85
7 ANÁLISE E APRESENTAÇÃO DOS RESULTADOS	86
7.1 CATEGORIAS DE ANÁLISE.....	86
7.1.1 Informações Gerais Sobre os Repositórios Dspace UFPR, Lume e RIUnB.....	86
7.1.2 Estruturas de Organização da Informação	99
7.1.3 Metadados Utilizados	113
7.1.4 Interoperabilidade Humana e Semântica	119
7.2 PROPOSTA DE ESTRUTURA DE ORGANIZAÇÃO DA INFORMAÇÃO PARA O REPOSITÓRIO DA UTFPR.....	126
CONSIDERAÇÕES FINAIS	132
REFERÊNCIAS.....	135
APÊNDICES	145
APÊNDICE A – Formulário para análise dos elementos de organização da informação em repositórios institucionais	146
APÊNDICE B – Roteiro de entrevista	148

1 INTRODUÇÃO

Este estudo se insere na linha de pesquisa Organização e Representação da Informação e do Conhecimento do Programa de Mestrado Profissional da Universidade Estadual de Londrina e tem como intuito possibilitar que, por meio de subsídios teóricos e práticos, seja possível contribuir para a organização da informação do repositório institucional da Universidade Tecnológica Federal do Paraná.

Sobretudo, porque é possível afirmar que cada vez mais buscam-se ferramentas capazes de armazenar e disseminar o conhecimento registrado de forma permanente e acessível, tendo em vista que produzir conhecimento é tão necessário quanto divulgá-lo. Entretanto, esta tarefa demanda mecanismos que possam assegurar a criação, a disseminação e a apropriação do conhecimento de forma contínua e legitimada pela comunidade.

Sob estes aspectos, Simon (1999) defende que um dos maiores objetivos de uma universidade consiste na realização da pesquisa, oportunizando a criação e descoberta da informação, sua transmissão por meio do ensino e das atividades de extensão e o registro do conhecimento produzido, por meio das publicações que são coletadas em bibliotecas.

Os sistemas de gestão da informação contribuem a estes objetivos e passam a obter lugar de destaque nas mais diversas instituições, especialmente nas instituições acadêmicas, tendo em vista as produções científicas ali produzidas e sua constante ampliação, fator que torna cada vez mais complexo o seu gerenciamento a sua preservação e recuperação.

Entende-se que as unidades de informação possuem a tarefa de subsidiar os processos de gestão do conhecimento registrado, sobretudo porque coube a elas, durante muito tempo a armazenagem e a mediação da circulação do material impresso, servindo de apoio à pesquisa, ao ensino e à extensão. Assim, as mesmas sempre estabeleceram condições de armazenagem e preservação do conhecimento produzido, fator que não se altera com os suportes digitais.

No entanto, as constantes modificações no acesso às informações e também as transformações no processo de comunicação científica, em especial com a aderência ao movimento de acesso aberto à informação, provocaram mudanças

consideráveis especialmente nas relações científicas, ampliando a comunicação entre pesquisadores, aproximando comunidades de interesses comuns e facilitando o compartilhamento das informações. Desta forma, à medida que a disponibilização dos conteúdos em acesso aberto vai sendo ampliada, as bibliotecas e/ou unidades de informação passam a aderir a este movimento e o apoiam principalmente no que tange aos critérios de organização das informações.

Viana, Márdeno Arellano e Shintaku (2005, p. 8) constata que “para as bibliotecas organizacionalmente os repositórios institucionais são uma resposta apropriada ao novo contexto da informação digital”. No entanto, estas ferramentas são bastante complexas, visto que compreendem diversas etapas, como: a formação e desenvolvimento do acervo que irão compor as suas comunidades; os processos de organização destas informações; a interface que permitirá o acesso da comunidade ao documento; a segurança de acesso aos documentos depositados e a preservação dos direitos autorais.

Neste aspecto, a relevância deste estudo envolve as questões teóricas aqui expostas e a prática profissional desta pesquisadora, que é servidora da Universidade Tecnológica Federal do Paraná, de forma a colaborar com os produtos e serviços desenvolvidos pelo Sistema de Bibliotecas (SIBI-UTFPR), setor ao qual se encontra vinculada. Especialmente, no que tange à motivação de ações visando à aprovação do edital FINEP/PCAL/XBDB n.01/2009, o qual possibilitou à instituição receber do Instituto Brasileiro de Informação Ciência e Tecnologia (IBICT), o *kit* tecnológico para implantação de seu repositório institucional¹.

Destaca-se que, embora recente como universidade, a UTFPR possui valiosa riqueza centrada no capital intelectual de sua comunidade acadêmica, que são explicitados por meio de suas produções intelectuais, registradas nos mais diversos formatos e tipologias documentárias.

Dentre eles pode-se destacar a produção de livros, capítulos de livros, folhetos, periódicos, artigos de periódicos, anais de eventos, monografias, teses, dissertações, herbários, patentes, *softwares*, dentre outros, sendo necessário identificar uma ferramenta capaz de ordená-los para que possam ser acessados e utilizados por toda a comunidade interna e externa à instituição.

Partindo deste pressuposto, identificou-se a necessidade de um estudo que pudesse analisar as estruturas de organização da informação existentes

¹ Disponível em: <<http://www.ibict.br/noticia.php?page=0&id=643>>

em repositórios institucionais de universidades brasileiras que viessem a nortear um modelo propício de estruturação da informação condizente às necessidades da Universidade Tecnológica Federal do Paraná.

Este estudo procurou responder a seguinte questão de investigação: Qual a estrutura de organização da informação existente em repositórios de instituições de ensino superior para o tratamento intelectual da informação?

A opção por estudar elementos de organização da informação em repositórios deu-se especialmente porque a UTFPR produz, nos seus diversos programas de ensino, grande quantidade de resultados de pesquisas, que em muitos casos, são pouco disseminados até mesmo na própria comunidade acadêmica, devido à quantidade de *campi* que possui nas mais diversas regiões do estado do Paraná, bem como em decorrência da falta de uma ferramenta de gestão que contemple tais informações.

Cada departamento de ensino ou programa de pós-graduação, na tentativa de reunir suas produções científicas na *web*, cria mecanismos próprios de organização, sem foco na padronização e na interoperabilidade dos dados. Dessa forma, a possibilidade de reunir toda a produção intelectual da instituição em um repositório certamente traria maior visibilidade para suas pesquisas.

No entanto, ao iniciar a estruturação de um repositório institucional, faz-se necessário identificar quais são as tipologias documentárias produzidas na instituição, analisar como ocorre o fluxo destes documentos e ponderar como deverão ser sistematizadas estas informações. Para que todos esses processos estejam em consonância e amparados institucionalmente, há a necessidade da definição de políticas institucionais de informação, que serão as diretrizes que definirão quais os procedimentos serão adotados para que haja a organização, preservação e disseminação do conhecimento produzido no âmbito da instituição.

Supõe-se que experiências de outras instituições de ensino superior que implantaram os seus repositórios institucionais sejam fundamentais para que novas ações possam ser propostas partindo da análise de modelos já existentes. Este estudo buscou analisar essas experiências, a fim de que possam ser apresentadas as diretrizes para a organização do conhecimento da UTFPR, possibilitando que ela tenha melhores condições de gerenciar e divulgar a sua produção e compartilhá-la com outras instituições e/ou pesquisadores.

Desta forma, este estudo foi norteado pelo objetivo geral de propor uma estrutura de organização da informação em repositórios institucionais como um parâmetro para a UTFPR. Para que fosse possível alcançar o objetivo geral, os desdobramentos desta pesquisa foram representados pelos objetivos específicos que consistiram em: identificar nos diretórios *OpenDoar* e *Roar* os repositórios institucionais vinculados a instituições de ensino superior; analisar as estruturas de organização da informação existentes; comparar suas estruturas de organização da informação e selecionar as estruturas de organização aplicáveis ao caso da UTFPR.

Buscou-se, para fins desta análise, pesquisar apenas os repositórios nacionais vinculados a universidades cujo software utilizado seja o DSpace. As razões que basearam este tipo de seleção visaram obter resultados específicos das realidades das universidades brasileiras, instituições com realidade similar a da UTFPR, objeto deste estudo.

Esta pesquisa pode ser definida como descritiva, cujo método escolhido foi o estudo de caso. A abordagem tem enfoque na análise qualitativa e visa retratar os resultados obtidos por meio da análise de suas características. Definiram-se como instrumentos para coleta de dados o formulário, para identificar os elementos de organização da informação presentes nos repositórios que forneceram as diretrizes para a implantação do modelo para a UTFPR e a entrevista que buscou identificar, junto ao comitê gestor dos repositórios, os elementos não explícitos.

Para analisar as estruturas de organização da produção intelectual existentes em repositórios de instituições acadêmicas utilizou-se o formulário como instrumento de coleta de dados, tendo em vista que as informações necessárias para responder às necessidades desta pesquisa estavam disponíveis *on-line*.

Ao observar a interface gráfica (desenho de sua estrutura) e sua forma de organização da informação, procurou-se identificar as categorias dos documentos, as formas de classificação, tratamento e armazenagem dos documentos, quais as linguagens documentárias utilizadas, a questão da interoperabilidade, os esquemas disponíveis, os recursos utilizados para facilitar o acesso à informação e quais os critérios utilizados pela instituição na definição de sua estrutura.

Optou-se por combinar o instrumento formulário à técnica de entrevista a fim de permitir que possíveis dúvidas pudessem ser esclarecidas junto aos gestores dos repositórios selecionados.

Partindo dos resultados obtidos pelos instrumentos de pesquisa, foi possível comparar os critérios de organização utilizados nos repositórios das instituições em estudo, de forma a identificar elementos aplicáveis ao caso UTFPR, levando em consideração a estrutura organizacional *multicampi*, como forma de apoiar a definição de critérios e políticas que norteiem todo o desenvolvimento de um mecanismo de gestão como um repositório.

Analisando todos os aspectos acima mencionados, julgou-se possível, elaborar uma proposta de estrutura de organização da informação para a UTFPR.

A fim de conduzir à exploração do tema em questão fez-se necessário investigar na literatura definições e conceitos para balizar teoricamente este estudo. No que se refere ao referencial teórico, buscou-se compreender e analisar temáticas que pudessem, em conjunto com a análise dos resultados responder ao problema de pesquisa proposto.

Para isso, os capítulos deste estudo encontram-se estruturados da seguinte forma: os capítulos de 1 a 5 constituem o referencial por meio das seguintes abordagens: 1) Introdução: apresenta o tema, o problema, os objetivos e a justificativa acerca da temática da pesquisa; 2) Caracterização da Universidade Tecnológica Federal do Paraná, busca contextualizar este estudo, apresentando os fatores que motivaram a implantação de um repositório institucional; 3) Movimento de acesso aberto à informação apresenta os movimentos que impulsionaram a filosofia de acesso aberto a produção e comunicação científica e o contexto dos repositórios institucionais; 4) Repositórios institucionais aborda a finalidade dos repositórios, os critérios para sua organização bem como a política de informação para a implantação, manutenção e preservação das informações depositadas 5) Elementos de organização da informação em repositórios institucionais: traz uma abordagem sobre a representação descritiva e temática dos documentos, servindo como base para estruturação das informações em um repositório institucional. O capítulo 6 aborda os Procedimentos metodológicos, no qual apresentam-se os instrumentos de coleta de dados, o universo da pesquisa e a população alvo. A Análise e apresentação dos resultados encontram-se no capítulo 7, sistematizado

nas seguintes categorias: 1) informações gerais sobre os repositórios; 2) estruturas de organização da informação; 3) metadados; 4) interoperabilidade. Neste capítulo, apresenta-se também a proposta de estrutura de organização da informação para o repositório da UTFPR.

2 UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ: OS FATORES QUE MOTIVARAM A IMPLANTAÇÃO DE UM REPOSITÓRIO INSTITUCIONAL

A Universidade Tecnológica Federal do Paraná (UTFPR) é a primeira universidade tecnológica do país, e está presente no Estado do Paraná em onze municípios: Apucarana, Campo Mourão, Cornélio Procópio, Curitiba (onde encontra-se a Reitoria), Dois Vizinhos, Francisco Beltrão, Londrina, Medianeira, Pato Branco, Ponta Grossa e Toledo, distinguindo-se das demais universidades federais, sobretudo por sua estrutura *multicampi*, que indica características que a diferenciam sob o ponto de vista cultural e de desenvolvimento regional. Ainda, “a gestão administrativa e acadêmica de uma estrutura *multicampi* como a universidade exige uma ação que seja ao mesmo tempo integradora e autônoma” (UNIVERSIDADE TECNOLÓGICA, 2009, p. 56).

A história da Universidade Tecnológica Federal do Paraná tem início com a criação da Escola de Aprendizes e Artífices, datada de 1909, cujo objetivo era fornecer trabalho e estudo às classes menos favorecidas da sociedade, ofertando ensino elementar e aulas de feitura de calçados e vestiário. Em 1937, transformou-se em Liceu Industrial do Paraná, quando passou a ofertar o ensino de primeiro grau nas atividades de alfaiataria, sapataria, marcenaria, pintura decorativa e escultura. Anos depois, em 1942, passa a ser a Escola Técnica de Curitiba, onde ministra o primeiro ciclo (ensino industrial, mestre, artesanal e aprendizagem) e o segundo ciclo (o ensino técnico e pedagógico).

No entanto, foi em 1959, como Escola Técnica Federal do Paraná, quando por meio de reformulação curricular e administrativa, obteve maior autonomia e descentralização. Os primeiros cursos superiores ofertados por esta instituição datam de 1973, com o curso de Engenharia de Operações nas áreas de construção civil e elétrica; somente em 1978, quando passa a Centro Federal de Educação Tecnológica do Paraná (CEFET-PR), começa a atuar com cursos superiores na área tecnológica. As atividades de pós-graduação *stricto sensu* surgem em 1989 com o Curso de Pós-graduação em Informática Industrial (CPGII).

A década de 1990 inicia marcada pela instalação da primeira unidade de ensino descentralizada no interior do Paraná, cuja cidade escolhida foi Medianeira. Em 1993, as unidades de Cornélio Procópio, Pato Branco e Ponta

Grossa são fundadas e, no ano de 1995, a cidade de Campo Mourão. Ainda nesta década, teve início o Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, em nível de Doutorado; bem como se deu início aos cursos superiores de tecnologia em todas as unidades descentralizadas.

No ano 2000, ocorre a implantação da unidade descentralizada em Curitiba, cidade que passa a responder por aproximadamente 50% de todas as atividades de ensino da instituição e em 2003, dá-se início ao primeiro curso de mestrado ofertado por uma unidade descentralizada, Mestrado em Engenharia da Produção, na cidade de Ponta Grossa, percorrendo uma trajetória de ensino de qualidade e buscando a melhoria da formação de seus estudantes.

De acordo com a Lei n. 11.184, de 07 de outubro de 2005, transforma-se na primeira Universidade Tecnológica do país, denominada Universidade Tecnológica Federal do Paraná (UTFPR), momento em que suas unidades descentralizadas passam a condição de *campi* (UNIVERSIDADE..., 2009, p. 3).

Segundo Silva (2006, p.73), a instituição passa a ser regida pelos princípios de enfatizar a formação de recursos humanos capacitados nas áreas de ensino ofertadas, proporcionar a vinculação com a tecnologia, o desenvolvimento cultural associado às atividades de ensino, pesquisa e extensão, a aproximação dos avanços científicos e tecnológicos com o cidadão-trabalhador e a organização descentralizada dos diversos *campi*, proporcionando maior inserção à realidade regional.

No ano de 2006, instala-se novo campus na cidade de Dois Vizinhos, em 2007 em Apucarana, Londrina e Toledo e em 2008 na cidade de Francisco Beltrão. Presente em onze municípios do estado do Paraná, no ano de 2009, a UTFPR comemora seu centenário.

A estrutura de uma universidade *multicampi* atenta às especificidades regionais deve ter condições de articular o saber científico à realidade na qual se insere. Schallenberger (2006, p.93) indica que a instalação de uma universidade tecnológica *multicampi*, no estado do Paraná, deve agregar à instituição “[...] contornos sócio-espaciais, refletidos pela incorporação do contexto econômico, político, cultural, histórico do seu universo de abrangência, nas funções que exerce”.

Os dados estatísticos do primeiro semestre de 2010, fornecidos pela Assessoria de Estatística e Pesquisa Institucional, vinculada à Pró-Reitoria de Planejamento e Administração da UTFPR, indicam que sua comunidade universitária é composta por 1.761 docentes, 840 técnicos-administrativos e 18.553 discentes (incluindo ensino médio, técnico integrado, pós-médio, cursos superiores de tecnologia e engenharias, como também cursos de pós-graduação: especialização, mestrado e doutorado).

A instituição possui onze programas de pós-graduação (Mestrado e Doutorado) ofertados nos *campi* Curitiba, Pato Branco, Ponta Grossa e Cornélio Procópio e conta atualmente com 624 alunos matriculados, sendo 542 nos programas de mestrado e 82 nos de doutorado.

No que diz respeito às suas unidades de informação (bibliotecas), pode-se mencionar que, assim como ocorria com as unidades de ensino, as bibliotecas atuavam de forma descentralizada. Entretanto, com as mudanças proporcionadas pelo avanço tecnológico, que possibilitou a cooperação de dados, e a necessidade de ampliar seus serviços, bem como a aquisição de um sistema gerenciador da informação, o *software* Pergamum, elas passaram a atuar de forma cooperativa, inicialmente, apenas nos aspectos relacionados ao tratamento intelectual da informação e posteriormente nos aspectos relacionados às políticas de informação institucional.

Embora ainda não formalizado institucionalmente, desde o ano de 2008, as bibliotecas passaram a denominar-se SIBI – Sistema de Bibliotecas da UTFPR, cuja coordenação encontra-se na biblioteca central “Prof. Rosário Farâni Mansur Guérios”, localizada na cidade de Curitiba.

Este sistema de bibliotecas é composto por onze departamentos de bibliotecas que possuem aproximadamente vinte bibliotecários atuando em todos os *campi* da UTFPR. Os departamentos de biblioteca estão subordinados à Secretaria de Assuntos Acadêmicos e à Direção de Graduação e Educação Profissional de cada campus.

Com a missão de promover o acesso, a recuperação e a transferência da informação atualizada para toda a sua comunidade, contribuindo para a formação profissional do cidadão, colaborando para o desenvolvimento científico, tecnológico e cultural do país, o SIBI UTFPR presta apoio ao ensino,

buscando criar uma infra-estrutura de informação que proporcione o aprimoramento intelectual de seus usuários (CARUSO et al., 2006).

Compete ao SIBI a tarefa de organizar e disponibilizar o conhecimento registrado, sendo que, atualmente utilizam-se como ferramentas para gerenciamento, armazenagem e disseminação dos conteúdos em texto completo, a Biblioteca Digital (BDTD) para Teses e Dissertações e o Sistema Eletrônico de Editoração de Revistas (SEER), para hospedar os periódicos produzidos pela instituição.

Cabe mencionar que, ao iniciar este estudo, realizou-se uma pesquisa institucional a fim de que fosse possível conhecer a estrutura desta universidade em relação às atividades de organização da informação já realizadas, bem como as necessidades existentes.

Esta etapa da pesquisa identificou que a UTFPR não possuía um servidor dedicado especialmente à organização e à disseminação de sua produção intelectual. Ainda, foi detectado que a produção intelectual institucional não possuía fluxo formal de depósito compulsório.

Assim, a partir do ano de 2008, o Repositório Institucional passou a ser o alvo dos trabalhos de uma comissão, inicialmente formada por interessados no assunto e posteriormente nomeada pela instituição, com intuito de definir estratégias de organização e disseminação da informação, especialmente porque o repositório além de atribuir valor aos serviços desenvolvidos pelo SIBI, oportunizará maior visibilidade da produção intelectual da Universidade.

No sentido de contextualizar a forma como a UTFPR disponibiliza suas produções intelectuais em texto completo na *web*, constatou-se que a faz por meio da BDTD ou de páginas das Coordenações de curso ou Programas de Pós-Graduação. Os documentos disponíveis no formato impresso são depositados nas bibliotecas do *campus* respectivo ao programa que motivou sua realização. No entanto, estas ações são isoladas, muitas vezes desenvolvidas sem critérios de seleção e/ou organização e sem um fluxo definido de submissão.

Quanto à determinação sobre depósito obrigatório das produções intelectuais dos docentes afastados para cursar programas de pós-graduação em outras instituições, pode-se mencionar que o fluxo de submissão destes documentos não é rígido, fator que não assegura o depósito obrigatório do conhecimento produzido.

Este fator demonstra que muito embora haja grande dificuldade para o estabelecimento de uma política institucional de informação, sua importância é ampliada, principalmente, à medida que é demonstrada a visibilidade proporcionada pelo acesso aberto às informações, bem como a organização das produções em um ambiente interoperável, gerando impacto tanto para o pesquisador que divulga suas produções, quanto para a instituição que a gerencia.

No contexto da organização da informação, o interesse em abordar sua aplicação em repositórios institucionais prevê o objetivo de subsidiar a estruturação do repositório institucional da UTFPR, devido às necessidades que esta instituição identifica, de reunir o conhecimento produzido.

A proposta deste estudo consiste em detalhar os métodos e instrumentos utilizados a fim de que seja possível identificar a estrutura ideal que servirá como norte para a implantação do repositório nesta instituição. Para contextualizar este tema, inicia-se a abordagem por meio do movimento de acesso aberto como forma de conduzir o texto para o surgimento dos repositórios institucionais.

3 O MOVIMENTO DE ACESSO ABERTO À INFORMAÇÃO: O CONTEXTO DOS REPOSITÓRIOS INSTITUCIONAIS

Não é novidade que os cenários de produção e comunicação do conhecimento passaram por alterações substanciais, principalmente em decorrência da velocidade com que as informações são disponibilizadas à sociedade diariamente. Sobretudo, porque estas modificações assinalam que o modelo tradicional de publicação científica tornou-se demasiadamente lento na comunicação dos resultados e incapaz de divulgar toda a produção existente, deixando de atender, com isso, de forma efetiva, o propósito para o qual fora criado.

Disseminar a produção científica e intelectual obtida em determinado assunto, a fim de oportunizar acesso às novas informações e, conseqüentemente, a geração de conhecimento, é uma das principais ações da comunicação científica. Meadows (1999, p.vii) declara que “a comunicação situa-se no próprio coração da ciência. É para ela tão vital quanto a própria pesquisa”.

Assim, o ato de compartilhar saberes, experiências e informações acerca das pesquisas em andamento, bem como o resultado de investigações concluídas e disponibilizadas aos seus pares para a avaliação e validação, objetivando a construção de um *corpus* de conhecimento, originou a prática da comunicação científica (MACHADO, 2006).

De acordo com Kneller (1980), a principal característica da comunicação científica está na interação entre os pesquisadores, na relação dialógica entre os conhecimentos e os interesses existentes, o que amplia sobremaneira a geração de conhecimento e a capacidade de inovação. Assim, ela permeia todo o processo de produção, por meio de canais formais e informais de comunicação.

Os canais informais são aqueles constituídos pelos contatos pessoais entre os pesquisadores que compartilham informações sobre uma área de interesse comum. São em geral estabelecidos em visitas técnicas e encontros científicos. Os canais formais são aqueles que apresentam a informação, registrada em algum suporte. De acordo com Lara (2006), a comunicação formal pode ser entendida como toda aquela publicada em livros e periódicos, sejam eles *on-line* ou

impressos, com o objetivo de disseminar os resultados de pesquisas de autores, instituições ou grupos de pesquisa.

Na visão de Mueller (2000), os tipos de comunicação produzidos pela ciência variam quanto ao formato, podem ser encontrados: relatórios, artigos, livros, palestras e outros; quanto ao suporte, podem estar em papel, meio eletrônico, em materiais multimídia, etc.; quanto a audiências, podem servir para comunicar entre os pares, entre estudantes, para o público em geral, etc.; e finalmente quanto à função, têm por objetivo informar, observar, registrar autoria e outros.

Pode-se afirmar que por um longo período, os periódicos no formato impresso consistiram, sem dúvida, uma das mais importantes formas de compartilhamento de conhecimento, comunicação de resultados de pesquisas e estabelecimento de propriedade intelectual. Para Barbalho (2005), o periódico desempenha um importante papel na comunicação científica, pois se trata do principal meio de veiculação de novos conhecimentos.

No entanto, dentre os periódicos, apenas os tradicionais detinham autoridade para a validação do conhecimento científico, sobretudo em função da revisão por pares, colocando em discussão a legalidade das publicações eletrônicas (MUELLER, 2006). Contudo, há de se destacar que os moldes de avaliação de conteúdos a serem publicados já estavam fundamentados e a filosofia aberta não demonstrava interesse em alterá-los.

Ainda, inúmeras barreiras tornaram este tipo de publicação menos acessível, sobretudo em decorrência do custo, geralmente alto, de suas assinaturas, distante da realidade financeira de muitas unidades de informação e até mesmo dos pesquisadores, fato esse que culminou na crise dos periódicos.

Mueller (2006, p. 31) afirma que desde a década de 1970 vinha se anunciando uma crise relacionada à publicação periódica, entretanto, foi em meados da década de 1980 que “a aparente estabilidade de que gozava o sistema de comunicação científica mundial foi abalada”, quando ocorreu a chamada crise dos periódicos.

Destaca-se que, muito embora países em desenvolvimento como o Brasil já enfrentassem essa dificuldade, a crise aconteceu, de fato, quando o problema atingiu as universidades norte-americanas.

Surge então um novo modelo, uma nova ordem mundial de desenvolvimento e acesso que pressupõe novas alternativas para a disseminação e

o compartilhamento de informação que oportunizam mudanças no formato das publicações tradicionais por meio do movimento de livre acesso à informação científica, que consiste nos conceitos de *open access* (acesso aberto ou livre) e de *open archives* (arquivos abertos), utilizando-se, na maioria dos casos, de tecnologias *open source* (*software* livre ou de código aberto). Torna-se possível afirmar que tal filosofia certamente trará impactos para a comunicação científica (SILVA; TOMAÉL, 2008).

Segundo Alves (2008), os *open archives* constituem-se numa interface que, por meio de mecanismos técnicos organizados, possibilita a interoperabilidade entre as máquinas que mantêm repositórios de dados, visando disponibilizar ampla e livremente o conhecimento científico já validado, criando assim um modelo inovador de comunicação científica.

Para Costa (2008, p. 219), o conceito de acesso aberto, baseado nas Declarações de Berlim, Bethesda e Budapeste foi consensualmente definido como “o acesso à literatura que é digital, *online*, livre de custos, e livre de restrições desnecessárias de *copyright* e licenças de uso”.

Já Kuramoto (2009), defende este conceito como acesso livre, destacando que se trata de algo “acessível em linha, acessível livre de custos, acessível imediatamente e acessível livre da maior parte das restrições relativas a direitos autorais e licenciamento”. Destaca ainda que, o acesso aberto/livre deverá essencialmente ser aplicado ao caso dos artigos científicos, publicados em periódicos e validado por pares, sendo opcional para comunicações, teses e dissertações, relatórios, *papers* e artigos não revistos. Não se aplica a materiais cujos objetivos sejam de obter receita/lucro, incluindo aqui livros, notícias e outros (KURAMOTO, 2009).

Refletindo acerca das definições dos autores mencionados, optou-se por utilizar o conceito de acesso aberto ao referir-se às questões provenientes da Iniciativa de Arquivos Abertos (OAI), que de acordo com Weitzel e Ferreira (2005), trata-se de “um movimento pioneiro, instituído por um grupo de pesquisadores, que buscaram rever o processo de comunicação científica a partir da integração de soluções tecnológicas e na sua implementação”. Partiram daí modelos de disseminação de documentos, dentre eles as bibliotecas digitais, o sistema de gerenciamento de acervos, a publicação de periódicos científicos e também os repositórios institucionais.

De acordo com Weitzel e Ferreira (2005), os fundamentais procedimentos e as principais ferramentas que apóiam a OAI são o autodepósito, a padronização dos metadados para a descrição das informações, o acesso livre e o protocolo OAI, que possibilita a interoperabilidade entre os arquivos, facilitando a recuperação e o compartilhamento de informações.

Contextualizando esta abordagem, Friedman (2007) afirma que o movimento aberto estabeleceu-se na academia entre pesquisadores e cientistas que, há muito tempo, se reuniam de forma colaborativa (inicialmente em redes privadas e posteriormente pela *web*), com objetivo de debater ideias e conceitos. Contrapondo-se ao modelo capitalista, eles buscavam desenvolver, adaptar e adequar sistemas caracterizados como “abertos”, com o objetivo de obter a aprovação de seus pares e divulgar seu conhecimento de forma irrestrita. Assim, percebe-se que comunidades com interesse comum de busca, desenvolvimento e colaboração constituem a tônica da auto-organização em busca de um modelo ideal de democratização da informação.

Este movimento resultou em iniciativas que consistem em modelos alternativos para a comunicação científica, uma proposta de disseminar conteúdos, ampliando a visibilidade do que é produzido, atendendo à demanda de pesquisa e criando canais diferenciados de acesso à informação, anteriormente privilegiados pelas editoras.

As primeiras iniciativas de acesso aberto surgiram na década de 1990, momento esse em que os primeiros periódicos eletrônicos começam a surgir. Paul Ginsparg, em 1991, implantou um sistema eletrônico (*ArXiv*) que permitiu uma das primeiras iniciativas na filosofia de acesso aberto, no intuito de que pesquisadores da área de física e demais interessados enviassem seus estudos para um repositório central, que pudesse permitir acesso a outros pesquisadores. Neste caso, a maioria dos trabalhos ainda não haviam sido avaliados “[...] os autores enviavam seus *préprints* para o Laboratório Nacional de *Los Alamos* ao mesmo tempo em que submetiam às editoras” (MUELLER, 2006, p. 31).

Foi a Convenção de Santa Fé, porém, realizada no Novo México em 1999, que estabeleceu os princípios básicos da nova filosofia para a publicação científica. Os mais importantes princípios estabelecidos foram: o autoarquivamento, que consiste no depósito da produção científica pelo próprio autor ou seu representante; a definição de um número mínimo de metadados e sua integração

com mecanismos de comunicação já existentes no meio científico; a transparência das críticas e sugestões provenientes do processo de revisão por pares sobre os documentos depositados nos repositórios (TRISKA; CAFÉ, 2001).

Destaca-se no apoio a este movimento a *Budapest Open Access Initiative* (BOAI), que ocorreu no ano de 2002, cujo objetivo consistia em “acelerar o progresso do esforço internacional de tornar artigos de pesquisa em todos os campos acadêmicos livremente disponíveis na *internet*” e definia duas estratégias baseadas na iniciativa de arquivos abertos e no protocolo de coleta de metadados OAI-PMH (*Open Archive Initiative-Protocol for Metadata Harvesting*), modelo baseado em padrões nacionais e internacionais de interoperabilidade (OPEN SOCIETY INSTITUTE, 2002).

Nesse contexto, surge duas estratégias de ação, ou seja, dois caminhos visando ao acesso aberto: a primeira estratégia, denominada *via dourada*, consiste na publicação de periódicos eletrônicos em ambientes de acesso aberto, e a segunda, conhecida como *via verde*, compreende o autoarquivamento de artigos de periódicos publicados ou aceitos para a publicação em periódicos especializados em repositórios digitais.

A *via dourada* consiste no movimento para que os periódicos publiquem seus artigos em acesso livre, enquanto a *via verde* está atrelada à implantação de repositórios institucionais por instituições de ensino e pesquisa, incitadas a estabelecer uma política com o preceito de que toda a produção científica de seus pesquisadores sejam autoarquivadas depois de validadas pelos pares (KURAMOTO, 2008).

Na visão de Arellano, Caregnato e Ferreira (2005), o protocolo OAI-PMH estimula o acesso aberto, por meio da disseminação da produção científica para acesso global e irrestrito, alinhando-se ao movimento de arquivos abertos e de livre acesso à informação e ao conhecimento em ciências e humanidades.

Posteriormente, no ano de 2003, a *Bethesda Statement on Open Access Publishing* definiu que as publicações de acesso aberto deveriam assegurar aos usuários que os autores e detentores de direito autoral cedam, a todos os utilizadores, direito livre, irrevogável e perpétuo de acesso a seus trabalhos; licença para copiar, usar, distribuir, transmitir e exibir os trabalhos publicamente, bem como elaborar e distribuir trabalhos derivados em qualquer meio digital, para qualquer

propósito responsável, sujeito à atribuição apropriada de autoria, assim como fazer um número pequeno de cópias impressas para uso pessoal.

Dessa forma uma versão completa do trabalho publicado, assim como de todo material suplementar, incluindo uma cópia da permissão citada na primeira condição, em formato eletrônico apropriado, deve ser depositada, imediatamente, após a publicação inicial em, pelo menos, um repositório *on-line* vinculado a uma instituição acadêmica, sociedade científica, agência governamental ou outra instituição bem estabelecida que busque permitir o acesso livre, a distribuição irrestrita, a interoperabilidade e o arquivamento em longo prazo.

A declaração de *Bethesda* apresentou, ainda, princípios visando envolver agências de fomento, cientistas, editores, bibliotecários e leitores na tentativa de acelerar a transição da publicação de acesso aberto.

A *Conference on Open Access to Knowledge in the Sciences and Humanities*, também de 2003, resultou na conhecida Declaração de Berlim, que compreende a Internet como um instrumento funcional para uma base global de conhecimento científico e especifica medidas que formuladores de políticas, instituições de pesquisa, agências de fomento, bibliotecas e museus precisam considerar.

O referido documento declara-se consonante aos movimentos anteriores ao endossar alterações significativas às práticas de publicação científica, contudo, integra-se ao *European Cultural Heritage Online* e agrega ao contexto do acesso aberto o patrimônio cultural, visando implementar um núcleo com infraestrutura permanente para garantir o livre acesso ao patrimônio cultural na Europa (INTERNATIONAL RESEARCH AND CULTURAL HERITAGE ORGANIZATIONS, 2003).

Motivado pela Declaração de Berlim, o Brasil lança, em 2005, o Manifesto Brasileiro em Favor do Acesso Aberto, documento que antecedeu outras iniciativas ocorridas no cenário nacional no mesmo ano: Declaração de Salvador, Declaração de Florianópolis e Carta de São Paulo (IBICT, 2005).

Em 2006, os participantes do Encontro Aberto “Acesso à Informação Científica: aspectos políticos, tecnológicos e diferenças disciplinares”, por meio de seus representantes, elaboraram a Carta aberta à Sociedade Brasileira para o Progresso da Ciência (SBPC), na qual solicitam que órgãos governamentais atuantes na área de ciência e tecnologia, bem como agências de fomento se

integrem no intuito de: sensibilizar os dirigentes das instituições quanto à importância do acesso livre à informação científica; promover as ações preconizadas pelo movimento brasileiro de acesso aberto à informação científica; apoiar a construção de repositórios digitais com vistas ao registro e à disseminação da produção científica das instituições brasileiras de ensino superior e de pesquisa; ampliar as fontes de informação existentes nas bibliotecas das instituições de ensino superior, em apoio ao ensino de graduação e pós-graduação, além da pesquisa.

Desse modo, “o Brasil caminha a passos largos para estabelecer uma sociedade do conhecimento com acesso totalmente livre e gratuito à informação científica” (IBICT, 2009), sobretudo porque, nos últimos anos, os benefícios decorrentes do movimento foram percebidos por autores, instituições, agências de fomento, bibliotecas e pesquisadores que, continuamente, aderem e contribuem para o fortalecimento do movimento, que parece ter alterado definitivamente a forma de comunicar os resultados das pesquisas.

A esse respeito, Costa (2006, p. 49) afirma que “a filosofia aberta tende a se constituir no fundamento inexorável da comunicação científica daqui por diante. Sua adoção tende, no entanto, [...] a se constituir em uma questão de tempo”.

4 REPOSITÓRIOS INSTITUCIONAIS: OS OBJETIVOS, OS CRITÉRIOS DE ORGANIZAÇÃO E A DEFINIÇÃO DE POLÍTICAS PARA SUA IMPLEMENTAÇÃO

Os impactos causados pela adesão ao movimento de acesso aberto à informação científica foram positivos para as unidades de informação, especialmente no que concerne ao acesso e ao uso das fontes de forma livre e gratuita, racionalizando os recursos financeiros e físicos, tão escassos para a maioria delas.

No caso dos periódicos, a adesão à filosofia aberta possibilitou que fosse ampliada a visibilidade, perceptível na quantidade de acessos aos artigos publicados e na demanda de submissão de *preprints* para a avaliação, além de propiciar significativa redução nos custos de produção. A viabilidade, nesse aspecto, torna-se clara pela crescente disponibilização de publicações em acesso aberto.

Os pesquisadores, aqueles que exercem indubitavelmente função primordial, uma vez que produzem, compartilham e consomem informações científicas, num ciclo ininterrupto e inesgotável, são beneficiados duplamente: como leitores pelo acesso livre de custos e da necessidade de deslocamento para a satisfação de suas necessidades de pesquisa e, também, como autores, pela ampla visibilidade proporcionada pela facilidade de disseminação, capaz de influenciar diretamente não só o índice de citação e o fator de impacto, como também seu potencial de difusão junto à sociedade como um todo.

Instituições e agências de fomento são beneficiadas pelos fatores anteriormente apresentados, além de cumprirem o seu papel de disseminar amplamente o resultado das investigações por elas mantidas e financiadas.

O êxito do movimento de acesso aberto à informação científica necessitou da anuência por parte de autores, institutos de pesquisas, universidades e editores, sendo imprescindível, contudo, o aval das agências de fomento, por meio do incentivo às publicações que o adotam, bem como no reconhecimento dessas fontes nas avaliações da produção científica.

Desta forma, o Brasil, gradativamente, se insere nesse movimento, com a adesão aos periódicos científicos eletrônicos e repositórios de acesso aberto, confirmando a mudança no sistema de comunicação científica (MORENO; LEITE; ARELLANO, 2006, p. 93). Por meio do IBICT, o país indica a aceitação e a

necessidade da organização do conhecimento científico em repositórios institucionais, a fim de ampliar e desenvolver a visibilidade da instituição, tanto no país quanto fora dele. Este órgão estimula o desenvolvimento das bibliotecas digitais e repositórios digitais, sendo agregador das Bibliotecas Digitais de Teses e Dissertações e incentiva a criação de repositórios nas instituições de ensino e pesquisa.

Especialmente no que tange à comunicação formal (produção científica), verifica-se que sua disponibilização é um importante indicador de qualidade dos programas de pós-graduação *stricto sensu*. Com isso, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) instituiu a Portaria n.13 de 15/02/2006, que vincula a avaliação dos programas de pós-graduação à disponibilização de arquivos digitais.

Já para a comunicação científica informal, o advento da informação aberta instituiu uma nova era, na qual as informações são multiplicadas e disseminadas de forma rápida em várias plataformas e suportes, dentre os quais se podem citar: as redes sociais, ferramentas *wiki*, *blogs* e *sites* pessoais, utilizados para estabelecer novos meios de compartilhamento de informações científicas na *web*.

Destarte, pode-se afirmar que o conhecimento produzido em uma instituição pode ser caracterizado como valioso instrumento de competitividade organizacional, cabendo a sua gestão adotar ferramentas que permitam seu armazenamento, tratamento e disseminação.

Neste sentido, cabe contextualizar os repositórios digitais como uma ferramenta que emerge do movimento de acesso aberto à informação apresentando um novo modelo de gestão da produção intelectual com o intuito de organizar, preservar e disseminar as informações, disponibilizando-as em texto completo. Assim, pode-se dizer que um repositório “é uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e prover o acesso apropriado” (VIANA; MARDERO-ARELLANO; SHINTAKU, 2005, p. 3).

A escolha por pesquisar a utilização de repositórios institucionais por instituições acadêmicas surge porque, nas universidades, a pesquisa ocorre em função dos cursos de pós-graduação *stricto sensu* que têm como uma das exigências, para fins de avaliação, apresentar à comunidade os resultados de

investigações produzidas por pesquisadores do corpo discente e docente, por meio da divulgação do conhecimento produzido.

Além de ampliar a visibilidade e o fator de impacto, tanto para o programa quanto para o pesquisador, a produtividade científica é um dos critérios de análise para a avaliação e atribuição de conceito aos programas de pós-graduação. Ademais, tornar público o trabalho é fator de extrema importância, visto que os resultados obtidos nas pesquisas pertencem a toda a sociedade sem distinção (TARGINO, 2000).

É possível afirmar que a aderência às mídias digitais possibilitou aos pesquisadores/docentes de instituições de ensino a condição de ampliar o escopo de conhecimento e permitir maior comunicação entre os acadêmicos e os pares. Portanto, publicar os resultados de pesquisas em periódicos eletrônicos passou a ser um dos diferenciais na promoção de um maior diálogo entre as áreas de conhecimento e os pares. Entretanto, alguns periódicos limitam o acesso às informações devido ao alto custo de suas assinaturas.

Algumas alternativas provenientes da filosofia de acesso aberto foram sendo aceitas pela comunidade científica, sendo os repositórios digitais uma delas. Especialmente no que pondera Café e colaboradores (2003), quando mencionam que disponibilizar as produções intelectuais em um repositório livre de custos, como os concebidos pela filosofia dos arquivos abertos, permite que o próprio autor tenha condições de gerenciar, por meio do autoarquivamento suas produções, dispondo-as a provedores de serviços nacionais e internacionais.

Costa (2006) reforça essa afirmação ao expor que toda a discussão do acesso aberto embasa-se em duas questões: a reação dos pesquisadores (autores) diante da característica comercial adotada pelas principais editoras de periódicos científicos e, em decorrência disso, a visão de que as pesquisas publicadas por tais editoras são, em grande parte, financiadas com recursos públicos e, portanto, deveriam estar publicamente disponíveis.

Dessa forma, têm-se os primeiros movimentos que sustentam a iniciativa de que todo resultado de pesquisa financiada com recursos públicos seja depositado em sistemas de livre acesso, o que alterou sobremaneira a comunicação científica, motivando cientistas e pesquisadores a desenvolverem estratégias a fim de promover o acesso, utilizando-se de ferramentas *web* para responder à demanda

existente no que tange à ampliação de acesso e disseminação dos resultados das pesquisas científicas (KURAMOTO, 2008).

De acordo com Café e colaboradores (2003), os repositórios digitais foram inicialmente, caracterizados como repositórios temáticos, possibilitando a organização de variadas tipologias documentárias dependendo do perfil da área de conhecimento, disponibilizando-as na *web*, permitindo a discussão entre os pares, geração de versões de um mesmo documento, autoarquivamento e interoperabilidade entre os repositórios.

Entretanto, em uma instituição, onde cada um de seus departamentos produz intelectualmente em determinada área do conhecimento, faz-se necessário estabelecer mecanismos de gestão capazes de evitar que o alcance a estas publicações limite-se somente àquela área de estudo.

A característica principal dos repositórios denominados institucionais está em agrupar o conhecimento registrado produzido por membros de uma instituição acadêmica em uma estrutura que permita a organização de diversas tipologias documentárias produzidas por áreas multidisciplinares. Segundo Café e colaboradores (2003, p. 4), um repositório institucional “agrega um conjunto avançado de serviços relativos a organização, tratamento, acesso e disseminação do conteúdo digital produzido por uma instituição”.

No que tange à comunidade acadêmica, Lynch (2003) defende que um repositório institucional pode ser definido como um conjunto de serviços que a universidade oferece para os membros da sua comunidade com vistas ao gerenciamento e disseminação do material digital criado pela instituição e pelos seus membros.

A implementação de um repositório visa atender às necessidades de dispor as informações de forma organizada, visando ao armazenamento, preservação e disseminação da produção intelectual de uma instituição no ambiente *web* (COSTA; LEITE, 2006, p. 1).

Estes repositórios estão inseridos no movimento a favor do acesso aberto e devem ser entendidos como um importante mecanismo de armazenamento ordenado de toda a produção intelectual de uma instituição, proporcionando maior visibilidade da produção científica, ou seja, da produção validada pelos pares, permitindo e facilitando o acesso e compartilhamento das informações ali depositadas.

Neste novo modelo de comunicação científica, não há impedimento financeiro, técnico ou legal, para utilização dos documentos, exceto aqueles inerentes ao acesso a *web*. O objetivo maior do autor é a visibilidade de suas produções que é ampliada pela forma de acesso.

Este modelo de gestão possibilitará aos profissionais da informação uma atuação diferenciada, resgatando antigas práticas para atender às novas exigências dos pesquisadores, que estão pautadas não somente na busca à informação, mas também na sua disseminação e visibilidade junto à comunidade científica. Segundo Leite (2009, p. 99), “bibliotecários tornam-se imprescindíveis mediadores entre a informação científica e seus leitores, atendendo às expectativas de quem a produz e de quem a utiliza”.

Ao que Sayão e Marcondes (2008, p. 135) acrescentam “[...] as organizações da área de conhecimento principalmente as bibliotecas, arquivos e museus – têm crescentemente migrado seus estoques de informação para repositórios digitais”. Essa tendência iniciou-se com as bibliotecas digitais, onde as informações eram disponibilizadas inicialmente no formato digital, sem a pretensão inicial de cooperação. A idéia de armazenagem e recuperação era a tônica inicial. Posteriormente, foram surgindo outras necessidades e estabelecidas redes de bibliotecas com temas comuns, acessíveis em um único ambiente. Nos dias atuais, os repositórios surgem na perspectiva de acesso à informação intelectual.

Tendo em vista os inúmeros suportes e tipologias documentárias produzidas em uma instituição, vale destacar que, muito embora o foco dos repositórios institucionais perpassa pela produção científica, amplia-se, na medida em que passa a gerenciar, toda a produção intelectual institucional que é representada pelas produções técnicas, tecnológicas, artísticas e culturais geradas no âmbito da universidade.

De acordo com a Resolução CES nº. 2 do Ministério da Educação (1998, p.32), cabe conceituar a produção intelectual como:

[...] realização sistemática da investigação científica, tecnológica ou humanística, por um certo número de professores, predominantemente doutores, ao longo de um determinado período, e divulgada, principalmente, em veículos reconhecidos pela comunidade da área específica.

Faz-se importante mencionar que os repositórios digitais são categorizados quanto aos seus conteúdos como temáticos (abrange temas específicos) ou institucionais (abrange as produções intelectuais de uma instituição). Quanto à tipologia documentária, podem ser monodocumentais ou multidocumentais; quanto à administração podem ser consorciados, centralizados, descentralizados e governamentais. Quanto aos objetos digitais, podem ser textuais e multimidiáticos; e quanto à natureza da informação, podem ser proveniente de fontes primárias e de agregadores que reúnem metadados de outros repositórios (SILVA; TOMAÉL, 2008, p. 128).

No modelo compartilhado de publicação de teses e dissertações, o IBICT iniciou como o agregador brasileiro que “por intermédio da BDTD (Biblioteca Digital de Teses e Dissertações), [...] expõe os metadados para o agregador internacional, o NDLTD (*Networked Digital Library of Thesis and Dissertations*)” (SILVA; TOMAÉL, 2008, p. 130).

Os repositórios são sistemas de informação de acesso aberto que implementam o conceito de autoarquivamento e interoperabilidade, sendo esse último, fator de extrema relevância para sua estruturação. Segundo Sayão e Marcondes (2008, p.134) “a interoperabilidade tem sido um dos itens mais críticos para quem pensa no desenvolvimento e operação de sistemas de repositórios e de bibliotecas digitais distribuídos funcionando em rede”.

Para as unidades de informação, a necessidade de compartilhar os dados já existentes e cooperar novos dados permeia o conteúdo dos documentos existentes nas bibliotecas tradicionais até os objetos disponíveis nas bibliotecas digitais. Com a aderência ao repositório institucional, este fator pode ser facilitado ao passo que o repositório possa ser um agregador de outros sistemas já existentes na organização.

Visando atender as necessidades de dispor as informações de forma organizada e no mesmo “ambiente”, os repositórios são abordados sob dois aspectos: a) repositórios institucionais – referem-se aqueles voltados ao armazenamento, preservação e disseminação da produção intelectual de uma instituição; b) repositórios temáticos - referem-se à produção intelectual de uma disciplina (COSTA; LEITE, 2006).

Neste sistema, o autor tem a possibilidade de disponibilizar as suas produções intelectuais e também ter acesso a outras pesquisas realizadas,

ampliando sua capacidade de geração de novos conhecimentos. Essa tendência se consolida mundialmente e amplia exponencialmente a criação de repositórios interoperáveis nacionais e estrangeiros, facilitando sobremaneira os processos burocráticos de uma instituição de nível superior (KURAMOTO, 2005).

A fim de identificar algumas iniciativas de outros países, percebe-se que muitas delas revelam-se eficientes no que diz respeito à visibilidade e competitividade proporcionada pela implantação de seu repositório institucional, dentre elas merecem destaque a Universidade do Minho, situada na cidade de Braga em Portugal e Universidade de Southampton, britânica, situada na cidade de Southampton no Reino Unido.

Embora ainda sejam poucos os repositórios institucionais acadêmicos implantados em universidades brasileiras, cabe mencionar que, além de facilitar o acesso à informação, o país passa a ser beneficiado na medida em que a proposta de interoperabilidade de repositórios brasileiros e estrangeiros possa acontecer, minimizando custos e ampliando sobremaneira o acesso aos documentos, favorecendo pesquisadores, instituições e a sociedade.

A implementação de repositórios institucionais está condicionada à escolha de diversas ferramentas e tecnologias computacionais. Com o propósito de demonstrar a estrutura que permeia os repositórios, Vicentini (2006) define como principais componentes à sua implementação: coleção ou conteúdo, equipe multidisciplinar e capacitada para o trabalho, padronização (metadados), tecnologia incluindo *hardware* e *software* (livre ou proprietário), flexibilidade para desenvolvimento e adaptação, facilidade de gerenciamento da coleção digital, linguagem de programação, utilização de protocolos de comunicação para importação e exportação de dados, digitalização, garantia de direito autoral e preservação do documento digital.

Neste caso, definir as ferramentas que serão utilizadas para a implantação de um repositório trata-se de uma tarefa que demanda muito mais planejamento e trabalho conjunto de analistas, bibliotecários e a gestão da instituição do que recursos financeiros, fator que dificultava a implantação de qualquer ferramenta de tratamento e organização da informação até poucos anos atrás. Atualmente, inúmeras alternativas de pacotes de software livre estão disponíveis na *web*.

Quanto aos softwares utilizados para repositórios, podem-se citar os seguintes: iTor, MyCoRe, o Archimene, CDSWare, Fedora (*Flexible Extensive Digital Object and Repository Architecture*), ARNO, Diva, Eprints (desenvolvido pela *Universidade de Southampton-Inglaterra*) e o DSpace (desenvolvido pelo *Massachusetts Institute of Technology-MIT* e pela *Hewlett-Packard – HP*), sendo que estas duas últimas ferramentas são as que possuem maior registro de repositórios implantados.

No Brasil, o IBICT direcionou estudos que possibilitaram a tradução destes dois softwares (DSpace e Eprints) para a língua portuguesa e passou a ofertar pacotes e suporte técnico para a instalação e uso dos mesmos (LEITE, 2009). Ademais, estudos desenvolvidos por Viana, Márdero-Arellano e Shintaku (2005) constataam que o software que melhor representa hoje a estrutura para implantação de repositórios institucionais é o DSpace.

Deste modo, este estudo traz maior enfoque no *software* DSpace, que é atualmente o mais utilizado internacionalmente para a implantação de repositórios. Trata-se de um *software* livre, configurado em ambiente *web*, que por meio de tecnologias de internet e robôs de buscas, permite a interoperabilidade entre diversos outros sistemas de publicações eletrônicas.

De acordo com Marques e Maio (2006, p. 4), o DSpace possui uma estrutura que reflete as estruturas organizacionais de informação, desta forma, sua configuração permite a representação das coordenações de cursos, programas, possibilitando a organização por área do conhecimento, de diversas tipologias documentárias, em um ambiente multidisciplinar, que é o objetivo de um repositório institucional.

Barton e Walker (2003) salientam que o projeto DSpace, além de estudar os repositórios institucionais, buscam identificar os seus controles de acesso, direitos autorais, versões digitais de documentos, aceite por parte da comunidade científica, bem como identificar seu grau de funcionalidade para a publicação.

De acordo com Viana, Márdero-Arellano e Shintaku (2005, p. 14), este sistema possui três camadas, organizadas em três pacotes distintos, sendo elas: Aplicação, Negócio e Armazenamento. Cada uma delas distribui os seguintes componentes: busca, interface *web*, provedor de dados do OAI-PMH, banco de dados, gerenciador de *workflow* de submissão e outros. A fim de possibilitar a

recuperação dos itens depositados, utiliza-se o mecanismo de busca Lucene e para disseminar os conteúdos utiliza-se o protocolo OAI-PMH possibilitando a interoperabilidade dos metadados. Cabe mencionar que O DSpace utiliza-se dos 15 elementos de metadados padrão Dublin Core, que podem ser usados isoladamente ou em associação com 46 qualificadores baseados no *Dublin Core Libraries Working Group Application Profile* (LAP). A utilização dos elementos dublin core e seus qualificadores é opcional, exceto o título, idioma e data de depósito.

Esta ferramenta permite, ainda, a alteração da configuração original do sistema, possibilitando a customização de sua interface gráfica, definição da estrutura dos metadados, bem como, a utilização de uma interface na língua portuguesa ou internacionalizada, conforme a necessidade de cada instituição (VIANA; MARDERO-ARELLANO; SHINTAKU, 2005).

No entanto, Sayão (2006, p. 138) indica que os objetos digitais, bibliotecas e arquivos digitais necessitam ser estruturados levando-se em conta os “conceitos, procedimentos e políticas que tornem tão longevos os seus estoques informacionais quanto eles são demandados pela sociedade de hoje, tendo como perspectiva o amanhã e o depois”.

Dessa forma, além da construção, comunicação e uso, faz-se necessária a preservação deste conhecimento, estes são os fatores imprescindíveis para possibilitar a qualidade de um sistema de informação. As ações e diretrizes que nortearão os fluxos da informação institucional e formas de organização serão determinadas pelo que pode se denominar política de informação, cujas características serão discutidas a seguir.

Editais que objetivam implantar repositórios nas instituições brasileiras são lançados pelo IBICT, permitindo que instituições de ensino recebam gratuitamente *software* e *hardware* para a implantação nas instituições. No entanto, são várias as exigências para que as instituições sejam aptas a receber os equipamentos, dentre elas, pode-se mencionar: a necessidade do aceite do responsável de maior nível hierárquico da instituição a fim de que o projeto não seja inviabilizado; a integração de diversos profissionais com interesse comum no acesso livre à informação visando à sensibilização dos demais envolvidos; bem como, a definição de uma política de informação institucional.

Seja qual for a ferramenta de gestão da informação a ser implementada, a necessidade da definição de políticas de informação institucionais é

algo imprescindível, pois são elas que determinarão os critérios que irão definir as suas estruturas de organização, o processo de gestão e a capacidade de preservação dos conteúdos nele armazenados.

Estas políticas atuam como um importante mecanismo de compartilhamento de informação e acesso ao conhecimento, possibilitando o “processo de transferência da informação, dentro de um contexto social que gere o ciclo de vida da informação” (AUN, 2003, p.68).

Sobretudo porque “abarcam amplitude maior de interesses, tais como: direitos autorais, acesso aberto e livre, privacidade, telecomunicações e conectividades, softwares livres, propriedade intelectual, governança e governo eletrônico” (SILVA; TOMAÉL, 2009, p. 18).

Uma política de informação pode ser entendida como base para uma ação, ou seja, uma forma dinâmica utilizada para estabelecer os mecanismos que garantam maior capacidade de tomada de decisão, por meio do estabelecimento de prioridades. O termo “política”, na visão de Zhaodong e Xaochu (1987, p. 61), tem a função de propor estratégias e ações a fim de atingir um objetivo e, conseqüentemente, um compromisso com este objetivo – sua implementação.

Além disso, Terra (2005, p. 198) declara que a política pode ser vista como um mandato, que protege tanto a instituição quanto os indivíduos que nela trabalham, facilitando o compartilhamento da informação e do conhecimento, reduzindo a insegurança dos indivíduos em aderir ao compartilhamento, fator que exige trabalho educativo e não somente a elaboração de normas a serem adotadas e procedimentos a serem seguidos.

Corroborando dessa afirmação Silva (1991) ao considerar uma política de informação, em termos ideais, aquela que integra, de forma participativa, os indivíduos e os avanços científico-tecnológicos. Para tanto, precisa ser flexível a ponto de atender a todas as demandas dos diferentes setores que compõem a sociedade e por isso deve ser planejada em conjunto, pelo Estado e pela sociedade.

A necessidade da definição de políticas de informação institucionais no Brasil é iminente, sobretudo pela necessidade de participar da sociedade da informação de forma mais positiva. Deve-se levar em consideração que uma política nacional de informação deve atender aos interesses sociais e culturais da nação, porém deve basear-se em experiências positivas vivenciadas por outros países, para que possa acompanhar o movimento da sociedade mundial (AUN, 2003).

Algumas iniciativas foram surgindo para suprir esta demanda, dentre as quais pode-se mencionar o Projeto de Lei 1.120/2007² que tramita na Câmara dos Deputados e prevê, em seu artigo 2º a discussão e o estabelecimento de uma Política Nacional de Informação Científica (ROLLEMBERG, 2007). O trâmite do referido projeto caminha para a aprovação. Na data de 12 de maio de 2010, o novo relator, o deputado Carlos Abicalil, apresentou parecer favorável à constitucionalidade, juridicidade e boa técnica legislativa do Projeto e de suas respectivas emendas, o que indica que certamente sua aplicabilidade movimentará a comunidade científica.

Até que seja definida uma Política Nacional de Informação da qual as Instituições contempladas pelos editais do IBICT derivariam suas Políticas Institucionais, Kuramoto no ano de 2009 propõe um modelo preliminar para a definição de políticas institucionais de informação³. Sugere-se que este documento seja um ponto de partida para cada uma das instituições, que aprovada por suas instâncias superiores, possa garantir a preservação dos documentos e a continuidade do repositório.

Leite (2009, p. 73) acrescenta que uma política de informação “deve refletir as decisões tomadas ao longo do planejamento e implementação do repositório, deve também estar em concordância com as políticas da biblioteca e da própria instituição”.

Considerando o posicionamento das autoras Tomaél e Silva (2007, p. 4-5), as políticas que determinarão a constituição do repositório deverão contemplar os seguintes aspectos: a) responsabilidade pela criação, implementação e manutenção do repositório; b) o conteúdo proposto e implementado; c) os aspectos legais relativos a documentos e licenças de *softwares*; d) padrões; e) diretrizes para preservação digital; f) políticas de acesso e uso; g) sustentabilidade e financiamento.

a) Quanto à responsabilidade pela gestão do repositório

Para o desenvolvimento de um repositório, o comprometimento e a colaboração entre diversos profissionais são imprescindíveis. A criação de um comitê gestor, composto por uma equipe multidisciplinar, é o primeiro passo a fim de definir as ações necessárias para sua implementação e desenvolvimento. Neste

² Disponível em: <http://www.camara.gov.br/sileg/Prop_Detalhe.asp?id=352237>

³ Disponível em: <http://kuramoto.files.wordpress.com/2009/04/pii_acessolivre.doc>

contexto, caberão, ao profissional da informação, as tarefas de estabelecer padrões de metadados para a organização dos registros de acordo com padrões bibliográficos e o incentivo ao depósito das produções intelectuais dos autores da instituição, entre outras tarefas, dentre as quais se pode mencionar a coordenação de implantação dos repositórios em suas instituições. Ademais, “estabelecer um repositório institucional indica que a biblioteca está mudando seu papel de custódia para contribuir ativamente na mudança do modelo de comunicação científica” (VIANA; ARELLANO; SHINTAKU, 2005, p. 8).

Neste contexto, caberá ao bibliotecário uma atuação que possa garantir um padrão de qualidade dos registros depositados, sobretudo daqueles cuja instituição tenha optado pelo autoarquivamento. Assim, tanto a organização quanto a manutenção dos conteúdos de um repositório passarão por processos criteriosos de avaliação e serão disponibilizados no repositório mediante aprovação deste profissional.

A relevância de manter um Comitê Gestor na instituição é relatado por instituições cujos repositórios já se encontram implantados, como o caso do Repositório Lume da Universidade Federal do Rio Grande do Sul, quando citam que a necessidade da criação de um Comitê Gestor está na possibilidade de estabelecer e acompanhar todas as ações e critérios que a instituição julgar viável para a disseminação de seus documentos (HOROWITZ et al., 2008).

A definição dos profissionais que irão compor este comitê deverá levar em consideração variadas especialidades, de forma que seja possível estabelecer uma distribuição de tarefas que contemple cada um destes profissionais nas atividades necessárias às etapas tanto de definição de ferramentas, customização do *software* e implementação quanto na manutenção do repositório após implantado.

Assim, profissionais de tecnologia da informação e da ciência da informação e comunicação são imprescindíveis, no entanto, considera-se também a necessidade de estabelecer representantes das classes docente e discente, a fim de que, como partícipes desta instituição, possam auxiliar no estabelecimento e definição dos documentos que irão ser contemplados, na divulgação e marketing do repositório, dentre outros.

b) Quanto ao conteúdo

Estatísticas de registro de repositórios institucionais têm crescido diariamente e, conseqüentemente, os conteúdos e tipologias documentárias neles depositados. Para que o repositório institucional possa cumprir com os objetivos propostos, cabe à equipe gestora identificar, desde a fase de implantação, quais são as necessidades da comunidade ao qual o repositório encontra-se vinculado. De acordo com Tomaél e Silva (2007, p. 6), “o conteúdo de um repositório institucional reflete ou demonstra o conhecimento, a pesquisa e o interesse de uma organização”.

Portanto, faz-se necessário um estudo das tipologias documentárias produzidas pela instituição a fim de que possam ser definidos quais os documentos farão parte do repositório. Inúmeras são as tipologias documentárias existentes, segundo Bellotto (2002, p. 20), as tipologias documentais são “a ampliação da Diplomática em direção à gênese documental, perseguindo a contextualização nas atribuições, competências, funções e atividades da entidade geradora/acumuladora”.

Neste aspecto, cabe definir qual o tipo de abordagem a instituição irá adotar para o repositório, sendo que ao optar pela abordagem rígida, o repositório deverá priorizar os documentos cuja qualidade já tenha sido validada por pares, como ocorre no caso dos artigos de periódicos. De acordo com Leite (2009, p. 28), esta abordagem indica que os repositórios devem, sobretudo,

[...] potencializar a comunicação científica formal visando especialmente a maximização dos impactos dos resultados das pesquisas por meio da disseminação ampla e irrestrita de artigos de periódicos científicos, principal argumento do movimento de acesso aberto à informação científica.

Para o autor supracitado, deve ser priorizada esta abordagem a fim de que o foco do movimento de acesso aberto, que é o de ampliar o impacto das pesquisas científicas, seja atingido.

No entanto, a abordagem flexível não surge para contrapor a abordagem rígida, mas para ampliar a quantidade de tipologias que um repositório pode contemplar, favorecendo a organização da produção intelectual (seja ela validada por pares ou não), incluindo imagens, sons, vídeos, objetos de aprendizagem dentre outros. Leite (2009, p. 29) argumenta que a abordagem flexível “reforça que o conhecimento científico não é produzido exclusivamente a partir

daquilo que já foi avaliado e é publicado formalmente, mas também daquilo que é veiculado informalmente”.

Definir qual tipo de abordagem a instituição utilizará a fim de constituir o conteúdo que irá integrar o seu repositório é uma das atividades mais relevantes para a etapa de implantação e deverá levar em consideração a necessidade que a instituição possui.

Também se deve levar em consideração a necessidade da definição dos formatos em que os documentos digitais serão depositados (doc, pdf, html, jpeg, gif, mpeg) fator que influenciará sobremaneira a customização da tecnologia de informação escolhida para o repositório (TOMAÉL; SILVA, 2007).

c) Quanto aos aspectos legais

Destaca-se que, muito embora a filosofia aberta apresente pontos extremamente positivos, em algumas instituições o movimento não obteve a aderência desejada, sobretudo em decorrência do questionamento/desconhecimento quanto ao direito autoral.

É necessário salientar que o fato de disponibilizar o documento em repositório institucional não confere a esta ferramenta a apropriação intelectual de seu conteúdo. Os repositórios, assim como as bibliotecas tradicionais, possuem tão somente a custódia do documento com objetivo de disseminação.

Muito embora seja possível afirmar que o conteúdo de um documento disponibilizado em acesso aberto possua maior visibilidade e acesso, a propriedade intelectual da obra continua sendo exclusivamente do autor. Assim, cabe a ele autorizar no momento do arquivamento quais as licenças que permite para acesso àquele documento, eximindo assim a responsabilidade da instituição acerca da divulgação do documento por ele permitido.

Há um mecanismo importante, denominado diretório *Sherpa-Romeo*, que possibilita a identificação das políticas de autoarquivamento estabelecidas por editores de periódicos científicos, servindo também de apoio aos gestores dos repositórios no auxílio aos autores das obras que serão depositadas.

Cabe aqui mencionar que dentre os seus objetivos, o repositório, como sistema de gestão da informação permite a disseminação das informações e também a proteção da propriedade intelectual da obra, possibilitando diferentes tipos de licenças sobre a utilização do material.

Neste contexto, surge o *Creative Commons*, como um novo modelo de gestão dos direitos autorais presente em mais de 40 países. Trata-se de uma iniciativa de Lawrence Lessig, defensor de opções alternativas de proteção dos direitos autorais, desde que não ofereça prejuízo aos usuários da informação produzida por estes.

De acordo com Jorente e Santos (2008, p. 120), o *creative commons* tem como objetivo “a definição de um espectro de possibilidades entre o direito de autor, totalmente proibitivo e o domínio público, ausência de garantias de direito, a manutenção de alguns direitos, enquanto se permitem alguns usos do trabalho criativo”.

Desta forma, busca-se soluções mais adequadas em meio às novas possibilidades de informação conhecidas nos dias de hoje, resguardando a autoria e facilitando o uso comum de trabalhos on-line, filmes, imagens, músicas, desde que marcados por uma licença do *Creative Commons*. Assim, faz-se importante mencionar que, ao escolher uma de suas licenças, o autor não deixa de ter direito sobre sua obra, o crédito ao autor não deixa de existir e, passa a ser definida em quais condições sua produção pode ser utilizada (JORENTE; SANTOS, 2008).

Apresentar a produção intelectual mediante uma licença como esta não significa que os direitos do autor sejam negados, e sim que há possibilidade de oferecer alguns destes direitos para alguns usuários, sob algumas condições. Ainda, o *Creative Commons* permite que suas licenças sejam combinadas e recombinadas fator que permite que cada autor opte pela licença mais adequada às suas necessidades e interesses (LEMOS, 2005).

Dentre as licenças permitidas pode-se citar as seguintes:

Licença de Atribuição: o autor permite a cópia, distribuição e execução de sua obra que é protegida por direitos autorais, bem como a de todas as obras derivadas dela, desde que seja dado o crédito estabelecido.

Uso não comercial : o autor permite a cópia, distribuição e execução de sua obra e as obras derivadas dela, somente para fins não comerciais.

Não a Obras Derivadas: o autor permite que outras pessoas copiem, distribuam e executem somente as cópias exatas de sua obra, porém, não permite estas opções para obras derivadas.

Compartilhamento pela mesma Licença : o autor pode permitir que outras pessoas distribuam obras derivadas da sua, desde que utilize de uma licença

idêntica à licença utilizada por sua obra. Desta forma, este tipo de licença se aplica somente a obras derivadas.

Recombinação (Sampling): Desenvolvida no Brasil, em conjunto com a Escola de Direito da Fundação Getúlio Vargas, com apoio do Ministério da Cultura. Nesta licença, o autor pode ou não autorizar a cópia livre, utilização e distribuição da obra. Neste caso, o autor permite o uso parcial ou recombinação de boa-fé da obra e o uso de técnicas de sampleamento, mesclagem e colagem, ou qualquer outra técnica artística. Esta autorização torna-se válida, desde que o original seja significativamente modificado, ou seja, que possibilite que uma nova obra seja criada. As obras derivadas são autorizadas para o autor que recriou a obra do autor original.

CC-GPL e CC-LGPL: Visam ao licenciamento de software. Como a licença anterior, estas também tiveram origem no Brasil, neste caso, visando incentivar a adesão ao uso do *software* livre no país. Estas licenças prezam os direitos do *software* livre, liberdade para: estudar o programa (acesso ao código fonte); executar o programa para qualquer finalidade; modificá-lo e aperfeiçoá-lo e distribuí-lo livremente.

O autor poderá ter a licença por meio de três formatos: 1) Licença para leigos (*Commons Deed*), trata-se de um resumo da licença de forma simples, completa e com os ícones relevantes, 2) Licença jurídica: trata-se de uma licença detalhada, que trará ao autor validação perante o judiciário e 3) Licença para máquinas: trata-se de uma versão que pode ser legível por máquinas, permitindo que mecanismos de busca identifiquem a obra e seus termos de uso.

Indica-se que ao optar pelo autoarquivamento dos documentos em um repositório, a instituição tenha definido em sua política a necessidade do autor documentar a permissão para disponibilização de seus conteúdos, evitando assim transtornos posteriores.

d) *Padrões*

Este aspecto abrange, além da plataforma tecnológica (escolha de *hardware* e *software*), o *workflow* (fluxo administrativo e de publicação de documentos), a interface e questões de usabilidade, aborda a definição dos metadados e a questão da interoperabilidade (humana e semântica).

Quanto à plataforma tecnológica a ser utilizada, constata-se que a definição de *hardware* e *software*, os quais irão definir a estrutura de um repositório, é uma etapa que deve ser trabalhada de forma simultânea com as demais etapas de implantação do repositório dentre elas, as políticas que o sustentarão. As funcionalidades inerentes a um *software* destinado a repositórios de acordo com Leite (2009, p. 58) devem:

[...] capturar e descrever conteúdos digitais por meio de interface de autoarquivamento, tornar público por meio da internet o acesso a esses conteúdos (ou quando necessário ao menos aos seus metadados), armazenar, organizar e preservar digitalmente conteúdos a longo prazo, compartilhar os metadados com outros sistemas na internet.

Embora seja possível identificar uma grande quantidade de *softwares* disponíveis, deve-se levar em consideração aqueles que atendam aos modelos de acesso aberto à informação, cujas características que os diferenciam dos demais consistem na escolha de *software* livre, com a possibilidade de disponibilização de texto completo, autoarquivamento dos documentos, disponibilidade de estatísticas de acesso e *downloads* dos documentos, embargo, interface de busca e navegação, gerenciamento de comunidades e coleções, interoperabilidade, disponibilidade de diferentes formatos de documentos, definição de padrões de metadados e flexibilidade para defini-los, preservação digital, linguagem de marcação XML, seleção de idioma (LEITE, 2009).

Cabe mencionar que a linguagem XML prioriza a descrição de estruturas semânticas, levando em consideração os significados contextuais, diferente da linguagem HTML, que é uma linguagem de marcação de textos utilizada para o desenvolvimento de páginas *web*. Ramalho, Vidotti e Fujita (2005, p.98) mencionam que a linguagem XML “[...] possibilita a estruturação dos dados por meio da definição de regras sintáticas para a análise e validação dos recursos”.

Os *softwares* livres são muito indicados para os repositórios devido ao não pagamento de licenças de uso, bem como, pelo uso que está sendo feito em diversas instituições, fator que já lhes confere a qualidade necessária para implantação e customização.

A definição dos metadados é a etapa que descreve a estrutura dos documentos depositados em um repositório possibilitando, além da descrição, a interoperabilidade entre os sistemas de informação existentes tanto na instituição

quanto fora dela. Entretanto, cada comunidade, cada área do conhecimento trata o termo metadado mediante uma perspectiva. De acordo com Leite (2009, p. 60) “Alguns o utilizam para se referir à informação compreensível por máquinas, enquanto outros o utilizam para designar registros que descrevem recursos de informação digital”.

Existem diversos tipos de padrões de metadados e eles são definidos de acordo com a exigência do documento que será descrito, sendo menos complexos para documentos que exigem menor descrição e mais complexos para aqueles que exigem maior nível de descrição. Leite (2009) afirma que o padrão de metadados mais utilizado pelas ferramentas de estruturação de repositórios é o Dublin Core.

Os padrões de fluxo de trabalho (*workflow*) é a etapa que inclui tanto o fluxo da gestão do repositório, quanto o fluxo que o documento percorre na instituição até seu depósito. Esta etapa definirá o percurso da produção intelectual desde sua explicitação até o momento em que é disseminada e ocorre de acordo com as políticas estabelecidas pela instituição.

Algumas instituições optam por centralizar em suas bibliotecas o arquivamento dos documentos, de forma a garantir critérios de qualidade da indexação dos metadados. No entanto, nada impede a definição do autoarquivamento, desde que as etapas de submissão (captura de metadados, carregamento dos arquivos e atribuição da licença) sejam rigorosamente respeitadas pelo depositante, bem como a análise de metadados pela equipe gestora do repositório de modo a “assegurar o seguimento das normas estabelecidas para descrição, da correspondência entre descrição e arquivos carregados e outros” (LEITE, 2009, p. 68).

A questão da interface diz respeito ao conteúdo, desenho e estrutura que o usuário vai encontrar ao acessar o repositório. Desta forma, trata-se de uma etapa que deve ser planejada cuidadosamente para atender satisfatoriamente às necessidades de recuperação, facilitando a inserção e o depósito por meio de mecanismos que promovam a aceitação do repositório perante a comunidade a que ele se destina. Assim, faz-se fundamental “a adoção de níveis diferenciados de interfaces, que incluam a submissão de processos para usuários visitantes e para a administração do sistema” (TOMAÉL; SILVA, 2007, p. 8).

O aspecto interoperabilidade surge em torno da necessidade de compartilhamento e cooperação de informações e faz com que ocorra a universalização da organização da informação. Sobretudo pelo aspecto que considera que um sistema de informação de uma biblioteca, deve “ser capaz de gerar processos que sejam interoperáveis com os sistemas que estão à sua volta” (SAYÃO, 2008, p.136). Fator que gerará maior visibilidade do documento e evitará a repetição do trabalho do bibliotecário.

De acordo com Sayão e Marcondes (2008), a interoperabilidade possui algumas faces: a interoperabilidade técnica, que diz respeito aos padrões de comunicação, transporte, armazenamento e representação de informação; a interoperabilidade semântica, que trata das ferramentas de representação da informação como os metadados, sistemas de classificação, as ontologias, de forma que seja possível sistematizar que um termo tenha o mesmo significado independentemente do recurso; e finalmente e não menos importante, a interoperabilidade humana, fundamental pois exige trabalho interdisciplinar, discussões em grupo, estabelecimento de ações comuns, bem como a questão legal no que tange às exigências de disponibilizar livremente a informação e a interoperabilidade internacional, visto que os padrões e normas diferem devido ao estilo e falta de fundamentação comum.

Dentre as etapas que integrarão uma política para implantação de repositórios, entende-se que o item “Padrões” é o que permite, por meio dos aspectos de definição de metadados e interoperabilidade, relacionar os assuntos até então abordados com a área de organização da informação por meio dos elementos de organização (temática e descritiva) permitindo uma investigação acerca dos critérios de organização da informações existente em repositórios institucionais.

e) *Preservação*

Desde a etapa de implantação de um repositório, é necessário o estabelecimento de critérios para a preservação digital a fim de administrar “o ciclo de vida do documento digital com ênfase no acesso de longo prazo à informação digital, conservada em sua integridade” (TAMMARO; SALARELLI, 2008, p. 194).

A preservação diz respeito tanto ao acesso aos documentos, quanto à preservação do *software* que abriga tais documentos. Assim, atualizar ferramentas

é tarefa imprescindível do gestor de um repositório que vise manter o acesso à informação em linha.

De acordo com Tamaro e Salarelli (2008, p. 195):

[...] a preservação tem a finalidade de garantir o acesso à informação digital e aos pontos de acesso a ela (metadados). [...] um conjunto de decisões estratégicas e de procedimentos, junto com outros aspectos que dizem respeito à gestão dos dados, ajuda a garantir que a autenticidade de um recurso sobreviva no tempo, apesar das transformações tecnológicas, com uma perda mínima de conteúdo informativo, de funcionalidade e acessibilidade.

f) Políticas de acesso e uso

Este item é o que define quais serão as políticas adotadas na instituição a fim de determinar quais usuários terão acesso a quais níveis do repositório. Ou seja, determinar responsáveis pela submissão dos itens, se poderão submeter a uma ou mais coleções no repositório, bem como definir os usuários que terão o trabalho de controle e verificação dos metadados inseridos – geralmente os bibliotecários.

A política de informação institucional é que irá definir se os usuários terão acesso ao autodepósito ou se a própria instituição fará este trabalho, bem como irá garantir a segurança dos metadados, os critérios de qualidade e permissão antes da liberação do documento na *web*. Enfim, o software escolhido pela instituição deverá permitir a possibilidade de delimitação dos níveis de acesso.

Já as políticas de uso do repositório devem definir quais os usuários terão acesso e a quais tipos de documentos.

g) Sustentabilidade e financiamento

A sustentabilidade de um repositório é também um fator de extrema importância e deve ser levado em consideração por sua equipe gestora. Desde a sua implantação, faz-se necessária a previsão dos custos que envolvem ou envolverão seu desenvolvimento, dentre os quais, encontram-se os recursos humanos e os equipamentos. De acordo com Tomaél e Silva (2007, p. 10) “A equipe de trabalho é o componente de maior investimento de um projeto de repositório institucional. O pessoal técnico competente é necessário para iniciar, customizar e manter o sistema”.

No que tange aos equipamentos, caso a opção da instituição seja a de utilizar de *software* livre, de igual modo, haverá igual necessidade de investimento em *hardware* e em processos de capacitação para a customização e outras adequações que se fizerem necessárias.

Desta forma, a instituição que implantar um repositório deverá definir em sua política institucional a garantia de permanência deste projeto, a fim de que os investimentos sejam valorizados e, mais do que isso, que os documentos sejam preservados.

Percebe-se que definir uma política de informação implica em uma atividade complexa e de extrema relevância, visto que, partindo dela, serão norteadas ações desde a implementação até as etapas que garantam a sustentabilidade dos repositórios institucionais.

Visando trabalhar sob esta perspectiva, este estudo buscou abordar especificamente o aspecto de organização das informações depositadas em repositórios, levando em consideração dois elementos mencionados na etapa Padrões: os Metadados e a Interoperabilidade (semântica e humana), os quais serão abordados no próximo capítulo que apresentará os procedimentos de organização da informação.

5 ORGANIZAÇÃO DA INFORMAÇÃO EM REPOSITÓRIOS INSTITUCIONAIS

Dias e Naves (2007) conceituam a organização da informação como a área que engloba as disciplinas, as técnicas, os métodos e os processos relacionados à descrição física e temática dos documentos dispostos em uma biblioteca ou sistema de informação, desenvolvendo instrumentos como códigos, linguagens, normas ou padrões que possuem o objetivo descrever os documentos, além de estabelecer estruturas físicas ou não, que possibilitem armazenar os documentos e os seus conteúdos.

Entende-se que por meio da organização das informações contidas em uma unidade ou sistema de informação amplie-se a possibilidade de disseminação da informação; visto que se trata de uma etapa que cumpre o importante papel de intermediar a relação entre o produtor (autor) e o consumidor (usuário), atuando como mediador da informação (GUIMARÃES, 2003).

Com o objetivo de apresentar os elementos de organização da informação e descrever os procedimentos e os instrumentos utilizados, faz-se necessário iniciar este capítulo com uma reflexão sobre a gestão da informação: “A gestão da informação se caracteriza, entre outras atividades igualmente importantes, pela reflexão sobre formas de ordenar o conhecimento” (ARANALDE, 2009, p. 87).

Partindo desta afirmação, pode-se inferir que um dos objetivos da gestão da informação está em classificar e organizar o conhecimento registrado nos mais diversos suportes, devido à necessidade de identificar, selecionar e disseminar a produção intelectual existente em uma instituição ou unidade de informação.

Sendo os repositórios institucionais uma ferramenta de gestão da informação, cumpre a este estudo o papel de identificar como as informações estão sendo representadas no que tange ao seu aspecto físico, bem como nos aspectos de conteúdo.

Introduzindo este tema, Barité (2001, p. 39) informa que não é recente a necessidade de estruturar instrumentos para armazenagem e recuperação de informações, necessidade essa que remonta ao século XIX. Com a sociedade da informação torna-se ainda mais necessário este trabalho, muito embora a dificuldade de “controlar” ou mesmo “padronizar” o formato das informações publicadas a todo o

momento ao redor do mundo tenha aumentado substancialmente, seja no formato impresso ou digital.

A organização da informação em ambientes digitais é uma área bastante explorada pela comunidade científica, sobretudo pela diversidade de entendimento e de linhas de raciocínio sobre o tema. Há pesquisadores que defendem a aderência aos métodos tradicionais e há aqueles que divergem dos primeiros, adotando outras formas de organização que incluem: modelagem de dados, arquitetura de dados e outras.

Faz-se necessário esclarecer as diferenças ou mudanças de terminologia empregada no que tange à organização da informação tradicional e à organização em ambiente digital. De acordo com Monteiro (2008) o acesso aos conteúdos é o que difere a organização da informação tradicional da organização em ambiente digital; no método tradicional, uma classificação indica uma localização física de um documento em formato impresso, geralmente armazenado na estante. Por outro lado, a classificação da informação em ambiente digital indica a existência de *links* como forma de acessar os documentos digitais, seja em uma biblioteca digital ou em um repositório institucional.

A estrutura proposta a seguir demonstra que o objetivo maior da organização da informação está em representar um documento de forma que seja possível diferenciá-lo dos demais itens existentes, facilitando assim o acesso e uso do mesmo.

Todavia, o ambiente digital poderá ter como base alguns dos procedimentos tradicionais de organização da informação, exigindo novas estratégias e rotinas de organização da informação, denominados arquitetura da informação.

O quadro abaixo procura indicar os procedimentos e instrumentos adotados na organização da informação em bibliotecas tradicionais e digitais, de forma a sistematizar os elementos que farão parte deste capítulo.

ORGANIZAÇÃO DA INFORMAÇÃO EM AMBIENTES TRADICIONAIS			ORGANIZAÇÃO DA INFORMAÇÃO EM AMBIENTES DIGITAIS (BIBLIOTECA DIGITAL E REPOSITÓRIOS)					
	PROCEDIMENTOS	INSTRUMENTOS			PROCEDIMENTOS	INSTRUMENTOS		
DESCRIÇÃO FÍSICA	Catalogação	AACR2		DESCRIÇÃO FÍSICA	Catalogação apresentada em elementos de metadados	AACR2		
		MARC				MARC		
	Tipologia documental	Quanto à natureza				Tipologia documental apresentada em estrutura hipertextual	DUBLIN CORE	
Quanto à forma		Quanto à natureza						
Quanto ao conteúdo		Quanto à forma						
DESCRIÇÃO TEMÁTICA	Classificação	LC	CDD, CDU	DESCRIÇÃO TEMÁTICA	Classificação apresentada em estrutura hipertextual	LC	Esquema de classificação	
		LN	Termo extraído do documento			LN	Termo extraído do documento	
	Indexação	LC	Cabeçalho de assunto		Indexação apresentada em hipertexto ou elementos de metadados	LC	Esquema de classificação	
			Tesauro				Cabeçalho de assunto	
			Taxonomia				Tesauro	
	Resumo	LN	Texto sobre o documento		Resumo	LN	Texto sobre o documento	

Quadro 1 - Organização da informação em ambientes tradicionais e digitais.

Fonte: Adaptado de Monteiro (2008, p. 50 e 53).

A classificação apresentada em modelo de hipertexto pode ser visualizada como uma estrutura que define a forma como as informações serão organizadas nos ambientes digitais; “tem por objetivo criar um ‘mapa’ abrangente dos dados organizacionais e em seguida construir um sistema baseado nesse mapa” (MCGEE; PRUSAK, 1995, p. 129).

A arquitetura da informação definirá o formato em que os dados serão apresentados na página onde está disponível, projetando condições e níveis de acesso. Miranda (2005) cita que isto ocorre por meio dos sistemas de organização, de rotulação, de navegação e busca.

As estruturas de organização, rotulagem, navegação e busca são apresentadas por Rosenfeld e Morville (1988), os quais citam que os esquemas de organização da informação são ordenados de acordo com elementos exatos e ambíguos. Os elementos exatos são aqueles que categorizam a informação de

forma clara, caracterizado como o esquema de organização que atende às necessidades do usuário que sabe o que quer. Os elementos ambíguos, trabalham com categorias subjetivas, indicado àqueles usuários que não sabem exatamente do que precisam. Os esquemas híbridos são considerados sistemas capazes de associar linguagem natural e termos controlados. Segundo Lancaster (2004, p. 272) “a utilidade do método híbrido é apoiada pelo fato de que, na maioria dos estudos realizados, as buscas em texto livre recuperaram alguns itens relevantes que não foram identificados por buscas com vocabulário controlado, e vice-versa”.

Os sistemas de navegação e busca atuam possibilitando melhor interface, ou seja, melhor interação entre o ser humano e a máquina. Os sistemas de rotulagem, de acordo com Curty e colaboradoras (2009, p. 4), podem ser definidos como “símbolos linguísticos que são utilizados para expressar um conceito”, tais sistemas permitem acesso ao conteúdo de um *site*, geralmente por meio de menus e barras de navegação, em forma de texto ou ícone.

O termo arquitetura da informação foi concebido e estudado por profissionais da engenharia e arquitetura, dentre os quais se pode destacar Wurman, Rosenfeld e Morville, McGee e Prusak e outros. No entanto, é possível localizar estudiosos da área da ciência da informação, como Straioto (2002), e Vicentini (2006) que se utilizam da mesma terminologia a fim de designar as estruturas de organização da informação no ambiente *web*.

Para Straioto (2002, p. 20), esta estrutura de informação “[...] refere-se ao desenho das informações: como texto, imagens e sons são apresentados na tela do computador, a classificação dessas informações em agrupamentos de acordo com os objetivos do site e das necessidades do usuário”.

São estes os fatores de estruturação indispensáveis à customização de um repositório institucional, visando a uma interface de fácil entendimento e acesso, de forma a proporcionar um perfeito entrosamento entre o usuário que busca e o sistema que fornece a informação. O quadro apresentado por Monteiro (2008, p. 72) procura esclarecer o que cada esquema representa.

ELEMENTOS			DESCRIÇÃO
ESQUEMAS	EXATOS	ALFABÉTICO	Apresenta informações ordenadas alfabeticamente
		CRONOLÓGICO	Apresenta informações ordenadas por datas ou períodos temporais
		GEOGRÁFICO	Apresenta informações ordenadas segundo aspectos espaciais
ESQUEMAS	AMBÍGUOS	TÓPICO	Organiza informações por assunto
	HÍBRIDOS	ORIENTADO A TAREFA	Organiza informações segundo processos, funções ou ações que usuários podem executar
		ORIENTADO AO PÚBLICO	Organiza informações de acordo com os tipos de usuário e seus respectivos interesses
		ORIENTADO A METÁFORA	Organiza informações a partir de metáforas conhecidas pelo usuário
		Combina elementos de organização exatos e ambíguos	
ESTRUTURAS	Hierárquica (<i>top-down</i>)		Estruturas taxonômicas com classes e subclasses, partindo de assuntos gerais para os mais específicos
	Base de dados relacional (<i>bottom-up</i>)		Estrutura baseada em registros com descrições de conteúdo, partindo de assuntos específicos para os mais gerais.
	Hipertexto		Estrutura que agrupa conteúdos de acordo com o nível de semelhança ou critérios pré-definidos, de forma não linear e com componentes interligados por <i>links</i> .

Quadro 2 - Elementos de organização da informação segundo Rosenfeld e Morville

Fonte: Monteiro (2008, p. 72).

Os elementos básicos de arquitetura da informação aplicados às bibliotecas digitais/repositórios permitem criar estruturas digitais e priorizam a organização descritiva, temática, representacional, visual e navegacional de informação que estejam em harmonia com o conteúdo informacional “[...] o contexto e o usuário, tendo objetivos bem definidos, adequando o dimensionamento e o direcionamento dos serviços e dos produtos informacionais aos usuários potenciais” (VIDOTTI; SANT’ANNA, 2006, p. 90).

Ainda Vidotti e Sant’Anna (2006, p. 80) salientam que o foco da arquitetura da informação está nos:

[...] serviços e produtos digitais, de modo a criar sistemas de seleção, descrição, armazenamento, representação, indexação, recuperação e disseminação de informações digitais que possibilitem a construção, o armazenamento e a disseminação do conhecimento.

De acordo com Fujita (2008, p.4), é importante salientar que:

[...] os avanços da área têm se pautado no fortalecimento de sua área teórica, processos, produtos e métodos a partir do conhecimento e reconhecimento da relevância científica e social adquirida ao longo dos tempos, tendo em vista resultados como os sistemas de classificação universais, tabela Cutter, a análise facetada, mudanças tecnológicas, etc.

Ao abordar as técnicas utilizadas e seus instrumentos de forma que seja possível atingir aos objetivos propostos neste estudo, faz-se necessário identificar cada uma das etapas que compõem a Organização da Informação, que na literatura da área de Ciência da Informação é apresentada de forma sinônima a: “tratamento da informação” e “representação da informação”, são elas: Tratamento Descritivo (Descrição Física), que representa os dados físicos de um documento e o Tratamento Temático (Descrição Temática), que se refere aos termos utilizados para expressar seu conteúdo, facilitando o acesso, etapa de fundamental importância para extração de conceitos que, traduzidos para a linguagem de indexação, irão compor os vocabulários controlados de uma especialidade.

De acordo com Galvão (2003), a linha de pesquisa organização da informação, consiste em representar a informação de forma que seja possível explicitar os dados que remetam tanto à autoria de um documento, quanto ao local onde fora produzido, bem como aos conteúdos contemplados por aquele documento.

Buscou-se ordenar os apontamentos deste capítulo conceituando inicialmente o termo “documento”, posteriormente apresentando o ciclo documentário e finalmente abordando os processos de descrição física e temática e os instrumentos utilizados.

Ao iniciar uma abordagem sobre organização da informação, faz-se necessário conceituar o termo documento, que de acordo com Medeiros (2009), pode ser definido como “todo registro de informação apresentado em forma impressa, visual, auditiva, tátil, eletrônica ou combinada, que pode conservar-se através do tempo e do espaço”.

Os documentos possuem características físicas, que são relacionadas ao suporte em que é apresentado, ao seu tamanho e peso; possui também características intelectuais, relativas a forma de difusão, originalidade, conteúdo, assunto e grau de elaboração, que diz respeito à autoria e finalidade a

que o documento se destina, sendo possível classificar os documentos em primários, secundários e terciários (FEITOSA, 2006).

Segundo o Núcleo de Informação, Pesquisas e Estudos em Educação Bibliotecária (2009), as tipologias documentárias são representadas por publicações avulsas (documentos primários e secundários) e por publicações seriadas e trabalhos em eventos. Os documentos primários são: apostila, capítulo de livro, dissertação, folheto, legislação, livro, material didático, monografia de curso de especialização, projeto, publicação técnica, relatório e tese. Os documentos secundários são: bibliografia, cadastro, catálogo, dicionário, enciclopédia, índice, glossário, guia, manual e tesouros. As publicações seriadas são os artigos em boletins, circulares, jornais, artigo de periódicos técnico-científicos e revistas, bem como os trabalhos publicados em eventos, os anais.

Há tipologias documentárias apresentadas em sistemas de hipertexto, trata-se da apresentação não-linear que possibilita ao leitor ampliar as suas possibilidades de busca dentro de um mesmo texto.

O artigo "*As we may think*", de Vannevar Bush publicado em 1945 apresenta um sistema denominado *Memex* capaz de organizar e recuperar a informação por associação entre assuntos, sistema que foi caracterizado como similar à mente humana. Depois disso, já em 1960, é desenvolvido o sistema *Xanadu*, cujo objetivo foi o de oportunizar um ambiente literário unificado. Esse sistema, criado por Ted Nelson, permitiu a definição do termo hipertexto como: "uma combinação de texto em linguagem natural com a capacidade do computador de fazer pesquisa interativa e exibição dinâmica de um texto não linear" (NONATO; LIMA, 2008).

Destarte, é possível mencionar que, no ambiente de hipertexto, as informações são interconectadas por *links* que associam os conteúdos de determinado tema. Nonato e Lima (2008, p. 198) ao referirem-se aos sistemas de hipertexto, indicam que "[...] consistem em uma abordagem de estruturação e manipulação de textos. Em tais sistemas os documentos são dispostos em uma base de dados repleta de conexões, formando uma rede hipertextual".

Alvarenga (2006) infere que existe uma diversidade de formato de documentos, dentre eles os textuais, os sonoros e os imagéticos, mostrando-se como um espaço de registro e recuperação de informação, fator que na mesma medida passa a necessitar de novos elementos que facilitem a recuperação. Desta

forma, neste estudo, a abordagem do termo documento, esta relacionada ao conhecimento explícito que, registrado e socializado passa a servir de base para a geração de novos conhecimentos.

Deve-se considerar que a forma estática, característica dos documentos impressos, lentamente vai sendo substituída pela dinâmica com que os conteúdos em formato eletrônico estão sendo intensificados diariamente, tornando cada vez mais necessária sua organização, visto que as tecnologias de informação proporcionaram inúmeras mudanças, não somente no formato de publicação do conhecimento, que passaram do formato impresso para o eletrônico; mas causaram também mudanças nas formas de tratamento, armazenagem e disseminação das informações.

Entretanto, deve-se levar em consideração que a tecnologia e os recursos oferecidos por ela dependem de conhecimentos, procedimentos e/ou capacidades necessárias à sua utilização. A própria etimologia da palavra tecnologia é suficiente para esclarecer o ponto em questão. “Segundo o dicionário Aurélio ‘tecnologia’ vem dos termos gregos *tecn(o)* + *-log(o)* + *ia* e se define como ‘o conjunto de conhecimentos, especialmente princípios científicos, que se aplicam a um determinado ramo de atividade’” (MERKLE, 2000, p.2).

Sobretudo porque a forma tradicional de organização, que utiliza de instrumentos e linguagens documentárias, raramente é encontrada nos ambientes digitais. Isso ocorre devido aos custos, à quantidade de informação disponível e também pela possibilidade de pesquisa por palavras/termos livres, que possibilitam a recuperação dos objetos digitais por meio de mecanismos de busca encontrados, por exemplo, nas páginas como as do *Yahoo*, *Wikipédia*, *Google*.

Toutain (2006, p.15) define o documento digital como um “registro de informação codificado por meio de dígitos binários”. Assim, todo material que pode ser lido e decodificado pode ser caracterizado como um documento.

De acordo com a abordagem de Tammaro e Salarelli (2008, p.3) “passou-se a considerar como documentos instrumentos, objetos, prédios, vestuário, alimentos, ou seja, tudo quanto se possa revelar como portador de significados”. Ou seja, para estas autoras, não se entende por documento somente aquilo que é legível por meio da palavra escrita, e sim a capacidade que o pesquisador possui de questionar um material cujo suporte permita a disseminação das informações.

Julga-se relevante mencionar que todo documento seja ele impresso ou em formato digital e que passa a fazer parte de um sistema de informação recebe um tratamento que visa selecionar, avaliar, analisar, traduzir e, finalmente, recuperá-lo, constituindo, assim, o denominado ciclo documentário. “As principais atividades do ciclo documentário envolvem a coleta, o registro, o tratamento intelectual, a pesquisa e a difusão” (FEITOSA, 2006, p. 19).



Figura 1 – O Ciclo Documentário – visão sistêmica
Fonte: Feitosa (2006, p.19)

Conforme a ilustração apresentada, a primeira etapa do ciclo documentário corresponde à atividade de identificar, selecionar e adquirir os documentos. A segunda, denominada atividades de processamento, relaciona-se à etapa que permitirá o tratamento documentário; e a terceira etapa, é caracterizada pelos produtos e serviços do sistema de informação planejados de acordo com a comunidade usuária, dentre os quais se pode citar as atividades de levantamentos e consultas bibliográficas, atividades de empréstimo, comutação documentária, entre outros (CERVANTES; FUJITA; RUBI, 2008).

Feitosa (2006, p. 20) apresenta um modelo esquemático que representa as atividades do ciclo documentário:

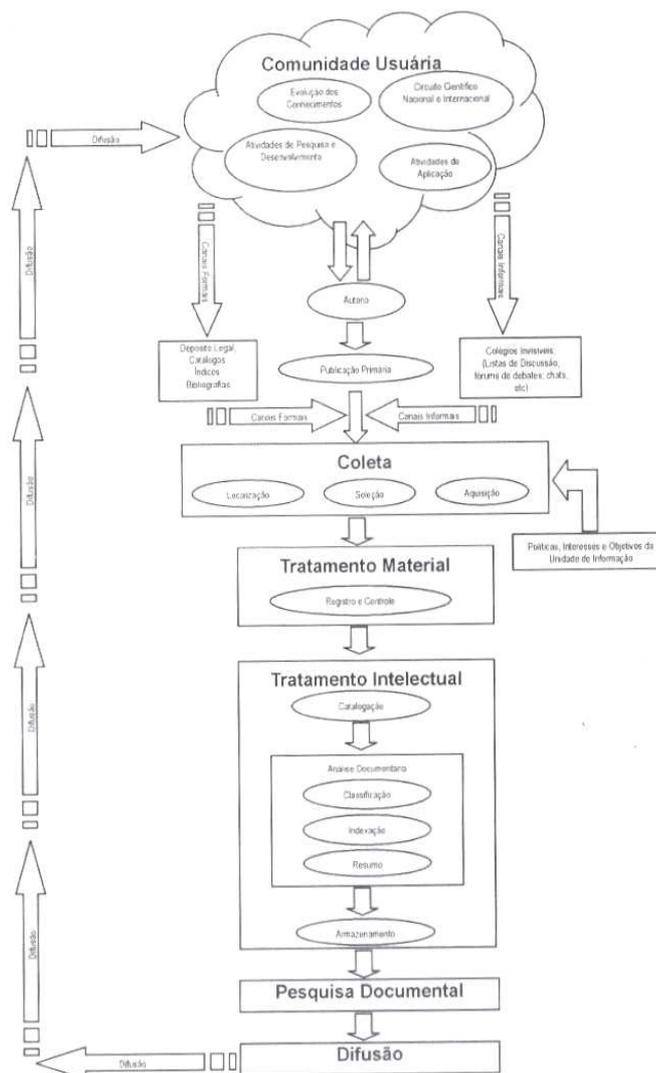


Figura 2 - Atividades do ciclo documentário
Fonte: Feitosa (2006, p.20).

Observando a figura, pode-se entender que a comunidade usuária é composta tanto por aqueles que produzem as informações, quanto pelos que a utilizam. É nesta primeira camada que ocorrem as pesquisas e por onde elas passarão a ser divulgadas, seja por meio dos canais formais ou informais de transmissão. A segunda etapa envolve: a coleta do documento que explicita o conhecimento produzido, sua localização, seleção e aquisição, que devem estar de acordo com as políticas da biblioteca e/ou unidade de informação. A próxima etapa envolve o tratamento material, que engloba o registro e controle deste documento que irá para a etapa de tratamento intelectual, que ficará responsável por sua catalogação, classificação, indexação e resumo para que possa ser o documento armazenado e disseminado. Desta forma é possível afirmar que o ciclo inicia e termina tendo como foco o usuário da informação.

Caracterizadas as principais atividades de um ciclo documentário como a coleta (seleção) e o tratamento material, faz-se necessário abordar os processos que determinarão o tratamento intelectual destes documentos permitindo a organização das informações para que estas possam ser disseminadas.

Neste estudo, optou-se por trabalhar com a identificação dos procedimentos e instrumentos que compõem o tratamento dos documentos, compreendendo operações como a Representação Descritiva (Catalogação) e Representação Temática (Classificação).

A representação descritiva é uma especialidade da área da organização da informação, também conhecida como análise descritiva, que está relacionada aos elementos como autoria, título, editora e demais informações facilmente identificáveis no documento (DIAS; NAVES, 2007); exige trabalho intelectual, realizado por profissionais da área de ciência da informação, tendo em vista a complexidade do uso de instrumentos que possam subsidiar a descrição dos elementos presentes no documento.

Um dos procedimentos utilizados na representação descritiva é a catalogação, entendida como uma forma de descrever ou representar os documentos, tradicionalmente encontrada em formato de fichas catalográficas, podendo ser denominada catalogação descritiva (quanto aborda o aspecto físico do documento), ou catalogação por assunto (quanto diz respeito ao conteúdo). Essa técnica requer tratamento realizado por especialistas, os quais possibilitarão ao usuário a localização da informação de que necessita.

Classificar é a atividade que permite caracterizar um item, possibilitando tanto sua individualização entre os demais itens documentários, como seu agrupamento por semelhança. “Cria alternativas de escolhas para seus usuários. Mas deve permitir que, feita a escolha do item, o usuário seja capaz de localizá-lo no acervo” (ANZOLIN, 2007, p. 3).

De acordo com Feitosa (2006, p.21), a catalogação ou descrição bibliográfica tem por objetivo representar um documento de forma que seja possível descrevê-lo “materialmente, de forma única e não ambígua, de modo a permitir sua identificação, localização e representação em catálogos ou em outros instrumentos que facilitem a sua localização física”.

Para que um documento seja catalogado e inserido em um sistema de informação, é necessária a utilização de um instrumento internacional de

normalização denominado AACR2 (*Anglo-American Cataloguing Rules*), código de classificação, que possui uma estrutura dividida em duas partes: a primeira delas aborda a descrição e a segunda os pontos de acesso como, títulos uniformes e remissivas.

Especificamente neste estudo, a abordagem da representação descritiva terá maior enfoque na organização da informação no ambiente digital, no qual se emprega “o termo ‘Metadado’ como designativo do produto da descrição documental e dos denominados pontos de acesso, resultantes dos processos da área da catalogação bibliográfica” (ALVARENGA, 2006, p. 93).

Este termo refere-se ao recurso eletrônico que descreve um documento em ambiente digital, possibilitando sua integração e seu compartilhamento; sua função é a de padronizar a entrada de dados em sistema de informação a fim de possibilitar uma recuperação de qualidade.

A respeito do ambiente digital, Monteiro (2008, p. 51) cita que “a organização da informação compreende a descrição dos documentos de acordo com seus aspectos físicos e temáticos sendo que, no ambiente digital, é realizada com metadados”.

Os metadados podem ser categorizados em: metadados para catalogação bibliográfica, aqueles que desempenham papel similar aos códigos de catalogação utilizados para descrever documentos (como o MARC); e em metadados para descoberta de recursos na *web*, aqueles que servem como suporte para os motores de busca (como o padrão Dublin Core).

Utilizada para descrever os documentos, a sigla MARC é proveniente de *Machine-Readable Cataloging* e pode ser definida como um padrão que tem por objetivo identificar, armazenar e disseminar informações bibliográficas em formato legível por máquina, de forma que diferentes computadores e programas possam compartilhar dos pontos de acesso a elementos que compõem a descrição bibliográfica. Pode-se dizer que este formato permitiu a substituição dos catálogos manuais, que referenciavam um objeto físico, vinculado geralmente a uma unidade de informação; os metadados referenciam objetos remotos, na maioria das vezes desvinculados de uma instituição ou ambiente físico.

Esse formato utiliza-se de campos, subcampos e indicadores, os quais possibilitam que os dados possam ser legíveis por máquinas. Sua estrutura é

apresentada por *tags* (etiquetas) de três dígitos numéricos que representam cada campo, conforme segue:

ELEMENTOS	DESCRIÇÃO
8XX	ENTRADAS SECUNDÁRIAS DE SÉRIE
0XX	INFORMAÇÃO DE CONTROLE, NÚMEROS E CÓDIGOS
1XX	AUTORIA (NOME PESSOAL, ENTIDADE, EVENTO)
2XX	TÍTULOS, EDIÇÃO E IMPRENTA
3XX	DESCRIÇÃO FÍSICA
4XX	SÉRIE
5XX	NOTAS
6XX	ENTRADAS DE ASSUNTO
7XX	ENTRADAS SECUNDÁRIAS (NOME PESSOAL, ENTIDADE, EVENTO, TÍTULO)

Quadro 3 – Elementos de Metadados MARC

Analisando a estrutura desse quadro, percebe-se grande semelhança com o formato de representar a informação nas tradicionais fichas catalográficas, utilizadas para representar o documento impresso, enquanto que, no formato eletrônico, são representados por metadados.

Já o padrão *Dublin Core* pode ser definido como uma linguagem utilizada para elaboração de classes particulares de declarações sobre recursos. “Nesta linguagem, há duas classes de termos: elementos (nomes) e qualificadores (adjetivos), que podem ser arranjados como um padrão simples de instruções” (MORI; CARVALHO, 2004, p. 13).

Inicialmente, o *Dublin Core* foi criado para auxiliar os mecanismos de busca a recuperar páginas *web*, desde então, evoluiu para um formato de intercâmbio e de recuperação de informação no espaço digital (MÉNDEZ RODRÍGUEZ, 2002).

Esse padrão é composto de quinze elementos semânticos que podem ser descritos “como o mais baixo denominador comum para a descrição de recurso (equivalente a uma ficha catalográfica)” (SOUZA; VENDRUSCULO; MELO, 2000, p. 93).

ELEMENTOS	DESCRIÇÃO
TÍTULO	Nome dado ao recurso
CRIADOR	Entidade originalmente responsável pela criação o conteúdo do recurso
ASSUNTO	Tema do conteúdo do recurso. Pode ser expresso em palavras-chaves e ou categoria recomenda-se o uso de vocabulário controlado.
DESCRIÇÃO	Relato do conteúdo do recurso. Ex: textos livres, sumário e resumo.
PUBLICADOR	Entidade responsável por tornar o recurso disponível.
COLABORADOR	Entidade responsável pela contribuição intelectual ao conteúdo do recurso.
DATA	Data associada a um evento ou ciclo de vida de um recurso.
TIPO	Natureza ou gênero do conteúdo do recurso.ex: texto, imagem,som,dados e software
FORMATO	Manifestação física ou digital do recurso. Ex: html, ppt, pdf, gif, xls
IDENTIFICADOR	Referência não ambígua (localizador) para o recurso dentro de dado contexto.
FONTE	Referência a um recurso do qual o presente é derivado.
IDIOMA	Língua do conteúdo intelectual do recurso.
RELAÇÃO	Referência para um recurso relacionado
COBERTURA	Extensão ou escopo do conteúdo do recurso; pode ser temporal e espacial
DIREITOS	Informação sobre os direitos assegurados dentro e sobre o recurso.

Quadro 4 – Elementos de metadados *Dublin core*

Analisando os padrões acima, pode-se dizer que o objetivo é descrever os dados para que possam ser recuperados. O *Dublin Core* é um padrão de fácil utilização devido ao pequeno número de atributos que apresenta. Importante mencionar que cada sistema de informação determinará quais campos serão utilizados em sua catalogação, visando sempre o compartilhamento de informações e a interoperabilidade que, para Méndez Rodríguez (2002), é concretizada com o trabalho humano de definir conceitos e transferi-los para metadados que, por meio de protocolos, possibilitem que sistemas de informação dispersos possam ter condições de acesso integrado aos seus conteúdos.

Dessa forma, os metadados possibilitam a otimização das buscas e recuperação da informação na *web*, descrevendo de forma ordenada um conjunto de dados; tarefa similar à catalogação tradicional, que propõe como produto final a elaboração de catálogos como forma de acesso à informação. Pode-se dizer que os “metadados correspondem a pontos de acesso temáticos, descritivos e de conteúdo dos documentos”, para que haja condições de compartilhamento de dados entre os sistemas de informação (ALVARENGA, 2006, p. 93).

Os metadados têm como objetivo a organização das informações nos sistemas digitais, fornecendo um novo conceito, uma nova abordagem de estruturação, em especial no que tange à possibilidade de compartilhamento das informações entre diversos sistemas de informação, fator que torna ainda mais

necessário o estabelecimento de padrões para representação destas informações em sistemas de recuperação.

Méndez Rodríguez (2002, p. 41) salienta que o termo metadado “remete a qualquer dado que descreve o atributo de um recurso, caracteriza suas relações, apoia sua recuperação, gestão e uso e existem em um contexto estritamente eletrônico”.

Para os profissionais da área da Ciência da Informação, o termo metadado é utilizado como “[...] um tipo de dado que é usado para descrever o conteúdo, a estrutura, a representação e o contexto de um conjunto de dados específico” (MÉNDEZ RODRÍGUEZ, 2002, p. 35). No domínio dessa área de atuação os metadados são aplicados especialmente a coleções pertencentes a bibliotecas digitais e/ou informações presentes nos repositórios de informação. No entanto, qualquer comunidade científica que utilize de sistemas de informação baseados em sistemas digitais necessita criar estruturas para organizar suas informações e isso ocorre no formato de metadados.

Os padrões de metadados mais utilizados são: IAFA (*templates*), MARC, *Text Encoding Initiative (TEI) Headers*, *Encoded Archival Description (EAD)*, *Computer Interchange of Museum Information (CIMI)*, *Government Information Locator Service (GILS)* e *Content Standards for Digital Geospatial Metadata*. Vale citar que cada padrão de metadado possui uma estrutura de combinação de elementos, levando em conta o público a que ele se destina (ALVARENGA, 2006).

Enfim, é necessário salientar que os metadados estão entre o processo do tratamento e a recuperação da informação na *web*. Trata-se de um novo modelo de gestão da informação, uma nova tendência de organização da informação, na qual as ferramentas de recuperação da informação na rede, independente dos dados, estão em um servidor local ou em rede e se fundamentam na automatização das tarefas tradicionais de organização do conhecimento (MÉNDEZ RODRIGUÉZ, 2002).

A representação temática, segundo Feitosa (2006), consiste na análise do conteúdo de um documento, possibilitando o entendimento do tema ou assunto abordado por ele, objetivando a sua disseminação. Assim, pode-se inferir que o maior objetivo de biblioteca ou unidade de informação é o de oportunizar a recuperação dos conteúdos dos documentos que ela possui, e isso ocorre por meio da análise de assunto, que é a primeira etapa do tratamento temático.

A análise de assunto, também conhecida como análise temática, análise documentária ou análise de informação, é a primeira etapa do Tratamento Temático e tem por função identificar as informações de um documento, de forma que possam ser extraídos os conceitos que representem o seu conteúdo. Dias e Naves (2007) citam que esse processo consiste na leitura do documento a fim de traduzir seu conteúdo.

Segundo Guimarães (2003), a análise de assunto implica o resgate de um processo cíclico de informação como alicerce do fazer documental, inicia-se na produção, passando pelo processo de tratamento ou organização, que possibilitará a recuperação, a disseminação e o uso da informação. Por se tratar de um processo cíclico, ao finalizá-lo inicia-se uma nova produção.

Após a análise do assunto, inicia-se a etapa de classificação do documento. De acordo com Vickery (1980, p. 23) “Classificar, na acepção mais simples do termo, é reunir coisas ou idéias que sejam semelhantes entre si, e separar as que apresentam diferenças”.

Dias e Naves (2007, p. 16) salientam que classificar consiste em analisar o conteúdo de um documento com o objetivo de determinar um lugar para seu armazenamento, organizando-o por assunto, e também utilizar um sistema de classificação bibliográfica para que possa representar os assuntos existentes em um documento.

Esta atividade é necessária, sobretudo, porque um conhecimento não possui uma estrutura definida. Assim, definir conceitos e categorias são a base para organização e representação do conhecimento. O conceito, de acordo com Cervantes (2009) trata-se da “unidade de conhecimento que compreende afirmações necessárias e verificáveis sobre um referente representada por uma designação”.

As categorias podem ser entendidas como conceitos de alta generalização que são classes que resultam da divisão do universo de conhecimento, de acordo com as características intrínsecas ou fundamentais de cada conceito (CAVALCANTI, 1978).

Nessa perspectiva, faz-se importante mencionar a teoria do conceito, que tem o objetivo de esquematizar a definição de um conceito, sua composição e relacionamentos. Essa teoria foi criada como uma abordagem de categorização temática para estruturação dos conceitos, por Ingetraut Dahlberg.

Esta afirma que qualquer organização do conhecimento deve ser baseada em unidades do conhecimento, que são nada mais que conceitos (LIMA, 2006).

Dias e Naves (2007, p. 66) citam que “a extração de conceitos tem como produto um assunto, que representa o conteúdo informacional de um texto”. Pode-se inferir que no momento da recuperação da informação, a indexação por assunto pode não apresentar o resultado esperado, devido a sua ambiguidade. Foskett (1973, p.40) menciona que:

O emprego de uma linguagem artificial nos permite usar a indexação de conceitos ao contrário de indexação de termos. Na indexação de termos, dependemos das palavras empregadas por cada autor a fim de obter as descrições de assuntos de que precisamos, na expectativa de que quando tentarmos localizar as informações para que possamos harmonizar os diferentes termos usados por diferentes autores.

Assim, para que a recuperação da informação ocorra de forma adequada, a indexação por conceitos permitirá maior padronização dos termos – por meio de conceitos – possibilitando localizar uma mesma informação, e que essa possa ser igualmente recuperada. Busca-se definir o termo conceito, conforme análise apresentada por Piedade (1983, p.35): “Conceito é a operação da inteligência através da qual se apreendem os caracteres essenciais daquilo que se conhece. É a representação mental do que se sabe, uma idéia, uma coisa, um julgamento, etc”.

Para Dias e Naves (2007, p. 64), “definir um conceito é uma operação verbal e lógica bem clara, na qual se usa uma série de idéias logicamente subordinadas para chegar a uma conclusão geral”.

Segundo Alvarenga (2001), a teoria do conceito se constitui de enunciados provenientes de reflexões e pesquisa, compreendendo a representação, comunicação e preservação de objetos e pensamentos, cujo conhecimento integra a filosofia, a linguagem, a psicologia, a comunicação e a ciência da informação.

Emergem nesse contexto a relevância de determinação dos conceitos presentes em textos, imagens, som e outras formas de representação, inclusive as disponíveis no ciberespaço. Alvarenga (2001, p. 8) cita que “o que na realidade se classifica em uma biblioteca tradicional ou digital não são os documentos, mas os conceitos contidos nesses documentos”.

O tratamento temático da informação tem como etapa a exigência de um trabalho intelectual do indexador ou catalogador, a leitura do documento a fim de que possa extrair conceitos e também definir termos em linguagem natural para que posteriormente, possam ser repassados para a linguagem de indexação.

Naves (2006) relata que a classificação do conhecimento iniciou-se com Aristóteles, que defendeu um modelo baseado em atributos das entidades materiais, que permaneceu aproximadamente por dois mil anos entre o período de (300aC a 1600dC). Esse modelo dividia o conhecimento em três partes: teórica (física, matemática e metafísica), prática (ética, política, economia e retórica) e produtiva (poesia e artes).

De acordo com Aranalde (2009, p.90), o filósofo alemão Immanuel Kant, contribui com a organização do conhecimento uma vez que “se dedica a rever a formulação aristotélica das categorias, reelaborá-las, justificá-las e apresentá-las em uma tábua de categorias em que elas se relacionam reciprocamente e adquirem sentido”.

Mais tarde, por volta do século XIX, com intuito mais subjetivo de arranjo, Bacon desenvolve um esquema de classificação baseado na memória e na razão. Já o filósofo grego Porfírio, apresenta como alternativa à classificação proposta por Aristóteles, um modelo de classificação denominado Árvore de Porfírio.

Durante muitos séculos, os métodos de organização do conhecimento foram dicotômicos, de acordo com os apontamentos de Campos (2004, p. 25), “na dicotomia encontram-se duas divisões no primeiro estágio, duas divisões de cada uma destas divisões são formadas no segundo estágio e assim por diante. A representação esquemática da dicotomia chama-se ‘Árvore de Porfírio’”.

De acordo com Gomes 1996, *apud* Tristão, Frachin e Alarcon (2004), existem vários tipos de relacionamentos entre conceitos, podendo-se destacar: relacionamento lógico (genérico/específico, analítico, de oposição), relacionamento ontológico (partitivo, de sucessão, material/produto) e relacionamento de efeito (causalidade, instrumental, descendência, genealógica, estágios da substância, antogênica).

Partindo destes pressupostos faz-se relevante destacar ainda a contribuição de Shiyali Ramamrita Ranganathan, que se dedicou a estudar os esquemas de classificação, contribuindo para a organização do conhecimento no que apresentou pontos positivos e negativos dos modelos propostos até o momento.

Opondo-se aos esquemas rígidos, propôs um sistema flexível – uma estrutura facetada denominada Classificação de Dois Pontos (*Colon Classification*). Assim, Ranganathan apresenta outras formas de subdividir o conhecimento, como todo/parte, propriedade/possuidor, ação/paciente e outros, aproximando-se mais das teorias de Aristóteles.

Ranganathan propõe, ainda, uma mudança ao modelo dicotômico de Porfírio, por meio da analogia com um tipo de figueira indiana chamada Árvore Baniana, no qual o pensador apresenta que este modelo é o mais viável a uma árvore de classificação visto que “do tronco original formam-se muitos outros troncos secundários, de tempos em tempos. Cada tronco foi considerado por Ranganathan como categorias. Assim, os domínios teriam um corte policotômico, e não dicotômico” (CAMPOS, 2004, p. 25).

As reflexões deste pensador apresentam o maior desenvolvimento teórico para a teoria da classificação no século XX (ARANALDE, 2009), no qual surgiram também outros estudiosos que apresentaram características mais empíricas para a organização do conhecimento, dentre eles Dewey, Cutter, Brown, Bliss e a *Library Congress*.

Os principais instrumentos utilizados para descrição temática dos documentos são a CDD e a CDU. A Classificação Decimal de Dewey (CDD), publicada em 1876, tem a autoria de Melvil Dewey, esta foi a primeira classificação baseada em números decimais para representação de assuntos; nesta classificação, as classes principais correspondem às áreas do conhecimento (LANGRIDGE, 2006, p. 84)

Por volta de 1880, a Classificação de Cutter (*Cutter Expansive Classification*) foi desenvolvida por Charles Ammi Cutter, com o objetivo de utilizar letras, contrastando com as demais teorias de classificação. Estrutura-se basicamente em três linhas, cabendo à primeira delas representar o assunto e a segunda representar o autor; já a terceira indica pela letra Y um trabalho sobre o autor ou obra indicado nas duas primeiras linhas, e iniciando com letras para indicar traduções para outras línguas. O instrumento utilizado na construção desta classificação é a Tabela Cutter. Esta tabela utiliza-se dos números a fim de representar sobrenomes de autores.

A CDU (Classificação Decimal Universal) surge em 1892, com Paul Otlet e Henri La Fontaine, que propuseram um controle por assunto, tendo como

base a CDD. Esses estudiosos inovaram a representação, utilizando-se de fichas catalográficas. Na CDU “a notação passou a ser mista (números, sinais, letras ou palavras)” (PIEDADE, 1983, p. 122).

Tanto a CDD quanto a CDU “[...] seguem os predicáveis de Porfírio, partem do geral para o específico, dividindo o conhecimento consecutivamente por diversas características, tomando por modelo a Árvore de Porfírio” (PIEDADE, 1983, p. 78). As teorias até então discutidas levavam em consideração que qualquer área do conhecimento era subdividida do geral para o específico, como no caso de gênero e espécie.

Langridge (2006, p. 11) é categórico ao afirmar que “[...] sem classificação não poderia haver nenhum pensamento humano, ação e organização que conhecemos”. Assim, a importância de classificar, categorizar, estruturar as informações é intensificada ainda mais com a explosão informacional, decorrente da Sociedade da Informação e das mudanças ocorridas com as possibilidades trazidas pelas tecnologias de informação e comunicação (TICs). Novos modelos de disseminação e compartilhamento da informação surgem, o que para Andrade (2006) passa a exigir ainda mais da classificação do conhecimento, de forma que o acesso ao usuário final seja facilitado, levando em consideração o usuário como um ser social e não um apenas como um receptor passivo de informações.

Lima (2006) cita que classificar é considerar tanto o contexto quanto o conteúdo de um documento, visando que à categorização de conceitos que auxiliem o usuário em suas buscas.

Já a indexação é uma atividade intelectual que visa compreender o texto e compor os elementos que irão representá-lo mediante as etapas de análise e identificação do assunto, identificação dos conceitos e adequação desses às linguagens de indexação. Dessa forma, indexar é incluir dados de um documento em um sistema de recuperação de informação, realizando sua identificação, os assuntos, bem como quais termos serão utilizados para sua representação (LIMA, 2006).

De acordo com Feitosa (2006, p. 21) “a indexação consiste em atribuir um ou vários descritores, que podem ter ligações entre si, para descrever o conteúdo do documento”. Consiste em analisar o assunto e traduzir os conceitos presentes no documento para uma linguagem controlada. Nonato e Lima (2008) descrevem a análise de assunto como uma atividade que ocorre em três etapas:

compreender o conteúdo do documento, identificar seus conceitos e selecionar aqueles que possam permitir a recuperação da informação. Já a tradução dos conceitos é a atividade que assegura a organização dos mesmos no momento da representação utilizando-se de instrumentos verbais ou simbólicos.

Piedade (1983) ressalta que, para tratar o tema de um documento, pode-se usar a indexação por palavras ou por conceito. A primeira é baseada na forma como as palavras aparecem no documento (linguagem natural); um exemplo disso são os índices que podem ser empregados para indexação por palavras, utilizando o título do documento. A indexação por conceito analisa o conteúdo temático do documento, determinando cabeçalhos de assunto a serem utilizados para o controle de homônimos e sinônimos e as ligações hierárquicas; isto é, trata-se de uma linguagem controlada. Para apresentar os conceitos, a linguagem controlada ou artificial os divide em pré-coordenados ou pós-coordenados. O índice formado pela indexação de conceitos compostos, ou seja, formado de vários conceitos são combinados no momento da indexação, caracterizando-a como pré-coordenados. Os pós-coordenados são constituídos de conceitos simples, tendo os resultados de conceitos compostos, somente no momento da recuperação.

O tratamento da informação é a atividade que tem por objetivo aproximar o usuário da informação. Essa atividade é inerente aos profissionais da informação. Segundo Fujita e Rubi (2006, p. 10), “o indexador tem a função primordial de compreender o documento e realizar uma análise conceitual que represente adequadamente seu conteúdo”.

Percebe-se que a tecnologia e todas as implicações decorrentes dela proporcionaram uma explosão de informação, e conseqüentemente ampliaram-se os registros e suas formas de disseminação por meio da *internet*. O que ocorre é que a tarefa de indexação dos dados está sendo realizada por máquinas. Caso esses trabalhos fossem executados por especialistas, o custo seria inviável levando-se em consideração a quantidade de informação a ser tratada.

Assim, apresenta-se a indexação automática que, de acordo com Robredo (1982, p. 238), é semelhante à leitura-memorização humana, baseando-se em “[...] cada palavra do texto com uma relação de palavras vazias de significado, previamente estabelecida, que conduz, por eliminação, a considerar as palavras restantes do texto como palavras significativas”; o que é relevante, sobretudo pelo

aspecto que ao indivíduo interessa o que as palavras representam no contexto em que ele estão inseridas.

Kuramoto (2006b, p. 118), no entanto, ressalta que “[...] os procedimentos de indexação automática têm se revelado pouco eficazes no sentido de possibilitar o acesso preciso à informação”. Esse fato se dá, em especial, devido ao uso de termos livres e/ou palavras, fator que demonstra que a indexação necessita de novos modelos. Se por um lado os procedimentos de indexação envolvem técnicas e métodos de reconhecimento e extração de informação de uma coleção de documentos, por outro esses procedimentos envolvem também a organização e a estruturação dessas unidades com vista a um acesso rápido e preciso da informação.

No método tradicional, o indexador utiliza-se de seu conhecimento tácito aliando-o ao uso de tesouros, vocabulários controlados e outros instrumentos para indexar um assunto. Isso demanda análise e pesquisa; a eficácia da recuperação da informação depende da qualidade da indexação. Já na indexação automática, os termos são extraídos, de forma genérica, do próprio conteúdo/documento, levando em conta sua importância e a quantidade de vezes que o termo se repete no documento (KURAMOTO, 2006a).

Dessa forma, infere-se que a pesquisa realizada em sistemas e/ou unidades de informação ocorre, na maioria das vezes, com palavras/termos da linguagem natural, que podem ser polissêmicas ou sinônimas, fator que pode inviabilizar o resultado, tendo em vista a recuperação inadequada do conteúdo pesquisado. Sobre isso Kuramoto (2006b, p. 127) destaca que, ao invés de indexar palavras, o sistema de indexação automática poderia identificar descritores “[...] ou seja, elementos que façam referência à realidade extralingüística do autor do documento”.

Enfim, pode-se inferir que na representação temática da informação, a linguagem de indexação objetiva estabelecer a padronização dos termos que irão descrever o conteúdo de determinado documento. A tarefa de definir assunto de um documento exige análise conceitual do documento.

A linguagem natural, de acordo com Lopes (2002), pode ser entendida como sinônimo de discurso comum, ou seja, a forma como a linguagem é usada habitualmente na fala e na escrita. Nas bases de dados, os termos do título e resumo representam esta linguagem.

Trata-se de uma linguagem não controlada, que consiste em utilizar a forma encontrada no próprio documento, fator que proporciona maior agilidade na indexação. Entretanto, a falta de uma padronização dos termos certamente afetará o resultado da busca pela informação exata.

Este tipo de linguagem abrange, sobretudo as terminologias atuais, que ainda não foram incluídas em linguagens controladas.

Os índices são produtos que se originam da indexação, com o objetivo de representar o conteúdo de documentos primários, facilitando o acesso ao documento indexado. Os índices podem estar presentes em bases de dados impressas ou eletrônicas, possibilitando pesquisas por índice, por operadores booleanos e outros.

Os tipos de índices existentes são: KWIK (referente à autoria e palavra-chave), índice alfabético, analítico, assindético, coordenado ou correlativo, corrente, cumulativo, de citações, KWAC (obtido pelas palavras de um título), KWIT (baseado na permutação selecionada), KWOC (palavras significativas em ordem alfabética seguidas pelo título completo), KWOT (palavras significativas são extraídas e colocadas em ordem alfabética), KWUC (arranjo feito pelas classes gerais da CDU e subdivisões dessas classes em ordem alfabética das palavras-chave dos títulos que constam em cada classe) (FEITOSA, 2006, p. 24).

O resumo, segundo Feitosa (2006, p. 21), “consiste na condensação do documento em linguagem natural”, sendo que seu principal objetivo está em sintetizar o conteúdo presente no documento, podendo ser breve ou detalhado, conforme a necessidade.

A linguagem é imprescindível para que o indivíduo se relacione com outros indivíduos e que comunique seus pensamentos e conhecimentos. O fato de nos comunicarmos e a forma como o fazemos têm por objetivo demarcar, significar e comunicar. Para que o conhecimento possa ser compartilhado, faz-se necessária uma forma de representação, denominada signo, objeto de estudo da Semiologia caracterizado por significante e significado. “Os signos da linguagem verbal têm uma importância tão grande para a humanidade que mereceram uma ciência só para si: a Lingüística” (PULCINELLI, 2005, p. 11).

Enquanto representação do pensamento, a linguagem é estudada por pensadores que procuram defini-la como princípios racionais e lógicos do ser humano. “O alvo que esses estudiosos querem atingir é a língua ideal – língua

universal, lógica, sem equívocos, sem ambigüidades, capaz de assegurar a unidade da comunicação do gênero humano” (PULCINELLI, 2005, p. 12).

Pode-se dizer que por “palavra” entendemos o processo concreto da comunicação inter-subjetiva com a ajuda de sinais sonoros. Por “linguagem” entendemos, em contrapartida, o sistema de regras semânticas e de regras gramaticais, abstraído do processo real da palavra. “[...] a palavra é a atualização da linguagem, e a linguagem é, pelo contrário, o abstraído do fenômeno geral da palavra – a sua potencialidade” (SCHAFF, 1964, p. 146).

A Ciência da Informação utiliza-se de linguagens que possibilitam o controle dos termos a serem indexados em um sistema de informação, minimizando a ocorrência do uso da linguagem natural que “pode ser definida como a linguagem do discurso técnico-científico, e, no contexto da recuperação da informação, o termo livre. Nas bases de dados, os campos de título e resumo registram os termos da LN” (LOPES, 2002, p. 2).

Dessa forma, a literatura da área de ciência da informação aborda que há necessidade de utilizar terminologias que tenham por objetivo o “controle” dos termos presentes em um documento.

A utilização de um vocabulário controlado caracteriza-se pelo acesso a uma lista de termos autorizados que o profissional deve utilizar no momento da indexação. O vocabulário controlado possui uma estrutura semântica destinada a controlar sinônimos, diferenciar homógrafos e para ligar termos similares. Os principais tipos são os esquemas de classificação (como a CDD e CDU), listas de cabeçalhos de assunto e tesouros (LANCASTER, 2004, p. 19).

Instrumentos como tesouros e cabeçalhos de assunto representam a linguagem alfabética, enquanto a linguagem simbólica é apresentada por códigos como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU) que representam o documento por meio de letras e números.

A Linguagem documentária visa ao controle terminológico, que permite a representação dos conceitos derivados da indexação sem ambigüidades, garantindo o emprego dos mesmos termos para expressar os mesmos conceitos. Dessa forma, pode-se afirmar que a linguagem documentária tem como característica a oposição à linguagem natural, com a finalidade de padronização visando facilitar a busca à informação.

A função da linguagem documentária, segundo Cintra e colaboradoras (2002, p. 16), é realizar o tratamento do conhecimento registrado, disponibilizando-os como informação, “em outras palavras, compete às linguagens documentárias transformar estoques de conhecimento em informações adequadas aos diferentes segmentos sociais”.

Desta forma, o ato de disponibilizar a produção intelectual em um sistema de informação não é suficiente para que este conhecimento possa ser compartilhado. Há necessidade de critérios de organização que visem preparar estes dados tendo em vista sua recuperação e a sua disseminação. Dessa forma, são definidas as políticas de indexação, para que haja condições de interoperabilidade entre os sistemas de recuperação da informação, levando-se em conta não somente a interoperabilidade técnica, mas, também, a interoperabilidade semântica.

Sayão e Marcondes (2008, p. 137) argumentam que a interoperabilidade pode ocorrer com a utilização de “ferramentas comuns e/ou mapeáveis de representação da informação, como esquema de metadados, classificações, tesouros e mais recentemente, ontologias”. Para tanto, é necessário o estabelecimento de políticas de indexação que possam garantir a qualidade desse trabalho. Importante mencionar que “a política de indexação não pode ser vista como uma lista de procedimentos a serem seguidos, mas sim uma filosofia a ser adotada pelo sistema de recuperação de informação” (FUJITA; RUBI, 2006, p. 49).

Além dos interesses institucionais e profissionais, ao definir uma política de indexação em uma instituição, deve-se levar em conta a necessidade dos usuários. Cabe ao profissional da informação definir os critérios de trabalho que deverão ser estabelecidos à exaustividade ou seletividade dos termos que representarão a qualidade da indexação de um documento em um sistema, em especial nas bibliotecas digitais e repositórios, ambientes onde os dados estão disponíveis e interoperáveis.

A linguagem controlada possibilita a elaboração de ferramentas e/ou produtos. São eles: os tesouros, as taxonomias e as ontologias.

Os tesouros organizam os conceitos, por meio de uma estrutura de termos de uma área do conhecimento semelhantes entre si, podendo apresentar-se de forma alfabética ou sistemática. A utilização de um tesouro em um sistema de informação é fundamental, tendo em vista a possibilidade que oferece ao usuário de

recuperar e ter a condição de escolher dentre os termos, qual o mais adequado à sua necessidade.

Foskett (1973, p. 41) define tesauro como:

Uma lista de termos em que se indica a sua classificação de acordo com as idéias que eles representam [...] esse tipo de lista só pode ser compilado através da indexação de conceitos, podendo se constituir na primeira etapa do desenvolvimento de uma linguagem de indexação completamente artificial.

O tesauro, na perspectiva de Gil Urdiciain (2004, p. 183), é “um vocabulário controlado e estruturado que é composto mediante a seleção de termos da linguagem natural”. Assim, pode-se dizer que é uma lista composta de descritores que, de forma estruturada, por relações hierárquicas, associativas e de equivalência, permitem o relacionamento com outras palavras também existentes no tesauro, ou seja, os conceitos. Para Campos (2006, p. 1) “conceito é, de fato, o ponto de partida para estabelecer as relações conceituais e determinar a forma verbal mais adequada para representá-lo”.

A taxonomia é utilizada para realizar um estudo teórico de uma classificação incluindo bases, princípios, procedimentos e regras.

Currás (2005, p. 53) a define como:

[...] uma aplicação *web* para gestão de linguagens documentárias orientada especificamente para o desenvolvimento de tesouros hierárquicos. Também podem ser utilizadas para desenvolver estruturas de navegação *web* ou articular um sistema de gerenciamento de conteúdo.

Taxonomias e tesouros são semelhantes na medida em que sistematizam o conhecimento segundo normas pré-estabelecidas, são sistemas pré-coordenados e pós-coordenados de linguagens documentárias e sistemas de indexação e recuperação da informação. Quanto às diferenças existentes entre eles, a mais visível é o uso de tecnologia da informação para o desenvolvimento de taxonomias. Desta forma, os profissionais da área de tecnologia utilizam-se mais das taxonomias, enquanto os profissionais da informação, dos tesouros (CURRÁS, 2005, p. 65).

A Ontologia procura suprir as necessidades relacionadas às questões semânticas e lexicais próprias da língua, fator que torna ainda mais difícil estruturar as informações em um sistema de informação. Assim, pode-se dizer que o

papel central da Ontologia é oferecer condições para que a interoperabilidade semântica de sistemas distribuídos ocorra, sobretudo, por ser possível afirmar que a *internet* revolucionou o acesso à informação, no entanto a máquina não dispõe de capacidade para processar o significado de uma informação, há sempre a necessidade de um ser humano para interpretá-la. Segundo Sayão e Marcondes (2008, p. 141) “as ontologias estabelecem uma compreensão compartilhada de um domínio de interesse para apoiar a comunicação entre seres humanos e agentes computacionais”.

É possível dizer que se trata de uma forma de linguagem de representação que pode ser “entendida” pelo computador, permitindo que indivíduos e máquinas atuem de forma cooperativa. Segundo Campos (2007, p. 118) “ontologia é o conjunto de conceitos padronizados onde termos e definições devem ser aceitos por uma comunidade no âmbito de um domínio, e tem por finalidade permitir que múltiplos agentes compartilhem conhecimento”.

O termo ontologia é utilizado no ambiente digital com funções similares às de classificação e indexação. Dias e Naves (2007, p.17) o definem como “[...] o trabalho de organização dos recursos eletrônicos com base em seus conteúdos, de forma a possibilitar a posterior recuperação dessas informações”.

Alvarenga (2001, p. 15) salienta que, segundo a teoria do conceito abordada anteriormente, aos seres concretos e abstratos correspondem os conceitos compostos de enunciados sobre esses mesmos seres e a esses conceitos são atribuídos seus respectivos nomes ou símbolos que, nos documentos textuais mais convencionais, acham-se representados por termos. Daí as relações entre ontologias, conceito e terminologia.

Finalizando a abordagem, cumpre salientar que a grande diferença existente entre a classificação e a catalogação é que a classificação restringe o documento a uma única localização física, enquanto que a catalogação e a indexação possibilitam determinar inúmeras entradas de assunto, seja em um catálogo ou em um índice, obedecendo apenas aos limites impostos por políticas de indexação existentes (DIAS; NAVES, 2007).

Presume-se que toda esta abordagem vise à recuperação e ao compartilhamento da informação, tendo em vista que todo o documento que recebe tratamento e armazenagem em um sistema de informação, certamente terá visibilidade e possibilidade de compartilhamento maximizada, devido aos

procedimentos adotados para sua indexação, objetivando melhor recuperação das informações.

Dias (2006, p. 70) destaca que:

Do ponto de vista temático à informação, pelo menos em face da tecnologia atualmente disponível, a busca eficiente de informação em acervos digitais não dispensa o auxílio de um sistema de recuperação da informação.

O autor supracitado relata que os mecanismos de busca são muito parecidos com os sistemas de recuperação da informação e ressalta que a combinação entre linguagem controlada e natural é o ideal para sistemas de recuperação da informação, visto que o uso apenas da linguagem natural em um sistema de informação inviabiliza a busca devido à ocorrência de situações de sinonímias, neologismo, polissemia, hiponímias e metáforas. Cintra e colaboradoras (2002, p. 71) acrescentam que: “Numa linguagem documentária, tanto a polissemia, quanto a ambigüidade devem ser neutralizadas, para que seja garantida a monossemia entre a forma do significante e a do significado”.

Os sistemas de recuperação da informação compõem-se pela entrada de dados, processamento e recuperação visando fornecer acesso às informações relevantes eliminando aquelas excedentes ao tema buscado. Os sistemas de recuperação da informação possuem a função de organizar e facilitar a localização de documentos. Segundo Dias (2006, p. 65) a função desses sistemas consiste em “[...] selecionar e adquirir os materiais de informação; fazer o tratamento, organização e armazenagem desse material; a função de dialogar com o usuário em torno de uma consulta para melhor defini-la, entre outras funções”.

Quanto aos modelos de recuperação da informação, destaca-se o *booleano*, modelo que opera com os conectores *e*, *ou* e *não*, por ser o mais utilizado em mecanismos de busca na *web*. No entanto, outros importantes modelos são também utilizados, dentre eles pode-se citar o modelo vetorial, o probabilístico e o de operadores *fuzzy* (FEITOSA, 2006, p. 32).

Piedade (1983, p. 10) destaca que “na avaliação de eficiência de um serviço de recuperação da informação, duas medidas estão diretamente ligadas aos seus índices, o coeficiente de revocação e o coeficiente de precisão”.

A revocação diz respeito à recuperação de itens pertinentes em uma coleção de documentos. Já a medida com que os itens relevantes são recuperados

denomina-se coeficiente de revocação. O coeficiente de precisão leva em consideração a exaustividade de elementos utilizados na indexação e a especificidade representada pelos descritores do documento solicitado. Desta forma, Lancaster (2004, p.189) menciona que “frequentemente se emprega um coeficiente de precisão junto com o coeficiente de revocação para se ter uma indicação do grau de eficiência ou discriminação atingido por uma consulta”.

6 PROCEDIMENTOS METODOLÓGICOS

De acordo com Gil (2008, p. 28), pode-se afirmar que quanto aos objetivos, este estudo é classificado como uma Pesquisa Descritiva, que visa a “descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis”. Segundo o autor, trata-se de um estudo que é indicado aos pesquisadores preocupados com a atuação prática.

Seu enfoque está na análise qualitativa por buscar observar o fenômeno estudado, destacando as características mais apuradas por meio de uma análise profunda (RAUPP; BEUREN, 2003). A pesquisa qualitativa possibilita que o pesquisador interprete os dados analisados podendo ou não incluir suas interpretações pessoais no momento da análise.

De acordo com Wolcott (1994) *apud* Creswell (2007, p. 186), a pesquisa qualitativa

[...] é fundamentalmente interpretativa. Isso significa que o pesquisador faz uma interpretação dos dados. Isso inclui o desenvolvimento da descrição de uma pessoa ou de um cenário, análise de dados para identificar temas ou categorias e, finalmente, fazer uma interpretação ou tirar conclusões sobre seu significado, pessoal e teoricamente, mencionando as lições aprendidas e oferecendo mais perguntas a serem feitas.

A fim de propor um delineamento para “o desenvolvimento da pesquisa, com ênfase nos procedimentos técnicos de coleta e análise de dados” (GIL, 2002, p. 43), torna-se possível classificá-la sob dois aspectos: como Pesquisa Bibliográfica, no que tange ao levantamento de fontes de informação que asseguram o conteúdo e revisão bibliográfica deste tema e também como Estudo de Caso, visto que consiste em pesquisar casos reais, definindo e esclarecendo seus limites, verificando o problema de forma integral (GIL, 2007).

De acordo com Martins (2006), esse tipo de pesquisa investiga um fenômeno real, no qual o pesquisador não tem controle sobre as ocorrências. Nesse sentido, o fato de compreender, descrever e interpretar o “caso” permite que o pesquisador analise o contexto real do objeto de pesquisa.

Para Chizzotti (2003, p. 102):

[...] o estudo de caso é uma caracterização abrangente para designar uma diversidade de pesquisas que coletam e registram dados de um caso particular ou de vários casos a fim de organizar um relatório ordenado e crítico de uma experiência, ou avaliá-la analiticamente, objetivando tomar decisões a seu respeito ou propor uma ação transformadora.

A escolha desse método pode ser defendida por Yin (2005, p. 32) ao afirmar que “[...] usaria o método de estudo de caso quando deliberadamente quisesse lidar com condições contextuais – acreditando que elas poderiam ser altamente pertinentes ao seu fenômeno de estudo”.

Quanto ao delineamento, optou-se pelo estudo de caso múltiplo por promover a ampliação de conhecimentos sobre o universo selecionado por meio do estudo de características das populações. Analisar múltiplos casos possibilitará a identificação de diferentes contextos, identificando fatores convergentes e divergentes entre eles, trazendo maior qualidade para os resultados deste estudo (GIL, 2007, p. 142).

Utilizando-se de múltiplos casos, faz-se necessário prever que as unidades selecionadas para o estudo possuam semelhanças, para que os resultados possam ser os almejados. O pesquisador deve, ainda, apoiar-se nos resultados obtidos e não apenas em deduções (VENTURA, 2007, p. 386).

A fim de delinear a abordagem pertinente à definição dos instrumentos a serem adotados com o intuito de verificar a aplicabilidade deste estudo, optou-se por identificar e analisar as práticas de outras instituições de ensino, cujos repositórios institucionais já se encontram implantados.

Inicialmente, realizou-se uma pesquisa nos diretórios *ROAR* e *OpenDoar* (no período de 12 a 13 de abril de 2010), período em que foram identificadas as instituições que fariam parte desta análise. Posteriormente, entre 14 de abril a 16 de maio de 2010, ocorreu a coleta de dados, baseada no instrumento Formulário; e no período de 08 a 10 de junho do mesmo ano, ocorreram as entrevistas com os gestores, que agregaram as informações necessárias à análise.

6.1 INSTRUMENTOS DE COLETA DE DADOS

Quanto aos procedimentos de coleta de dados, identificou-se que na busca por maior entendimento acerca do problema de pesquisa, a análise qualitativa permitiu selecionar os participantes mais indicados para responder a questão de partida. “Isso não sugere necessariamente amostragem aleatória ou seleção de um grande número de participantes e locais, como geralmente vemos na pesquisa quantitativa” (CRESWELL, 2007, p. 190).

De acordo com Barros e Lehfeld (2006, p. 89), “a coleta de dados significa a fase da pesquisa em que se indaga e se obtém dados da realidade pela aplicação de técnicas utilizando para isso de instrumentos de coleta de dados”. Neste caso, optou-se por associar mais de uma técnica para atingir o objetivo desejado, fator que poderá atribuir ainda mais qualidade aos resultados obtidos. Neste estudo, optou-se por dois instrumentos: o formulário e a entrevista.

Por meio do formulário, buscou-se identificar os elementos disponíveis no repositório selecionado, enquanto a entrevista permitiu maior contato entre o entrevistado e entrevistador, agregando informações pertinentes e não explícitas na estrutura do repositório. Esta técnica tem por objetivo compreender o significado que os entrevistados atribuem a determinadas “questões e situações, em contextos que não foram estruturados anteriormente, com base nas suposições e conjecturas do pesquisador” (MARTINS, 2006, p. 27).

Quanto à operacionalização deste instrumento, optou-se pela entrevista estruturada, na qual as questões são previamente estabelecidas. Este instrumento possibilitou maior contato entre entrevistador e entrevistado, tornando possível identificar “dados que possam ser utilizados em análise qualitativa, ou seja, os aspectos considerados mais relevantes de um problema de pesquisa” (BARROS; LEHFELD, 2006, p. 91).

As maiores vantagens na aplicação da entrevista como instrumento de coleta de dados está na facilidade que o entrevistador possui em formular e reformular a questão caso seja necessário, visando a um melhor entendimento. Ainda, o entrevistador poderá observar as práticas do profissional entrevistado, enriquecendo assim sua análise.

Entretanto, o entrevistador deve ter ciência de que este instrumento irá depender maior tempo na análise. Assim, a fim de facilitar o acesso aos entrevistados, infere-se que é possível coletar dados da seguinte forma: por meio de anotações, gravação, internet e telefone. A entrevista por telefone “é útil para complementar questões pendentes, tirar dúvidas ou obter informações simples, que ajudem a completar o quadro em análise”. Os benefícios obtidos por este tipo de entrevista consistem na agilidade do acesso aos entrevistados que se encontram distantes geograficamente (DUARTE; BARROS, 2009, p. 77).

A entrevista por telefone diminui custos, agiliza o procedimento, possibilita que o agendamento ocorra em momento apropriado para o entrevistado, no entanto, não possibilita que o entrevistador visualize as condições e as características do ambiente onde se encontra o entrevistado. De acordo com Gil (2008, p. 113) “mais recentemente, porém, as entrevistas por telefone passaram a ser mais aceitas como procedimento adequado para pesquisa em ciências sociais”.

A validação dos dados obtidos em uma pesquisa qualitativa poderá ocorrer por meio de critérios de fiabilidade (independência das análises meramente ideológicas do autor), credibilidade (garantia de qualidade relacionada a exatidão e quantidade de observações efetuadas), constância interna (independência dos dados em relação a acidentalidade, ocasionalidade, etc) e transferibilidade (possibilidade de estender as conclusões a outros contextos) (CHIZZOTTI, 2003).

6.2 UNIVERSO DA PESQUISA

Este tópico visa representar o universo da pesquisa, sua extensão e as formas de seleção dos sujeitos, principalmente pela falta de possibilidade em “obter as informações de todos os indivíduos ou elementos que compõem o universo ou população que se deseja estudar” (BARROS; LEHFELD, 2006, p. 86).

A maioria dos repositórios existentes encontram-se registrados nos diretórios OpenDoar e ROAR, assim, a extensão da amostra é representada neste estudo, pelo total de instituições acadêmicas que possuem repositórios institucionais neles registrados.

Ao selecionar os repositórios, buscou-se obter informações relevantes de natureza de estruturação e organização dos documentos ali contidos, visando identificar a aplicação de políticas de informação nas instituições as quais se encontram vinculadas, especificamente no que tange as suas estruturas de organização da informação.

Assim, a seleção dos elementos deu-se de forma não-probabilística que de acordo com Barros e Lehfeld (2006, p. 88) “são compostas muitas vezes de forma acidental ou intencional”. Neste caso, ocorreu a amostra intencional, cujas estratégias de seleção, para fins deste estudo, consistiram nas seguintes: buscou-se selecionar os repositórios institucionais nacionais vinculados a instituições acadêmicas de ensino superior, cuja tipologia documentária depositada contemplasse variados documentos e que, o *software* utilizado fosse o DSpace.

Em 12 de abril de 2010, realizou-se o levantamento no diretório *OpenDoar* do qual obteve-se o resultado de cinco repositórios conforme seguem: 1) Biblioteca Digital da Ação Educativa, 2) Biblioteca Digital de Monografias (BDM) da Universidade de Brasília, 3) Brasiliana da Universidade de São Paulo, 4) DSpace da UFPR - Universidade Federal do Paraná e 5) Repositório Institucional da Universidade de Brasília (RIUnB).

Analisando cada um destes repositórios, optou-se por eliminar da análise dois deles: a Biblioteca Digital Ação Educativa por tratar-se de um repositório temático na área de Direito e por não pertencer a uma instituição acadêmica e o Repositório Brasiliana da USP, por sua missão de reunir documentos provenientes de outras instituições, não possuindo a característica de um repositório institucional. Desta forma, o diretório *OpenDoar* forneceu a este estudo a seleção de três repositórios: BDM UNB, DSpace UFPR e RIUnB da UNB.

Na sequência, realizou-se o levantamento no *ROAR*, visando identificar os repositórios com as mesmas características selecionadas no primeiro diretório. Diferente do mecanismo de busca do *OpenDoar* que seleciona mais de um item e permite seu cruzamento, o *ROAR* permite apenas busca simples. Assim, optou-se por selecionar os repositórios brasileiros; fator que resultou em setenta repositórios nacionais cadastrados. Para refinar a busca, realizou-se a seleção por tipo de repositório (institucional, especificamente os vinculados a universidades) e a seleção do *software* (DSpace) ocorreu por meio da análise nos sites de cada um dos setenta repositórios registrados neste diretório.

Constatou-se que muitos dos repositórios registrados são periódicos e/ou bibliotecas digitais de teses e dissertações. Importante mencionar que, neste estudo, as bibliotecas digitais de teses e dissertações não foram selecionadas para análise, defendendo o posicionamento de que os repositórios institucionais permitem a existência de uma variada tipologia documentária, ou seja, seu objetivo é o de reunir, organizar e disseminar toda a produção intelectual institucional, independente da tipologia em que se apresenta.

Além destes fatores, Leite (2009, p. 35) demonstra que um repositório institucional diferencia-se de uma biblioteca digital por gerenciar exclusivamente o conteúdo produzido na própria instituição, por possuir padrões de interoperabilidade e devido a seu *software* ser desenhado em função dos sistemas de comunicação científica, fatores que necessariamente não são encontrados nas bibliotecas digitais.

Analisando estes apontamentos, foram considerados como objetos de análise, o Lume Repositório da UFRGS, o DSpace da UFPR e o RIUnB Repositório Institucional da UNB.

6.3 POPULAÇÃO ALVO

Após selecionar os repositórios que possuem características de repositórios institucionais, foi possível definir o total de elementos que constituíram o *corpus* de análise desta pesquisa: o Repositório RIUnB da Universidade de Brasília, Repositório LUME da Universidade Federal do Rio Grande do Sul e o Repositório DSpace da Universidade Federal do Paraná. Acredita-se que a população aqui selecionada possui condições de apresentar resultados que atendam às necessidades deste estudo em termos de qualidade de suas estruturas, proporcionando evidências de contextos diferenciados e possibilitando a elaboração de uma pesquisa de melhor qualidade.

7 ANÁLISE E APRESENTAÇÃO DOS RESULTADOS

Este capítulo apresentará a análise dos resultados sistematizada em quatro categorias: Informações gerais sobre os repositórios; Estruturas de organização da informação; Metadados; e Interoperabilidade (semântica e humana).

A combinação de dois instrumentos de pesquisa foram fatores positivos a esta análise, visto que possibilitaram a identificação e coleta inicial dos dados referentes às estruturas de organização dos repositórios por meio do formulário e o esclarecimento de questões ali não explícitas, pela entrevista. Devido à relevância dos dados obtidos, julgou-se necessário apresentar alguns trechos dos relatos verbais, que estão destacados no texto em itálico.

7.1 CATEGORIAS DE ANÁLISE

7.1.1 Informações Gerais Sobre os Repositórios Dspace UFPR, Lume e RIUnB

A análise do DSpace UFPR, Repositório Institucional da Universidade Federal do Paraná⁴, foi realizada em 16 de maio de 2010 e buscou, além de identificar as estruturas de organização da informação, verificar a existência de políticas e/ou diretrizes que pudessem definir os critérios de organização do repositório. No entanto, no período de análise, não foi possível observar nenhuma documentação disponível. Este possui atualmente o total de cinco mil, oitocentos e cinquenta e sete itens depositados e é composto pelas seguintes tipologias documentárias: Livros, Teses e Dissertações, Relatórios técnicos, Vídeos, Imagens e Sons.

O aceite do Comitê Gestor em colaborar com esta pesquisa, compartilhando os conhecimentos quanto aos trabalhos realizados na implementação do software DSpace na UFPR, ocorreu por meio da Diretoria do

⁴ Disponível em: <<http://DSpace.c3sl.ufpr.br/DSpace/community-list>>.

Sistema de Bibliotecas da instituição. Este contato possibilitou a contextualização dos objetivos deste estudo, a troca de informações a respeito da estrutura de informação da instituição em estudo, bem como a obtenção de alguns elementos não identificados anteriormente, na análise do *site*, pelo Formulário.

Inicialmente, buscou-se identificar o nível de participação do respondente no repositório, a resposta obtida indicou que sua participação é de Coordenadora do repositório, mas que esta atribuição é dividida entre ela e um docente do Centro de Computação Científica e Software Livre (C3SL), que responde pela parte de tecnologia em parceria com seus bolsistas, enquanto a bibliotecária e sua equipe responsabilizam-se pela definição do esquema de organização da informação e pelos critérios de metadados a serem adotados.

O trabalho com o repositório DSpace UFPR, teve início em 2004, ano em que ocorreram os primeiros testes, a customização de sua interface e os depósitos dos primeiros documentos. No ano de 2005 ocorreu sua disponibilização para o público em geral, assim, pode-se dizer que oficialmente, a UFPR tem trabalhado com o DSpace há seis anos.

Quanto aos objetivos que levaram à sua implantação, identificou-se que o principal deles foi *“proporcionar visibilidade da produção científica e intelectual da UFPR”*. De Este projeto viabilizou o acesso aberto, reunindo a implantação e gestão da Biblioteca Digital da UFPR, as revistas, bem como vídeos da UFPRTV e a integração dos sistemas da Biblioteca Digital ao Catálogo geral do Sistema de Bibliotecas, com busca unificada, detecção de coerência, consistência e sincronia entre as bases⁵.

De acordo com a entrevistada, a implantação do repositório veio ao encontro da necessidade de ampliar a visibilidade de produção científica e intelectual da UFPR. *“Este projeto partiu de reuniões realizadas no Centro de Computação Científica e Software Livre, que geraram a expectativa e necessidade da criação de um repositório que pudesse disponibilizar toda a riqueza intelectual produzida na instituição. Partindo desta iniciativa, esta demanda foi encaminhada ao Sistema de Bibliotecas, tendo em vista que quem guarda esta demanda é a Biblioteca. Só foi alterado o suporte. A necessidade de tratamento e armazenagem continua a mesma, e continua sendo trabalho da Biblioteca”*.

⁵ Disponível em: <<http://www.ufpr.br/adm/templates/index.php?template=1&Cod=2628>>.

Uma das principais decisões tomadas pela equipe que deu início ao repositório foi a definição do *software* que iria gerenciá-lo. Optou-se por um *software* livre, para que a instituição tivesse maiores condições de alterá-lo e customizá-lo sem que houvesse a necessidade de dependência de sistemas proprietários. De acordo com a entrevistada, o E-prints, Fedora e Kepler também foram testados, entretanto, a avaliação dos docentes e dos bolsistas do C3SL, que realizaram os testes, considerou o DSpace o mais adequado para atender às necessidades da UFPR.

Quanto à definição de um Comitê Gestor, constata-se a existência desta equipe desde a fase inicial do repositório, porém, informalmente. Conforme mencionado anteriormente, esta equipe é composta por profissionais da área de tecnologia, docentes, bolsistas e bibliotecários que dividem, além da responsabilidade, as tarefas no repositório, sendo que a coleta das teses e dissertações competem ao Sistema de Bibliotecas e os vídeos, à UFPRTV e ao C3SL.

Buscando definir as atribuições desta equipe, de acordo com a gestora, compete a eles “*manter a casa em ordem, administrar o banco de dados, alimentar a biblioteca digital, ou seja, manter a Biblioteca Digital atualizada e no ar ininterruptamente para garantir a visibilidade ininterrupta da produção científica e intelectual da UFPR*”. Sendo que compete especificamente aos bibliotecários, “*a atribuição dos metadados, o controle de autoridades e a submissão dos itens, porque no nosso repositório não há autodepósito*”.

Quanto à existência de políticas de informação, constatou-se que a UFPR não as possui de maneira formalizada. No entanto, as atribuições que dizem respeito ao repositório são especificadas institucionalmente, mas não ocorre de forma explícita. A política ainda está em construção, tendo em vista que não se trata de algo estático, renova-se constantemente. Neste aspecto, a definição das comunidades, de acordo com a explicação da respondente “*são criadas pela demanda, e as demandas são crescentes*”.

No que tange aos procedimentos adotados pela instituição para definir quais as tipologias integrarão o repositório, bem como quais publicações poderão fazer parte dele – se somente as publicações geradas na instituição ou se produções de servidores da instituição em outras universidades constatou-se que “*Isto está definido na política do Sistema de Bibliotecas, essa atribuição cabe às*

bibliotecas”. Neste caso, vale mencionar que a UFPR opta por disponibilizar tanto a produção gerada em seu âmbito, quanto em outras instituições, desde que os autores sejam da comunidade acadêmica da UFPR.

Conduzindo esta análise para as questões mais específicas da área, buscou-se conhecer a estrutura do Sistema de Bibliotecas da UFPR. Pode-se mencionar que esta instituição possui uma sede administrativa (Biblioteca Central), treze bibliotecas universitárias e uma de ensino médio. Em Curitiba estão localizadas doze delas e as demais se encontram nos municípios de Palotina, Pontal do Paraná e Matinhos. Este Sistema de Bibliotecas possui cinquenta e nove bibliotecários, oitenta assistentes (oito deles possuem formação em Biblioteconomia) e vinte e cinco bolsistas sênior (sete deles possuem formação em Biblioteconomia). A Diretora do Sistema de Bibliotecas responde tanto pelo SiBi quanto pelo repositório.

Identificou-se que o sistema gerenciador de informações da biblioteca da instituição é o *software* Sophia, sendo importante mencionar a existência da interconexão dos dados nele depositados com o repositório.

Ao indagar a questão do acesso aberto, buscou-se identificar se as publicações depositadas no repositório atendem a esta filosofia. Constatou-se que todos os documentos existentes no repositório estão em acesso aberto. Não há documentos em situação de embargo, ou seja, não possui depósitos de documentos com permissão de visualização apenas de seus metadados e não do texto completo. Esta situação ocorre geralmente por solicitação do autor e/ou instituição, devido a questões de direitos autorais, patentes e outros. No entanto, na UFPR, tudo o que está disponível no repositório encontra-se disponível em texto completo.

Quanto ao respaldo aos autores, a UFPR optou por trabalhar de forma inversa aos modelos tradicionais, ou seja, só preenchem documentação aqueles usuários que não irão depositar seu trabalho, ou seja, somente as exceções (como ocorre no caso de patentes). De acordo com a entrevistada *“Quando uma pessoa entrega o seu arquivo ela já aceita a sua inclusão no repositório. No caso de patente ou outro item que tenha algum problema que impeça a sua difusão e disseminação, o coordenador do programa de pós-graduação e o autor é que nos solicitam (solicitam à biblioteca) por escrito que não será realizado o depósito, evita-se desta forma o acúmulo de papel desnecessário”*.

Esta medida está respaldada no ofício 101634N3 – PRPPG, datado de 27 de junho de 2004 e na Resolução CEPE nº 65/09 (o depósito legal é tratado no Artigo 67 §2º e Artigo 81 item III) que estabelece normas gerais únicas para os cursos de pós-graduação *stricto sensu* (mestrado e doutorado) da UFPR. O ofício citado é destinado a todos os coordenadores de curso, indicando que a partir da criação do repositório toda a produção científica e intelectual produzida na UFPR ou por seus membros, obrigatoriamente, deve ser depositada em papel e meio digital.

Desta forma, caso haja exceções, estas devem ser analisadas. Assim, no caso das publicações científicas, é necessária a entrega do documento impresso que será disponibilizado na biblioteca e também do documento no formato digital, que irá ser armazenado no repositório. E a resolução diz que “*Será exigido o seguinte número de exemplares (impressos e digitais) definitivos: um para a Biblioteca Central, um para a Biblioteca Setorial e um exemplar para cada membro de banca examinadora, incluindo os suplentes*”.

O segundo repositório analisado, o LUME⁶, pertence à Universidade Federal do Rio Grande do Sul, possui esta denominação por representar a manifestação de conhecimento, saber, luz, brilho e trata-se do portal de acesso às coleções digitais de documentos gerados no âmbito da universidade e de outros documentos que, por interesse da instituição sejam passíveis de armazenagem neste ambiente, visando a sua preservação e disseminação.

O Lume utiliza-se do software DSpace, customizado pela interface XML Manakin, compatível com o protocolo OAI. O padrão de metadados utilizado é o Dublin Core e o sistema de identificador persistente utilizado é o CNRI Handle. De acordo com as informações divulgadas na página inicial do repositório, esta ferramenta consiste em reunir, preservar, divulgar e garantir acesso a documentos acadêmicos, científicos, artísticos e administrativos gerados na universidade. Ampliando seu escopo, busca reunir coleções históricas e demais documentos que embora não produzidos na UFRGS, são considerados relevantes a sua comunidade.

Embora não denominadas como “políticas”, constatou-se a existência de regulamentações que disponibilizam os ofícios e portarias que respaldam as ações desde a criação da Biblioteca Digital até a criação do comitê gestor do repositório; bem como, regulamentações que fornecem as instruções para submissão de documentos, até mesmo a resolução que determina a obrigatoriedade

⁶ Disponível em: <<http://www.lume.ufrgs.br>>.

de depósito de teses e dissertações. Este fator pode conferir permanência e preservação de sua estrutura, tendo em vista a existência de documentos que definem as ações inerentes ao seu funcionamento, bem como indicam o fluxo dos documentos dentro da instituição. Disponibiliza também informação aos autores. Acredita-se que as dificuldades encontradas no momento do depósito do documento possam ser minimizadas com a leitura destes documentos. Proporciona ainda, mecanismos de relacionamento entre o usuário e a gerência do repositório, de forma que seja possível oferecer esclarecimentos quando necessário.

O Lume disponibiliza as seguintes tipologias documentárias: Fotografias, Artigos de Periódicos, Teses e Dissertações, Trabalhos de Conclusão de Curso (Especialização), Trabalho de Conclusão de curso (Graduação) e Trabalhos em Eventos, em formato texto, imagem, vídeo e áudio, e em sua maioria em acesso livre.

A entrevista com a integrante da comissão gestora identificou que a sua atribuição no repositório é de integrante da equipe técnica/operacional, na função de bibliotecária, que em conjunto com profissionais de tecnologia da informação mantém os documentos digitais em acesso aberto.

Identificou-se também que a criação do Lume ocorreu partindo da ampliação da Biblioteca Digital de Teses e Dissertações, existente na instituição desde o ano de 2001. Desta forma, como repositório, sua disponibilização ocorreu em janeiro de 2008, com o objetivo de reunir a produção científica e intelectual da instituição, proporcionando ampla visibilidade.

Questionou-se quanto às etapas de criação do repositório na UFRGS; de acordo com a gestora *“A criação partiu de uma iniciativa do Centro de Processamento de Dados (CPD) e do Sistema de Bibliotecas da UFRGS (SBU), que identificaram a necessidade de ampliar as tipologias de documentos existentes na Biblioteca Digital de Teses e Dissertações, especialmente devido às demandas internas”*.

Esta equipe optou inicialmente por estudar alternativas de *softwares* já existentes, dentre eles pode-se citar o *Nou-rau*. Entretanto, após participarem de um evento, tomaram conhecimento do *DSpace* e constataram que o mesmo estava sendo amplamente utilizado por várias instituições do mundo. Levaram em consideração, sobretudo, que o seu desenvolvimento tenha sido realizado no MIT, fator que os levou a direcionar ações para esta iniciativa. Importante mencionar

alguns esclarecimentos da respondente: *“Desde o início da automação, a UFRGS, optou por utilizar normas e padrões internacionais, visando facilitar o intercâmbio e a migração de dados e padrões entre os sistemas para evitar a perda de informações e re-trabalho”*.

Há um Comitê Gestor institucionalizado para o Lume, desde 20/07/2009, entretanto, segundo a entrevistada *“o Lume foi sendo consolidado gradualmente e, em julho de 2009 a UFRGS instituiu seu comitê gestor por meio de portaria do Reitor”*. Segundo a gestora *“as atividades iniciaram-se pelo Sistema de Bibliotecas que eventualmente visitava a Pró-Reitoria de Pós-Graduação, no intuito de normalizar as questões das teses e dissertações, dos trabalhos de conclusão de curso. Este trabalho visou consolidar inicialmente a ferramenta, ou seja, “o produto” – a fim de que o entendimento sobre sua relevância fosse sendo esclarecido de forma gradativa.*

Constatou-se que a composição da equipe gestora do LUME foi definida em parceria com o reitor, procurando contemplar as grandes áreas do conhecimento, através da nomeação de indivíduos que pudessem trabalhar com as comunidades e produções diferenciadas, objetivando, assim, agregar quem realmente pudesse contribuir. Desta forma, *“docentes das diversas áreas de conhecimento, da secretaria de ensino a distância, analistas de sistemas, bibliotecários, diretora da biblioteca central e do CPD, compõem atualmente o Comitê Gestor do repositório”*.

Definidas as pessoas que fariam a gestão desta ferramenta institucional, buscou-se identificar as principais atribuições deste Comitê. De acordo com a gestora do Lume, inúmeras são as atividades realizadas, dentre as quais pode-se destacar *“a definição da política institucional, a avaliação da implantação de novas comunidades e coleções no repositório, dentre outras”*.

Quanto aos bibliotecários *“o grande papel é a coleta dos documentos de produção intelectual e o processamento técnico dos mesmos (indexação e catalogação)”*. Assim, cabem às bibliotecas do SBU o recebimento e processamento e registro dos metadados no SABi (Sistema de Automação de Bibliotecas), visto que a instituição ainda não adota política de autoarquivamento, principalmente, pela dificuldade em delegar o registro dos metadados para o depositante, uma vez que na prática é difícil que isto funcione neste primeiro momento.

A UFRGS já possui políticas de informação para o Lume, cabe salientar que o comitê definiu uma política geral de informação e que as especificidades seriam tratadas oportunamente. De acordo com a respondente “*a política está pronta, mas ainda não tramitou pelos órgãos competentes*”.

Quanto à política para a inserção de conteúdos no repositório, identificou-se que o Lume optou, desde o início, pela inclusão de documentos produzidos, também, fora do seu âmbito, mas que sejam de interesse institucional. Desta forma, isto se reflete até no seu nome Lume Repositório Digital ao invés de Repositório Institucional.

Buscando identificar a estrutura desta instituição, obteve-se a informação de que a Biblioteca Central da UFRGS é um órgão suplementar da universidade vinculada diretamente à Reitoria e possui a missão de coordenar e supervisionar, de forma sistêmica, as trinta e três bibliotecas da universidade. O SABi utiliza o software Aleph 500 para gerenciar as atividades e serviços oferecidos pelas bibliotecas da UFRGS a sua comunidade usuária. O Lume faz a coleta de metadados no SABi através de um programa desenvolvido para este fim.

Procurou-se verificar se os itens disponíveis no Lume estão na totalidade disponíveis em acesso aberto, ao que se constatou que alguns documentos são de acesso restrito, como ocorre com algumas teses e dissertações. O acesso restrito é atribuído quando não foi possível localizar o autor para solicitar o preenchimento do formulário de autorização e é permitido apenas aos usuários que possuem o cartão de identificação da universidade, de acordo com a respondente “*a procuradoria geral da UFRGS é bastante criteriosa, assim, todas as teses e dissertações depositadas no LUME possuem a autorização do autor, bem como os artigos de periódicos possuem a autorização de seus editores*”.

O terceiro repositório analisado, o RIUnB⁷ é vinculado à Universidade de Brasília com objetivo de ser um conjunto de serviços oferecidos pela Biblioteca Central para a gestão e disseminação da produção científica e acadêmica da Universidade de Brasília.

A análise deste repositório foi realizada entre o período de 14 a 27 de junho de 2010, momento em que possuía o total de quatro mil, quatrocentos e quatro itens depositados, contemplando as seguintes tipologias documentárias: teses, dissertações, artigos de periódicos, anais de eventos, livros e capítulos de

⁷ Disponível em: <<http://repositorio.bce.unb.br/>>.

livros. Buscou, além de identificar as estruturas de organização da informação, observar e analisar a existência de políticas e/ou diretrizes que pudessem definir os seus critérios de organização. No entanto, no período de análise, não havia documentação disponível.

A última entrevista, destinada a obter informações junto à Universidade de Brasília, foi realizada com a servidora que responde pela função de Chefe do Serviço de gerenciamento de informação digital da UnB, que inclui o RIUnB – Repositório Institucional, implantado no ano de 2005. De acordo com ela os objetivos deste repositório são *“Armazenar, preservar, divulgar e garantir acesso à produção científica e acadêmica da Universidade de Brasília em formato digital; proporcionar visibilidade à produção científica da instituição; apoiar as atividades de pesquisa e criação do conhecimento científico; apoiar o processo de ensino-aprendizagem por meio do acesso facilitado ao conhecimento”*.

A fim de iniciar os apontamentos acerca da estrutura do RIUnB solicitou-se à coordenadora do repositório que comentasse a respeito das etapas que permitiram a criação do repositório na UnB. Segundo a entrevistada, a iniciativa surgiu de uma docente do Departamento de Ciência da Informação da Universidade que participou, no ano de 2003, de uma Conferência na Universidade do Minho, em Portugal, onde bibliotecárias do MIT apresentaram um trabalho sobre o desenvolvimento de repositórios institucionais e do software DSpace. Assim, no ano de 2004, este projeto foi apresentado ao Reitor da UnB, que aprovando a iniciativa possibilitou a formação de uma parceria entre o Departamento de Ciência da Informação e Documentação (CID) e a Biblioteca.

Já em 2005, deu-se a institucionalização do projeto com a sua implementação no servidor da BCE com o aval da Reitoria e do Decanato de Pesquisa e Pós-Graduação (DPP) da UnB. Para dar início às atividades, optou-se pela definição de um projeto piloto, sendo escolhido um departamento que apoiou a iniciativa e começou divulgando seus trabalhos científicos. Assim, em 2007, iniciou-se a submissão à comunidade da Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação - FACE.

No ano de 2008, a BCE recebeu a colaboração do Centro de Seleção e de Promoção de Eventos - CESPE/Banco Internacional de Objetos Educacionais para a escolha do conjunto de metadados, padronização de dados e a

concessão de bolsistas. No dia 16 de setembro de 2008, deu-se o lançamento do Repositório Institucional da UnB.

Após conhecer o histórico de implantação deste repositório, questionou-se a respeito da definição de um Comitê Gestor. De acordo com a entrevistada, a UnB não possui um comitê gestor, toda a gestão do repositório é realizada pela Biblioteca Central em parceria com as Faculdades e Institutos da Universidade. *“São delegados professores que ficam responsáveis por fazerem a divulgação do repositório em seus departamentos. Esses professores funcionam como intermediários entre a biblioteca e o professor que irá ter seus documentos depositados”*. De acordo com a gestora apesar de não haver um comitê formalizado, todo o processo de divulgação e parceria com os departamentos, faculdades e institutos da UnB foi feito desde o início do projeto do Repositório, desta forma permitiu a integração de docentes, técnicos em informática e bibliotecários.

Questionou-se a respeito das atribuições delegadas aos bibliotecários da UnB, ao que se constatou que *“todo o gerenciamento do repositório é feito pelos bibliotecários da BCE, desde a coleta dos documentos, passando pela submissão e revisão dos metadados”*.

É possível afirmar a existência de uma política de informação institucional na UnB, entretanto, encontra-se em fase de aprovação pelas instâncias competentes. Quanto a inserção de conteúdos, a gestora relata que se optou por disponibilizar apenas produção institucional *“os conteúdos do Repositório Institucional da Universidade de Brasília devem possuir as características de serem científicos ou academicamente orientados; produzidos, submetidos ou patrocinados pela Universidade de Brasília ou por membros da comunidade acadêmica; devem estar em formato digital; devem estar prontos para serem disseminados amplamente na rede ou, conforme exceções, acessível a níveis determinados; o autor deve estar habilitado a garantir à instituição o direito de preservar e distribuir o trabalho por meio do repositório mediante as condições estabelecidas no termo de autorização do autor”*.

Acrescenta que só são depositados documentos produzidos pelo corpo docente da Universidade, salvo as teses e dissertações. Neste caso, são depositados documentos produzidos dentro ou fora da universidade desde que seja produção de um docente.

A fim de contextualizar o Sistema de Bibliotecas da UnB, constatou-se que o mesmo fora criado para apoiar o ensino, a pesquisa e a extensão, sendo composto pela Biblioteca Central, Biblioteca de Pós-Graduação e pelo Campus Avançado Asa Sul. Possui como *software* gerenciador o sistema Pergamum que permite a interconexão de dados com o repositório, ao que destaca a respondente que *“não é utilizado, pois a Biblioteca não faz indexação de artigos de periódicos. Para os livros esse campo pode ser utilizado. Para as teses e dissertações e colocado o link da BDTD”*.

A fim de esclarecer se todos os itens depositados no repositório estão disponíveis em acesso aberto, identificou-se que o RIUnB possui itens restritos, *“entretanto a maioria absoluta é documentos de acesso aberto. Normalmente os documentos estão restritos a pedido do autor por motivos de direito autoral na maioria dos casos”*.

Apresentados os dados da primeira categoria de análise – que corresponde às informações gerais a respeito de cada um dos repositórios a fim de identificar sua estrutura desde a sua criação até os processos que garantam sua permanência julgou-se necessário apresentar uma síntese das informações obtidas, por meio do quadro a seguir:

REPOSITÓRIO	DSPACE UFPR	LUME	RIUNB
IMPLEMENTAÇÃO	2004 ocorrem os testes e customização e 2005 é lançado oficialmente. Atualmente possui 5.751 itens.	Surge como ampliação da BDTD da UFRGS criada em 03/06/2005, entretanto oficialmente como repositório data de janeiro de 2008. Atualmente possui 22.495 itens	Inicia a mobilização institucional juntamente com o IBICT no ano de 2003, em 2005 implementa o servidor, em 2007 inicia o projeto piloto e em 16/09/2008 lança o repositório. Atualmente possui 4.404 itens.
POLÍTICAS DE INFORMAÇÃO	Não há política formalizada (está em construção)	Política formal ainda em trâmite pela instituição para sua aprovação. Inicia regulamentação em sua página inicial	Existe política formal, tramitando pelas instâncias superiores para a aprovação
TIPOLOGIAS DOCUMENTÁRIAS EXISTENTES	Livros, Teses e Dissertações, Relatórios técnicos, Vídeos, Imagens e Sons. *Todos os documentos estão em acesso aberto	Fotografias, artigos de periódicos, teses, dissertações, trabalhos de conclusão de curso e trabalhos publicados em eventos. *Alguns documentos (em especial aqueles que foram migrados do formato impresso) possuem acesso restrito, devido à falta de autorização do autor.	Teses, Dissertações, Artigos de periódicos, Anais de eventos, Livros e Capítulos de livros. *Há restrições de acesso, geralmente a pedido do autor, entretanto, a maioria absoluta dos documentos encontram-se em acesso aberto.
OBJETIVOS	“Proporcionar visibilidade à produção científica e intelectual da UFPR”	“Reunir a produção científica e intelectual da UFRGS, proporcionando ampla visibilidade”	“Armazenar, preservar, divulgar e garantir o acesso à produção científica da UNB em formato digital; proporcionar visibilidade; apoiar as atividades de pesquisa e geração de conhecimento científico e apoiar o processo de ensino e aprendizagem por meio do fácil acesso ao conhecimento”
COMITÊ GESTOR	Existe informalmente. É composto por analistas de sistemas, bibliotecários e bolsistas	Existe formalmente desde 20/07/2009. É composto por docentes, profissionais da secretaria de ensino à distância, analistas de sistemas e bibliotecários.	Não há comitê gestor. A gestão está a cargo da Biblioteca Central em parceria com as Faculdades e institutos.
BIBLIOTECAS	Possui sistema de bibliotecas composto por 13 bibliotecas, que utilizam o Sophia como sistema gerenciador de acervo, que permite a interconexão dos metadados com o repositório.	Possui sistema de bibliotecas composto por 33 bibliotecas, que utilizam o ALEPH 500 como sistema gerenciador de acervo, que permite a interconexão dos metadados com o repositório.	Possui sistema de bibliotecas composto pela biblioteca central, biblioteca de pós-graduação e Campus Avançado Asa Sul.
ATRIBUIÇÕES DOS BIBLIOTECÁRIOS	Compete aos bibliotecários a atribuição dos metadados, o controle das autoridades e a submissão dos itens. Não há autodepósito.	Processar os documentos, indexar, catalogar, coletar os dados e a submissão dos itens. Não há autodepósito.	Compete aos bibliotecários todo o gerenciamento do repositório, como coletar os documentos, a definição e revisão dos metadados e a submissão. Não há auto arquivamento.

Quadro 5 – Informações gerais sobre os repositórios Dspace UFPR, Lume e RIUnB

Fonte: Elaborado pela Autora.

Constata-se que obter informações gerais a respeito das etapas que procederam a implantação do repositório em cada uma dessas instituições é uma oportunidade ímpar de troca de informações, especialmente, por meio da entrevista, que proporciona maior aproximação com o tema, podendo avaliar o percurso desde a etapa de seleção do *software*, definição da equipe que irá operacionalizar e apoiar o desenvolvimento do repositório, a sensibilização junto à comunidade acadêmica, bem como as decisões a respeito de quais tipologias o integrarão.

Acredita-se que a definição de uma política que norteie as ações de uma ferramenta de gestão, como o repositório, deve ser uma das primeiras ações da instituição que implantar um repositório, no entanto, vê-se que sua elaboração ocorre concomitantemente com a implantação do repositório, por não ser uma ferramenta estática e por estar sempre em desenvolvimento, cabendo à instituição defini-la mais rígida ou flexível conforme a necessidade.

No caso dos três repositórios estudados, observa-se que, desde a implantação até a etapa de disponibilização dos conteúdos, há um período de muito trabalho, no qual as decisões devem ser tomadas levando em consideração a instituição como um todo e não apenas áreas específicas.

Como toda decisão institucional demanda aprovação de várias instâncias, percebe-se que, se por um lado os processos administrativos diminuem o ritmo e desenvolvimento do repositório, por outro, acabam envolvendo mais e mais instâncias da universidade, tornando-se instrumento democrático de discussão, no qual as decisões são tomadas geralmente em grupos compostos por profissionais das mais diversas áreas, proporcionando constante troca de informações e geração de novos produtos e conhecimentos.

Assim, cabe mencionar que em apenas um dos repositórios analisados pode ser encontrado um Comitê Gestor formalizado pela instituição. Nos demais casos, esta comissão existe informalmente, as atribuições são divididas e as decisões compartilhadas, objetivando sempre a divulgação e relevância deste mecanismo de gestão da informação.

A relevância de um projeto como um repositório é demonstrada em cada um dos objetivos apresentados pelas instituições que já os possuem: reunir, armazenar, tornar visível, garantir acesso permanente, apoiar a geração de conhecimento, incentivar a pesquisa e outros. Observando cada um destes objetivos, constata-se que o papel das bibliotecas das universidades, aqui representadas, por seus Sistemas de Bibliotecas é essencial.

Nas três instituições estudadas, são bibliotecários os responsáveis pelo repositório, na função de Equipe técnica ou Coordenação do repositório, cujas atribuições vão além das etapas de divulgação da ferramenta junto às instâncias administrativas; passam pela customização do *software*, a definição da estrutura, a seleção dos metadados, o controle para que o documento seja preservado, bem como a representação temática e descritiva de seus conteúdos.

7.1.2 Estruturas de Organização da Informação

A estrutura de organização do DSpace UFPR é composta hierarquicamente em: comunidades e coleções, sem a utilização da estrutura de

sub-comunidades. A estrutura de alta generalização (comunidades) é representada pelas tipologias documentárias existentes (Livros, Teses e Dissertações, Sons e Imagens). Os conteúdos destas comunidades estão subdivididos sob uma estrutura denominada: coleções, conforme demonstra a figura:

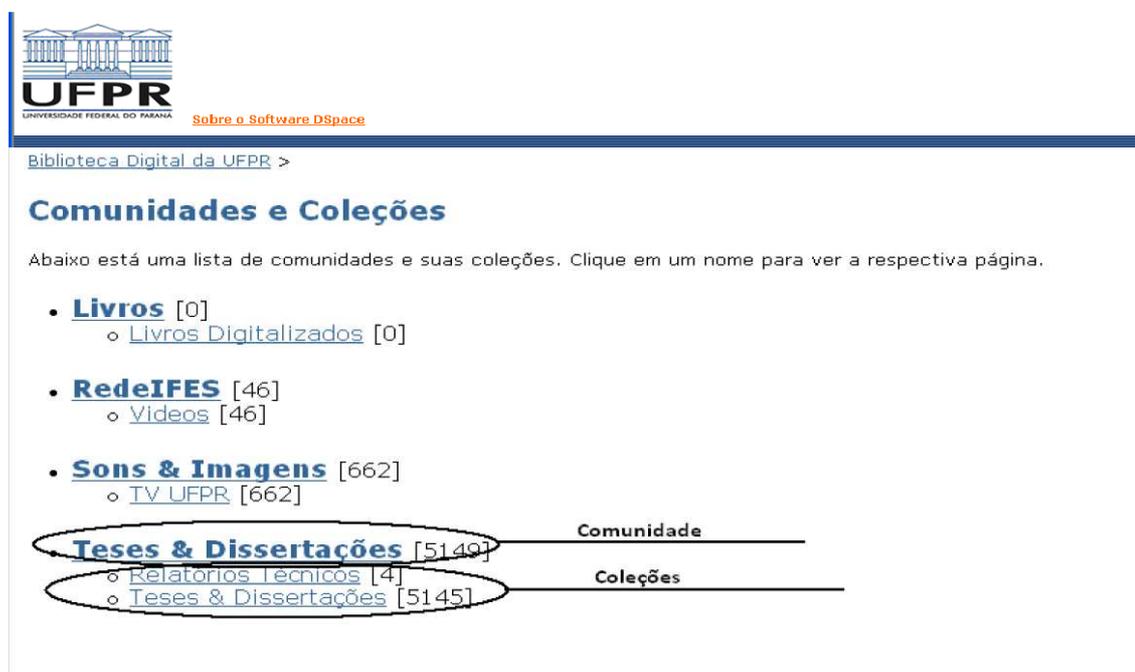


Figura 3 - Estrutura de organização da informação DSpace UFPR

Há quatro coleções existentes neste repositório, que correspondem também à tipologia documentária nele depositada: a Comunidade livros apresenta a coleção livros digitalizados; a Comunidade Rede IFES apresenta a coleção dos Vídeos que podem ser permutados entre as instituições federais de ensino; a Comunidade Sons e Imagens apresenta a coleção TV UFPR e a Comunidade Teses e Dissertações apresenta duas coleções: a primeira Relatórios Técnicos, que organiza a documentação do departamento de informática e a coleção Teses e Dissertações, que concentra maior número de itens depositados, totalizando cinco mil, cento e quarenta e nove documentos submetidos.

Neste repositório, a Comunidade Teses e Dissertações é a que possui o maior número de itens submetidos. Entretanto, cabe destacar que a UFPR utilizou-se da denominação “Teses e Dissertações” tanto para caracterizar a comunidade, quanto para caracterizar a coleção, conforme demonstra a figura a seguir:



 **UFPR**
UNIVERSIDADE FEDERAL DO PARANÁ [Sobre o Software DSpace](#)

[Biblioteca Digital da UFPR >](#)

Comunidades e Coleções

Abaixo está uma lista de comunidades e suas coleções. Clique em um nome para ver a respectiva página.

- **Livros** [0]
 - [Livros Digitalizados](#) [0]
- **RedeIFES** [46]
 - [Videos](#) [46]
- **Sons & Imagens** [662]
 - [TV UFPR](#) [662]
- **Teses & Dissertações** [5149]
 - [Relatórios Técnicos](#) [4]
 - [Teses & Dissertações](#) [5145]

Figura 4 - Uso do mesmo termo na comunidade e coleção DSpace UFPR

Outro fator a ser mencionado é que, no período de análise, a página inicial do DSpace UFPR apresentava uma interface em processo de customização, sendo possível encontrar ainda palavras na língua inglesa, fator que pode vir a ser indicativo de que este repositório encontra-se em fase de testes e/ou estruturação sendo exemplificado pela frase disponível em sua página inicial: “DSpace está funcionando”.

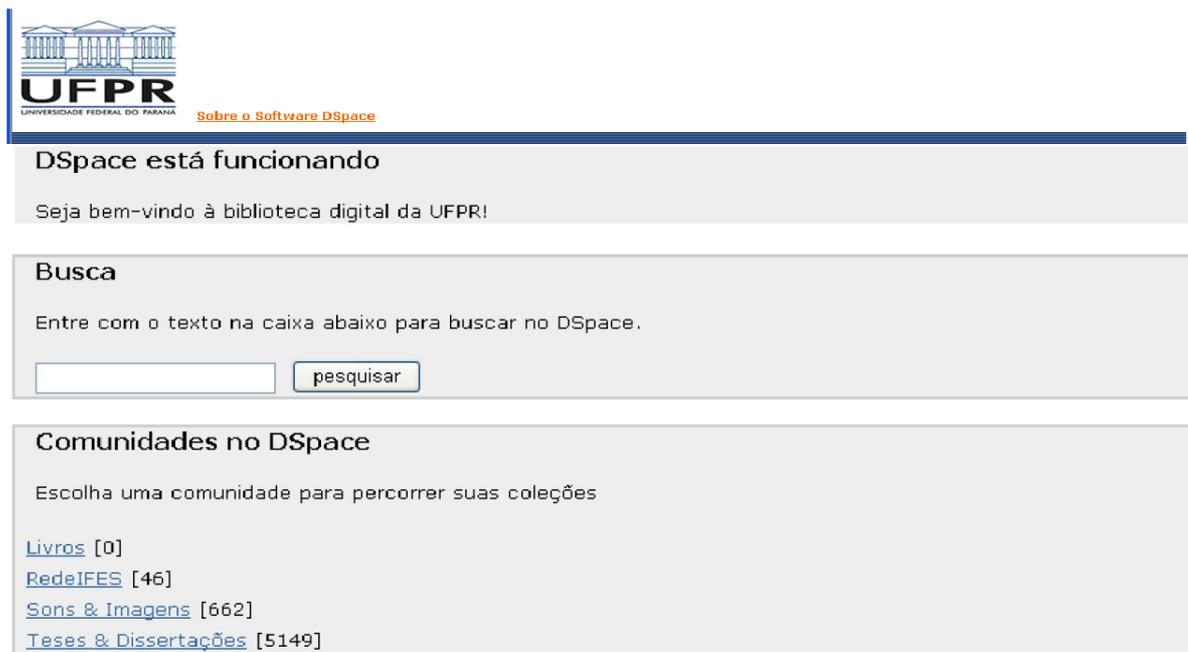


Figura 5 - Customização da página inicial DSpace UFPR

Ao iniciar a identificação da categorização da informação no repositório buscou-se, junto à respondente, uma definição para sua estrutura. De acordo com ela *“o repositório é organizado por tipologias documentárias, ou seja: teses e dissertações e vídeos. Agora, começam abrir novas comunidades, de acordo com a necessidade”*. Assim, questionou-se quais os critérios para a definição de uma comunidade, sub-comunidade e coleção, tendo como resposta da gestora a seguinte afirmação: *“priorizar a produção intelectual e científica”*.

Quanto aos instrumentos utilizados para representação temática e descritiva dos documentos depositados no repositório, de acordo com a respondente, o DSpace UFPR não permite ainda o autodepósito pelos autores, fator que confere às bibliotecas a submissão dos itens ao repositório e a indexação dos documentos e de seus metadados no software Sophia (gerenciador da biblioteca), utilizando-se do CatBib e Bibliodata. *“Utilizamos o MARC, AACR2, CDD e CDU – no caso dos dois últimos instrumentos, optamos por utilizá-los de forma mista. Assim cada um deles atende melhor cada uma das áreas de conhecimento existentes. Quanto ao vocabulário controlado, utiliza-se a Rede Bibliodata e a Library Congress (LC)”*.

No que tange à existência de políticas de indexação para garantir a qualidade dos itens, bem como os instrumentos utilizados para indexação dos conteúdos para o repositório institucional, identificou-se que a UFPR possui o

processamento técnico semi-centralizado, onde parte do trabalho é realizado pela Biblioteca Central e parte pelas bibliotecas setoriais. A responsabilidade pela estruturação das atividades, bem como pela padronização e elaboração de manuais, compete a um grupo de trabalho composto por servidores do SiBi, tanto da Biblioteca Central quanto das setoriais, que definem as planilhas de metadados a serem indexados para a descrição de cada uma de suas tipologias documentárias e os manuais de trabalho que possibilitam a padronização das ações dos bibliotecários.

Ao buscar obter informações a respeito dos pontos positivos e passíveis de melhoria no que tange à estrutura adotada pela instituição para organizar sua produção intelectual em repositórios institucionais, observou-se que *“o principal retorno está no alto índice de qualidade das produções da UFPR que passou a ser visível a partir da criação das bibliotecas digitais”*. Entretanto, os pontos passíveis de melhoria dizem respeito à parte técnica - que envolve o Sistema de Bibliotecas, especialmente no que tange aos metadados e ao controle das autoridades, a respondente menciona que *“como houve aproveitamento dos dados que já existiam no formato impresso, certamente houve perda de qualidade na atribuição dos metadados. Percebe-se que atualmente, a responsabilidade é muito maior, pois a qualidade dos metadados está visível em rede. Entretanto, manter este controle é muito difícil, especialmente em uma estrutura semi-centralizada como a UFPR”*. De acordo com a respondente uma das principais atribuições do Grupo de Trabalho é melhorar a qualidade do controle da base de dados de autoridades.

A estrutura de alta generalização do segundo repositório analisado, o Lume, encontra-se organizada por tipologia documentária, dentre elas: fotografias, artigos de periódicos, teses, dissertações, trabalhos de conclusão de curso e trabalhos publicados em eventos, totalizando vinte e dois mil, quatrocentos e noventa e cinco itens depositados, incluindo além das produções intelectuais da UFRGS, também produções de outras instituições, desde que atendam às necessidades de informação da comunidade acadêmica, conforme demonstra a figura 6:

Lume Repositório Digital
Universidade Federal do Rio Grande do Sul

Lume → Lista de comunidades

Apresentação | Instruções aos autores | Regulamentação | Contato | Ajuda

Pesquisar Enviar

[Pesquisa avançada](#)

Comunidades

Selecione uma comunidade para percorrer suas coleções.

- ☒ [Acervo Fotográfico \[1327\]](#)
- ☒ [Artigos de Periódicos \[2830\]](#)
- ☒ [Teses e Dissertações \[12420\]](#)
- ☒ [Trabalhos de Conclusão de Curso de Especialização \[739\]](#)
- ☒ [Trabalhos de Conclusão de Curso de Graduação \[1337\]](#)
- ☒ [Trabalhos de Eventos \[3842\]](#)

Percorrer

Todo o repositório

- [Comunidades e coleções](#)
- [Anos](#)
- [Autores](#)
- [Títulos](#)
- [Palavras-chave](#)

Meu cadastro

- [Login UFRGS](#)
- [Login não UFRGS](#)

Estatísticas

- [Downloads](#)
- [Itens mais baixados](#)

O autor é titular dos direitos autorais dos documentos disponíveis neste repositório e é vedada, nos termos da [lei](#), a comercialização de qualquer espécie sem sua autorização prévia.

Figura 6 - Estrutura de organização da informação do Lume Repositório Digital

É possível afirmar que a organização do repositório Lume contempla a estrutura de comunidades, sub-comunidades e coleções. Vinculadas à comunidade Teses e Dissertações, há duas sub-comunidades: Teses e Dissertações defendidas na UFRGS, contendo doze mil duzentos e trinta e três itens, e a sub-comunidade Teses e Dissertações não defendidas na UFRGS composta de cento e quarenta itens:

Apresentação | Instruções aos autores | Regulamentação | Contato | Ajuda

Teses e Dissertações

[Estatísticas](#)

Teses, dissertações e trabalhos de conclusão de mestrado profissional produzidos na UFRGS e fora dela.

Pesquisa geral

Autor

Orientador

Título

Palavra-chave

Ano

- Todos
- 2010
- 2009
- 2008

Subcomunidades

- [Teses e Dissertações defendidas na UFRGS \[12330\]](#)
- [Teses e Dissertações não defendidas na UFRGS \[140\]](#)

Entradas recentes

[A elaboração de tarefas de compreensão oral para o ensino de português como](#)

Figura 7 - Sub-comunidades do Lume Repositório Digital

Percebe-se que, mesmo ao optar por manter suas comunidades por tipologias documentárias, a instituição preocupou-se em manter as coleções, em especial aquelas que possuem características científicas como os artigos de periódicos, teses e dissertações, trabalhos de conclusão de curso e trabalho em eventos, ordenados por área do conhecimento conforme demonstra a figura 8:

The screenshot displays the Lume Repositório Digital website interface. At the top, there is a header with the UFRGS logo and the text 'LUME Repositório Digital Universidade Federal do Rio Grande do Sul'. Below the header, a navigation bar includes links for 'Apresentação', 'Instruções aos autores', 'Regulamentação', 'Contato', and 'Ajuda', along with search buttons 'Pesquisar' and 'Enviar'. The main content area is titled 'Artigos de Periódicos' and includes a sub-header 'Artigos científicos publicados em periódicos nacionais e estrangeiros por docentes e servidores técnico-administrativos da UFRGS.' A search form is present with fields for 'Pesquisa geral', 'Autor', 'Título', and 'Palavra-chave'. There are also dropdown menus for 'Origem' (with options: Todos, Periódico estrangeiro, Periódico nacional) and 'Ano' (with options: Todos, 2010, 2009, 2008). An 'Enviar' button is located below the search form. To the right, there is a sidebar with navigation options: 'Percorrer', 'Todo o repositório' (with links for Comunidades e coleções, Anos, Autores, Títulos, Palavras-chave), 'Esta comunidade' (with links for Anos, Autores, Títulos, Palavras-chave), 'Meu cadastro' (with links for Login UFRGS, Login não UFRGS), and 'Estatísticas' (with links for Downloads, Itens mais baixados). At the bottom of the page, there is a list of collections with their respective counts: Ciências Agrárias [687], Ciências Biológicas [500], Ciências da Saúde [1285], Ciências Exatas e da Terra [182], Ciências Humanas [265], Ciências Sociais Aplicadas [319], and Engenharias [121].

Figura 8 - Organização das coleções do Lume Repositório Digital

O repositório Lume vincula suas coleções às suas comunidades, de diferentes formas, a comunidade Acervo fotográfico: possui três coleções que correspondem à estrutura administrativa da universidade (Escola de Educação Física, Instituto de Física e Museu Universitário). As comunidades Teses e Dissertações, Artigos de Periódicos, Trabalhos de Conclusão de Curso de Especialização e Trabalhos publicados em Eventos, possuem as suas coleções ordenadas de acordo com as nove áreas de conhecimento (Ciências Agrárias, Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Ciências

Humanas, Ciências Sociais Aplicadas, Engenharias, Lingüística, Letras e Artes e Multidisciplinar).

Cabe mencionar que a comunidade denominada Trabalhos de Conclusão de Curso de Graduação possui uma coleção diferenciada. Neste caso, os itens não estão subdivididos por área do conhecimento e sim por entrada recente dentro de uma coleção também denominada “Trabalhos de Conclusão de Curso”.

Figura 9 - Uso do mesmo termo na comunidade e coleção Lume

Dentre as comunidades existentes no Lume, a que apresenta a maior quantidade de itens depositados é a de Teses e Dissertações, com doze mil, quatrocentos e vinte itens; seguida da Trabalhos de eventos, com três mil, oitocentos e quarenta e dois itens, o que nos permite inferir que o maior número de itens depositados corresponde àqueles cuja determinação é de obrigatoriedade de depósito.

Um fator interessante observado neste repositório é a possibilidade de identificar as estatísticas de *download* por comunidade, sub-comunidade e coleções, e os itens mais baixados – por coleção, fator de extrema relevância no intuito de demonstrar junto à comunidade a abrangência e visibilidade oferecida na

disponibilização das produções intelectuais em repositórios.

Os apontamentos acerca da categorização estabelecida para o repositório Lume, demonstram que sua forma de organização é diferente da encontrada usualmente em outros repositórios DSpace, que priorizam as estruturas administrativas.

A gestora explica esta decisão ao afirmar que *“nossa estrutura contraria o que se vê normalmente em outros repositórios – especialmente aqueles que refletem em uma estrutura organizacional da instituição à qual pertencem, acreditamos que esta forma de organização dificulta a busca para o usuário externo. Quem sabe futuramente possamos pensar diferente, mas neste momento, optamos por tipo de documento tendo em foco o usuário. Cada comunidade tem uma estrutura que atende melhor tanto os usuários internos quanto os externos”*.

Ao questionar a respeito dos critérios para a definição de comunidades, sub-comunidades e coleções, identificou-se que a avaliação para inclusão de novas comunidades e coleções no Lume leva em consideração os seguintes aspectos: inicialmente se a produção é da UFRGS e se o proponente possui interesse em disponibilizar livremente este documento na rede, a questão do conteúdo proposto, sua adequação aos objetivos do repositório, seu formato, bem como o comprometimento dos proponentes em manter o fluxo contínuo de novos objetos digitais, mantendo o repositório atualizado.

Quanto aos instrumentos utilizados para a descrição dos documentos no repositório, identificou-se que cada uma das trinta e três bibliotecas que compõem o Sistema de Bibliotecas da UFRGS realiza a indexação dos documentos utilizando instrumentos próprios e específicos da área da ciência da informação, como os vocabulários controlados e os tesouros. Detalhadamente a respondente indica que *“utilizamos o AACR2 nível dois de descrição. No SABI utilizamos o formato MARC para descrição bibliográfica e, no Lume, o padrão de metadados é o Dublin Core, assim, fizemos uma tabela de compatibilização dos campos MARC-Dublin Core. No caso de teses e dissertações, trabalhos de conclusão de curso de especialização e de graduação é incluído o resumo e as palavras-chave em todos os idiomas disponíveis no documento”*.

Esta iniciativa visa ampliar o alcance e visibilidade do documento, dentro e fora do país. Percebeu-se neste repositório o interesse em ampliar a disseminação, possível por meio da indexação a ampliação das possibilidades de

busca e recuperação da informação. No entanto, de acordo com a entrevistada “o uso da linguagem natural facilita a indexação, mas é fato que apresenta problemas, é necessário que ainda ocorram melhorias neste sentido”.

Constatou-se que não há uma política de indexação para o Lume nem para o SBU. Cada biblioteca utiliza instrumentos próprios, conforme a(s) área(s) do conhecimento que abrange. Entretanto, algumas coleções como: Acervo fotográfico do Museu universitário, da Escola de Educação Física e do Instituto de Física possuem um padrão de registro de metadados diferenciado, tendo em vista que os próprios responsáveis pelas comunidades preenchem os metadados. Estas coleções são incluídas no sistema Acervo Fotográfico e posteriormente coletadas pelo Lume. De acordo com a entrevistada “a equipe técnica se reuniu e definiu os metadados para estas coleções, baseados no AACR2. Embora a equipe destas comunidades alimente os dados em outro sistema, da mesma forma como o Lume coleta dados no SAbi, ele coleta neste outro sistema, compatibilizando dados, para atender ao padrão Dublin Core. Assim, é possível manter um padrão de qualidade na inclusão dos metadados”.

Procurou-se observar a existência de pontos positivos e outros passíveis de melhoria no que tange à estrutura adotada pela instituição para organizar sua produção intelectual, sendo que, de acordo com a gestora, a avaliação é muito positiva “acredito que a organização está boa. Temos o SAbi como instrumento de coleta e registro da produção intelectual da UFRGS desde 1989. Quanto aos fatores a serem melhorados com relação ao Lume, considero a necessidade de melhorar os canais de divulgação e o fluxo dos documentos”.

O terceiro repositório analisado, o RIUnB, diferencia-se tanto do DSpace UFPR quanto do Lume, visto que sua estrutura de organização é composta hierarquicamente por comunidades, sub-comunidades e coleções. A generalização (comunidades) reflete a estrutura administrativa da instituição, generalizando-a em vinte e cinco especialidades que permitem a ordenação das produções intelectuais dos acadêmicos e servidores vinculados às faculdades, institutos e a biblioteca da UnB.

Nas comunidades, os conteúdos são subdivididos em outras duas estruturas: sub-comunidades e coleções. Desta forma, é possível encontrar no

repositório comunidades que possuem sub-comunidades e coleções e outras que possuem apenas coleções, conforme apresenta a ilustração abaixo:

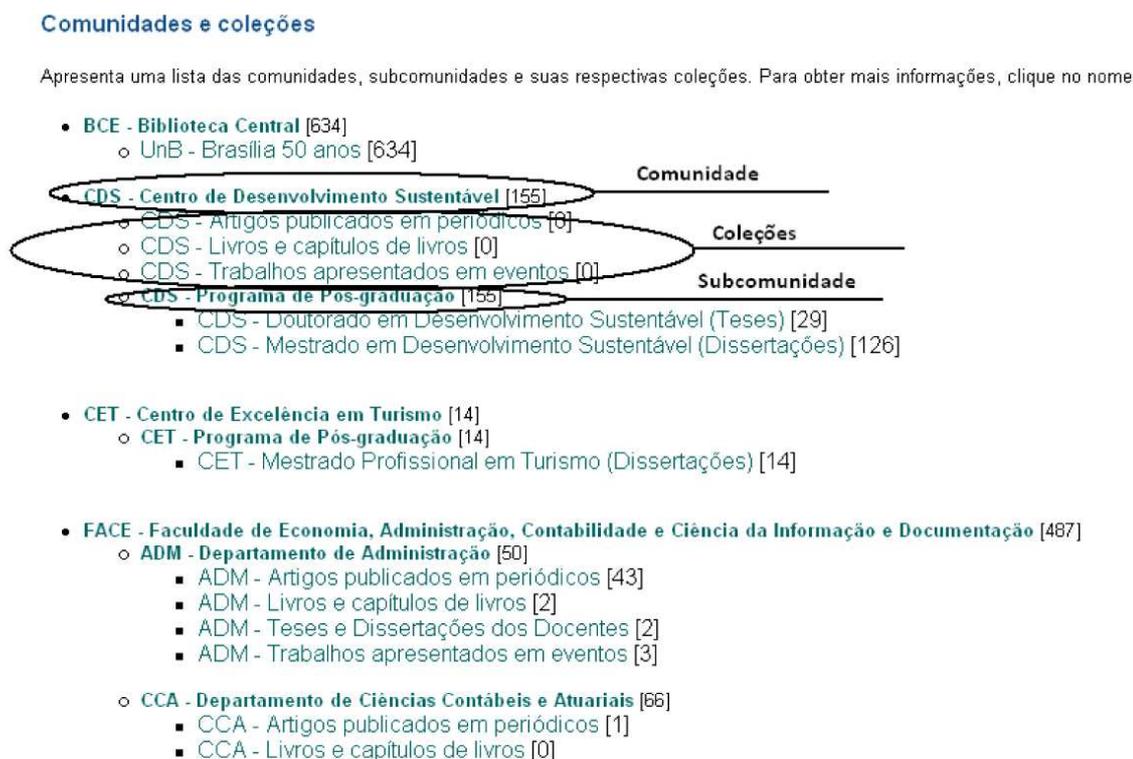


Figura 10 - Estrutura de organização da informação do RIUnB

Isto se deve ao fato de que as setenta e cinco sub-comunidades existentes no RIUnB correspondem aos programas de pós-graduação e aos departamentos subordinados às faculdades mencionadas nas comunidades. Dessa forma, não existem sub-comunidades nos departamentos e/ou programas que não possuem programas de pós-graduação.

Porém, as suas trezentas e uma coleções estão presentes em todas as comunidades, pois é nesta estrutura que as tipologias documentárias estão definidas e ordenadas.

A análise da estrutura deste repositório identificou que o maior número de itens depositados está nas coleções pertencentes aos programas de pós-graduação (mestrado e doutorado) cuja tipologia trata-se de Teses e Dissertações. Há também grande quantidade de depósitos de artigos de periódicos, fator que pode conferir ao repositório uma característica voltada ao depósito de documentos com caráter científico, ou seja, validado por pares.

Constatou-se, na entrevista com a gestora, que os critérios adotados para que esta estrutura (comunidade, sub-comunidade e coleção) possa ser criada, vão ao encontro da estrutura da UnB. Desta forma, de acordo com a respondente *“caso seja criado um novo departamento, faculdade, ou instituto, suas respectivas comunidades e sub-comunidades serão criadas. As coleções são criadas pelos tipos de documentos, ou seja, só será criada uma nova coleção se ficar estabelecido que haverá um novo tipo de documento para entrar no repositório”*.

Quanto ao aspecto da representação temática e descritiva adotada pela instituição, constata-se a utilização dos metadados, seguindo o padrão Dublin Core que foi estabelecido pensando nos tipos de documentos que o repositório abriga, bem como o uso de instrumentos como os códigos de catalogação, classificação e os vocabulários controlados.

Assim, a UnB optou por manter políticas que resguardem a qualidade da indexação dos materiais por meio de um manual produzido em parceria com os bibliotecários da catalogação a fim de padronizar as entradas de autor, orientador e assunto. Para indexação utilizam-se do vocabulário controlado do Pergamum, combinando-o com a linguagem natural.

Questionando a respeito dos pontos positivos e passíveis de melhoria na estrutura adotada pela instituição para organizar sua produção intelectual em repositórios institucionais, obteve-se como resposta que *“é necessário melhorar a divulgação do serviço. Algo imprescindível para qualquer instituição que adote um sistema de repositórios é o apoio da gerência da instituição, no que diz respeito ao processo de institucionalização. O repositório tem que ser institucionalizado para que suas políticas sejam seguidas por qualquer administração que possa vir, evitando que cada gestão queira inserir aquilo que acha melhor. Na UnB o processo de institucionalização está caminhando”*.

Dando sequência à análise, o quadro a seguir representa a segunda categoria de análise que corresponde às estruturas de organização encontradas em cada um dos repositórios em estudo.

REPOSITÓRIOS	DSPACE UFPR	LUME	RIUNB
ESTRUTURA DE ORGANIZAÇÃO	Utiliza a estrutura do DSpace: Comunidades e coleções. Sua estrutura de alta generalização reflete as tipologias documentárias existentes	Utiliza a estrutura DSpace: Comunidades, Sub-comunidades e Coleções. A estrutura de alta generalização (comunidade) está organizada por tipologias documentárias	Utiliza a estrutura DSpace: Comunidades, Sub-comunidades e Coleções. A estrutura de alta generalização está organizada por estrutura administrativa
COMUNIDADE	Possui 4 comunidades: 1) Livros 2) Rede Ifes 3) Sons e Imagens 4) Teses e Dissertações	Possui 6 comunidades: 1) Acervo Fotográfico; 2) Artigo de periódico; 3) Teses e Dissertações; 4) TCC Graduação 5) TCC Especialização 6) Trabalho em Evento	Possui 25 comunidades que representam a maior estrutura acadêmica (Faculdades, Institutos, etc.)
SUB-COMUNIDADE	NÃO POSSUI	Possui apenas duas, vinculadas à Comunidade Teses e Dissertações: 1- Defendidas na UFRGS 2- Não defendidas na UFRGS	Possui 75 sub-comunidades que representam as estruturas hierárquicas menores da instituição
COLEÇÃO	Possui 5: 1) Livros digitais 2) Vídeos 3) TV UFPR 4) Relatórios Técnicos 5) Teses e Dissertações	As coleções são ordenadas por área de conhecimento, exceto a Comunidade Acervo fotográfico, que possui as seguintes coleções: 1- Escola de Educação Física 2- Instituto de Física 3- Museu Universitário	Possui 301 coleções, subordinadas às comunidades e sub-comunidades, organizadas pelas seguintes tipologias documentárias (Livro, capítulo de livro, artigo, teses, dissertações e Trabalho em eventos)
CRITÉRIOS PARA DEFINIÇÃO	Priorizar a produção intelectual e científica da UFPR, sendo que as comunidades surgem de acordo com a demanda	Quando há procura pela criação de uma nova Comunidade, o comitê gestor avalia o conteúdo proposto, o formato dos documentos e, sobretudo, o comprometimento que proponente possui em manter o fluxo contínuo de submissão de novos documentos.	As comunidades e suas respectivas sub-comunidades e coleções são criadas, quando é criado um novo departamento, de acordo com a estrutura da UnB
INSTRUMENTOS UTILIZADOS	Formato MARC; AACR2; CDD e CDU; Rede Bibliodata; Vocabulário da Library Congress	Formato MARC compatibilizado para o padrão Dublin Core; AACR2 nível 2;	Formato Dublin Core AACR2 Utilizam vocabulário controlado do Pergamum combinado com linguagem natural
POLÍTICAS DE INDEXAÇÃO	Como o SIBI possui o processamento técnico semi-centralizado, uma equipe de trabalho define os metadados, disponibilizando-os em uma planilha, que define os campos e formato de indexação padronizada para atender a cada uma das tipologias documentárias.	Não há política formal de indexação nem para o Lume nem para o SBU	Existem manuais de padronização que são desenvolvidos por bibliotecárias no intuito de garantir a qualidade na padronização de autores, orientadores e assuntos.
OBSERVAÇÕES	Observou-se como ponto positivo, o alto índice de qualidade que a UFPR possui nos documentos disponibilizados e como pontos a serem repensados, destaca-se a questão da qualidade dos metadados, fator que vem sendo trabalhado nesta universidade	A estrutura de organização do LUME está muito boa. A utilização do SABI como instrumento oficial de coleta e registro da produção intelectual da UFRGS traz resultados positivos. No entanto é necessário ampliar os canais de divulgação e fluxo dos documentos.	Como ponto positivo cabe mencionar a institucionalização do repositório na UnB, fator que garante sua permanência e sustentabilidade. Entretanto, para ganhar mais adeptos necessita melhorar sua divulgação.

Quadro 6 – Síntese das estruturas de organização da informação dos repositórios

Fonte: Elaborado pela autora.

A estrutura do DSpace UFPR contempla as tipologias documentárias produzidas na instituição e também fora dela, desde que sejam produzidas por membros da comunidade acadêmica. Ao visualizar a estrutura das comunidades

deste repositório, presume-se que o mesmo tenha o interesse de disponibilizar em formato digital, além dos documentos científicos, as produções intelectuais como sons e vídeos produzidos na instituição.

Já o repositório Lume optou por manter sua estrutura de acordo com as tipologias documentárias. Esta forma de organização, aparentemente, mostra-se eficiente ao facilitar a seleção, desde a primeira camada hierárquica (comunidade), de qual tipo de documento se quer buscar a informação, especialmente, para a comunidade externa à instituição. As coleções se apresentam ou por departamento da universidade ou por área do conhecimento. Cabe mencionar ainda um aspecto relevante neste repositório, a disponibilização da documentação que respalda sua criação, a obrigatoriedade de depósito dos documentos produzidos na instituição, bem como instruções aos autores. Ainda, no que diz respeito aos direitos autorais, desde a página inicial do repositório, há indicação de que o autor é o titular dos direitos autorais dos itens submetidos ao Lume, sendo que, cada item está licenciado sob uma licença *creative commons*.

Quanto ao RIUnB, sua estrutura de alta generalização (comunidades) corresponde à estrutura acadêmica/administrativa da UnB, é um fator que minimiza esforços do usuário que conhece esta estrutura, ou seja, o usuário interno da instituição. Entretanto, percebe-se que o usuário externo à instituição terá condições de obter os resultados necessários às suas pesquisas, sobretudo, pela grande quantidade de recursos de recuperação da informação disponíveis.

Quanto à definição destas estruturas, o DSpace UFPR, estrutura seu repositório de acordo com as solicitações que lhes são feitas, ou seja, são criados de acordo com a necessidade que vai surgindo na instituição, fator que ocorre também no repositório Lume, que diferencia-se no sentido que cada uma das solicitações devem ser aprovadas pelo Comitê Gestor, o qual avalia a relevância e a continuidade de submissão de documentos, tendo em vista a dinamicidade característica de um repositório institucional. Quanto ao RIUnB, esta definição é da própria universidade, visto que o repositório reflete sua estrutura administrativa e acadêmica. Desta forma, se surgem novas coordenações, departamentos, cursos, conseqüentemente surgirão novas comunidades, sub-comunidades e coleções no RIUnB.

De acordo com as informações obtidas, comum aos três repositórios analisados, é o fato de o maior número de itens depositados estar atrelado à

tipologia Teses e Dissertações, o que nos faz refletir que os repositórios concentram os resultados de pesquisas científicas geradas nas universidades. Santos, Teixeira e Pinto (2005) afirmam que, nos últimos anos, tem aumentado o número de implantações, principalmente na comunidade universitária, com objetivos orientados à gestão, ao armazenamento, à preservação e à divulgação do trabalho intelectual produzido por essa comunidade.

Inferre-se sob este aspecto, que muitas instituições ampliam assim, as suas Bibliotecas Digitais, acrescentando a elas novas tipologias documentárias, novos formatos de documentos, sobretudo aqueles validados por pares.

Quanto aos instrumentos utilizados, não se pode dizer que somente o padrão Dublin Core é utilizado, visto que muitos dados são cooperados de sistemas gerenciadores de bibliotecas que se utilizam em todos os casos aqui analisados, do padrão MARC, AACR2, CDD, CDU, vocabulários controlados da LC (*Library Congress*) ou do sistema gerenciador da biblioteca combinados com a linguagem natural.

No entanto, sabe-se que a qualidade da indexação é fator de extrema relevância para a recuperação das informações, assim, percebe-se que cabe às bibliotecas definir políticas, planos de trabalhos, manuais, planilhas para oportunizar a padronização na entrada dos metadados, resultando em maior qualidade nos resultados de pesquisa.

Os pontos positivos encontrados em todos os repositórios estão na integração das bibliotecas aos demais órgãos da universidade, fator que possibilita maior entendimento a respeito das atribuições de um bibliotecário em uma instituição de ensino. Este profissional pode contribuir na definição das estruturas de organização, tendo em vista os conhecimentos relativos às categorizações, relevância, hierarquia, etc. Percebe-se que, embora diferentes em suas estruturas, os repositórios obtiveram os objetivos esperados: visibilidade, melhoria na qualidade do que é produzido. Entretanto, para que ele mantenha-se em constante crescimento, é necessário que consiga ser institucionalizado, ou seja, conhecido e reconhecido por todas as instâncias da universidade, para tanto necessita de grande divulgação e controle dos fluxos dos documentos.

7.1.3 Metadados Utilizados

A fim de identificar os metadados utilizados para descrever os documentos, realizaram-se pesquisas rápidas e avançadas às diferentes tipologias documentárias com o objetivo de identificar o padrão adotado para cada tipo de material.

Desta forma, por meio da captura das telas (registro completo de metadados) de cada uma das tipologias documentárias existentes no repositório Dspace UFPR, Lume e RIUnB procedeu-se a análise dos metadados utilizados para a descrição de seus documentos.

A primeira comunidade do DSpace UFPR, denominada: Livros, possui apenas uma coleção: livros digitais, entretanto não há itens para análise.

A comunidade Rede IFES abriga a coleção de Vídeos, que armazena imagens. Os metadados definidos para sua descrição em formato completo são: data de acesso, data de disponibilidade, data de emissão, identificador persistente, título e a coleção em que este item está disponível, bem como o formato em que o arquivo encontra-se disponível, neste caso MPEG.

A comunidade Sons e Imagens, cuja coleção denomina-se TV UFPR, apresenta os seguintes campos de descrição: autor, colaborador, data de acesso, data de disponibilidade, data de criação, data de emissão, identificador persistente, tipo de material, língua, editora, fonte, resumo, título e coleção, bem como o formato em que o arquivo encontra-se disponível, neste caso *Unknown*.

A comunidade Teses e Dissertações apresenta as seguintes *tags* para representar a coleção “Relatórios Técnicos”: autor, data de acesso, data de emissão, identificador persistente, descrição, tipo, formato, idioma, título e coleção onde se encontra o item. O formato de disponibilização do item é o *Postscript*.

Já a coleção “Teses e Dissertações” apresenta os seguintes campos: autor, colaborador, data de acesso, data de disponibilidade, data de criação, data de emissão, identificador persistente, resumo, tipo, idioma, protocolo OAI, url relacionada, título e coleção a que pertence o item. O formato de disponibilização do item é o PDF. A análise desta coleção permitiu observar que a informação a respeito do orientador foi inserida no campo relativo a resumo.

Cabe mencionar que o item formato do documento relacionado à manifestação física ou digital do recurso (ex: html, ppt, pf, gif, xls), não está localizado dentro dos campos Dublin core, mas em tela separada.

Por meio de pesquisa ao catálogo virtual do Sistema Sophia (sistema gerenciador de informação das bibliotecas da UFPR), foi possível identificar a existência de interoperabilidade entre este sistema e o repositório por meio do campo: “localização da fonte”, ou seja, há a inclusão desta etiqueta, que permite remeter para o *link* do documento disponível no repositório.

Realizada a análise dos metadados utilizados em cada uma das tipologias, pode-se supor que a instituição não utiliza um formulário padrão para a entrada dos metadados, sendo que, dependendo do documento, faz-se necessário utilizar determinados campos para sua descrição. Alguns apontamentos são necessários após a identificação e observação deste repositório, sobretudo nos seguintes aspectos: não há campo para descrição de palavra-chave que indique o assunto dos documentos das comunidades “vídeos e sons” e “relatórios técnicos” neste repositório; o campo “orientador” aparece apenas no campo resumo; as informações doc source: oai: ufpr.br e doc. source aparecem para descrever a interoperabilidade entre o sistema de informação da instituição e o repositório e, no segundo caso, para relacionar onde o item encontra-se na *web*.

Um aspecto relevante a ser mencionado diz respeito ainda à definição do esquema de metadados a serem utilizados em cada uma das tipologias existentes no repositório. Esta é uma tarefa que diz respeito ao Sistema de Bibliotecas, visto que a padronização não mudou, o que alterou foi o suporte onde a informação passou a ser disponibilizada. De acordo com a respondente “*A informação já existia, o advento das tecnologias de informação alterou o suporte, mas não o conteúdo a ser analisado, dessa forma, continua sendo tarefa da biblioteca e utilizamos para isso o padrão MARC*”.

Quanto à existência dos campos, pode-se mencionar que determinadas comunidades, como Sons e Vídeos, por se tratar de tipologias cujos conteúdos resultam em termos repetitivos, excluem-se alguns dos campos, como se identificou na etapa de coleta de dados com o instrumento formulário, no que diz respeito à inexistência do campo 650 (assunto), nestes tipos de materiais, quando o autor e/ou título e/ou assunto se repetem. De acordo com a respondente, este fator ocorre com o intuito de “*evitar a repetição de informação, tendo em vista que a*

recuperação irá ocorrer de igual maneira, levando em consideração as palavras do autor ou título, por exemplo”.

Questionou-se se foram utilizados os campos de metadados pré-estabelecidos pelo *software* Dspace, ou se houve a necessidade da inserção de outros campos para descrição das produções no repositório. Neste aspecto, a respondente informou que este procedimento é realizado por meio do Sophia (sistema gerenciador da Biblioteca), o que indica que o tratamento da informação não ocorre no DSpace, ou seja, no padrão Dublin Core e sim no Sophia, que utiliza o padrão MARC.

Levando em consideração a importância da indexação dos termos para a recuperação da informação, buscou-se identificar a existência de uma definição quanto ao número de descritores utilizados para representação de assunto de cada obra no DSpace UFPR. A resposta da gestora do repositório indicou a utilização de vocabulários controlados para a definição das palavras que representarão o assunto, não havendo um limite de termos a serem utilizados, fica a cargo do indexador a definição da quantidade de descritores que melhor represente a obra. No entanto, utiliza-se também de termos livres para representação de um documento, como ocorre, geralmente, no caso dos resumos e dos títulos.

Cabe salientar que como não há autodepósito, quem realiza o registro e controle dos metadados é a própria biblioteca, o que permite a utilização de vocabulários controlados para a indexação de seus conteúdos.

Analisando a questão dos metadados no Lume Repositório Digital, constata-se que a tipologia Fotografias apresenta os seguintes campos de metadados: autor, data do documento, data de submissão do arquivo, ano, identificador persistente, formato de extensão, descrição, formato do documento, local, assunto, título, tipo de documento, identificador (referente ao sistema de onde o documento é proveniente) e o histórico de descrição (resumo da obra).

Para a descrição dos Artigos de periódicos, são utilizados os seguintes campos: autor, data de submissão do artigo, ano de publicação, ISSN, identificador persistente, formato do arquivo, língua, assunto, título, tipo de documento, identificador.

As Teses e Dissertações e os Trabalhos de Conclusão de Curso de Graduação e Especialização apresentam os mesmos campos de descrição dentre eles: título, link para texto completo, autor, orientador, data, nível (mestrado ou

doutorado), instituição, palavra-chave, resumo (em todas as línguas que estiverem no documento).

Já os trabalhos publicados em eventos contemplam os metadados de autor, data de submissão, data de publicação, identificador persistente (link para acesso), formato e extensão do documento, periódico, assunto, título, tipo de material, evento onde foi publicado e número de identificação.

Um fator de extrema relevância a ser analisado neste estudo diz respeito a definição dos metadados. Desta forma, questionou-se como ocorreu a definição dos campos que seriam utilizados nas tipologias documentárias disponíveis no Lume e obteve-se como resposta que a opção foi utilizar os campos do formato MARC, adotados no SABI, para cada tipo de documento e compatibilizá-los com o formato Dublin Core, partindo dos quinze elementos básicos, tendo sido necessário qualificar alguns deles, para atender às especificidades de cada coleção.

Quanto à indexação, no que diz respeito à existência de uma definição ao número de descritores a serem utilizados para representação do assunto de cada obra, identificou-se que a UFRGS utiliza instrumentos distintos para a indexação, conforme a biblioteca, porém, não há um número definido de termos a serem atribuídos a cada obra. Observou-se que o registro e controle dos metadados, até o momento, estão a cargo das bibliotecas. Cabe ressaltar o apontamento da entrevistada *“Por enquanto nenhum autor submete diretamente os seus documentos para o LUME, mas é provável que futuramente adotaremos o autoarquivamento”*.

Identificou-se que o RIUnB utiliza o padrão de metadados Dublin Core e o protocolo OAI-PMH. As Teses e Dissertações apresentam os dezoito campos de metadados em formato completo, são eles: orientador - cuja entrada se dá pelo sobrenome, nome; autor – entrada pelo sobrenome, nome; data de acesso; data de disponibilidade do item (no padrão dia, mês, ano e horário); data de emissão (no padrão mês e ano); identificador de citação no formato: autor. título. ano, número de páginas, tipo de material, instituição, cidade e ano, identificador persistente, número de folhas, resumo na língua portuguesa, resumo na inglesa, língua em que está o texto, direitos de acesso (neste caso indicando acesso aberto); título e subtítulo do documento, tipo de documento, neste caso tese; descrição do documento: inclui vínculo a instituição e finalidade, palavra-chave, país e coleção a que este documento pertence.

A tipologia Livros abrange os seguintes padrões: autor: entrada sobrenome, nome, data de publicação, data de disponibilidade do item (no padrão dia, mês, ano e horário); data de emissão (no padrão mês e ano); identificador de citação (no formato autor, título, ano); número de páginas; tipo de material; instituição; cidade e ano; identificador persistente; resumo na língua portuguesa, língua em que está o texto; direitos de acesso (neste caso indicando acesso aberto); título e subtítulo do documento; tipo de documento (neste caso livro); descrição do documento (inclui número de folhas, palavra-chave, editores) e coleção a que este documento pertence.

Os capítulos de livros possuem 16 campos: autor, data de acesso, data de disponibilidade, data de emissão, identificador da citação, identificador da página (on-line), descrição, resumo, língua, editor, direitos de acesso, título, tipo de documento (neste caso: capítulo de livro), palavra-chave, país, coleção.

Os trabalhos em eventos disponibilizam os seguintes metadados: autor: entrada sobrenome, nome, data de acesso, data de disponibilidade do item (no padrão dia, mês, ano e horário); data de criação; data de emissão (no padrão mês e ano); Identificador de citação (no formato autor, título, ano); número de páginas; tipo de material; instituição; cidade e ano; identificador persistente; descrição do documento (inclui número de folhas), resumo na língua portuguesa e em língua estrangeira; língua em que está o texto; direitos de acesso (neste caso indicando acesso aberto), título e subtítulo do documento, tipo de documento (neste caso trabalho em eventos); palavra-chave; país e coleção a que este documento pertence.

Finalmente, os artigos, apresentam os seguintes elementos autor (entrada sobrenome, nome); data de publicação; data de disponibilidade do item (no padrão dia, mês, ano e horário); data de emissão (no padrão mês e ano); identificador de citação (no formato autor, título, ano); número de páginas; tipo de material; instituição; cidade e ano; identificador persistente; descrição do documento (inclui número de folhas); resumo na língua portuguesa e na língua inglesa; língua em que está o texto; direitos de acesso (neste caso indicando acesso aberto); título e subtítulo do documento; título alternativo; tipo de documento (neste caso: artigo de periódico); palavra-chave; país e coleção a que pertence este documento.

Cabe mencionar que o item “formato” relacionado à manifestação física ou digital do recurso (ex: html, ppt, pf, gif, xls), não está localizado nos campos Dublin Core. Aparece em tela separada, como demonstrado a seguir:



Meu espaço | Editar perfil | Fale conosco

RIUnB
repositório Institucional da Universidade de Brasília

Página inicial | Comunidades e Coleções | Data de publicação | Autor | Orientador | Título | Assunto | Ajuda

RIUnB >
CDS - Centro de Desenvolvimento Sustentável >
CDS - Programa de Pós-graduação >
CDS - Doutorado em Desenvolvimento Sustentável (Teses) >

Utilize este link para identificar ou citar este item: <http://hdl.handle.net/10482/1890>

Arquivo	Descrição	Tamanho	Formato
2008_RicardoJoseCaleboMarra.pdf		5,68 MB	Adobe PDF ver/abrir

Registro completo

Campo Dublin Core	Valor	Língua
dc.contributor.advisor	Mota, José Aroudo	-
dc.contributor.author	Marra, Ricardo José Calebo	-
dc.date.accessioned	2009-10-07T16:28:10Z	-
dc.date.available	2009-10-07T16:28:10Z	-
dc.date.issued	2008-12-15	-
dc.identifier.citation	MARRA, Ricardo José Calebo. Critérios de relevância para classificação de cavernas no Brasil. 2008. 393 f. Tese (Doutorado em Desenvolvimento Sustentável)-Universidade de Brasília, Brasília, 2008.	en
dc.identifier.uri	http://hdl.handle.net/10482/1890	-

Figura 11 – Indicação do formato do arquivo disponível no RIUnB

Realizada a análise dos metadados utilizados em cada uma das tipologias, pode-se supor que a instituição utilize um formulário padrão para a entrada dos metadados, sendo que, dependendo do documento, faz-se necessário ou não detalhar determinados aspectos. As principais diferenças identificadas na observação a este repositório deram-se nos seguintes aspectos: não há um número padrão de palavras-chave a serem utilizadas na descrição de um documento; o campo “orientador” aparece apenas na descrição de teses e dissertações; as tipologias Livro e Capítulo de livros apresentam o campo “editores”, não presente nos metadados das demais tipologias; o campo “título alternativo” aparece somente na tipologia Artigos.

Observou-se que a instituição definiu seu esquema de seleção dos campos conforme suas necessidades, em concomitância com os documentos a serem inseridos no repositório, seguindo o padrão Dublin Core e “*procurou-se utilizar*

aqueles metadados que já são utilizados pela comunidade acadêmica que possuem repositórios estabelecidos para não destoar muito do que já havia sido feito”.

De acordo com a entrevista, os metadados pré-estabelecidos no DSpace não contemplavam todos os documentos que seriam inseridos no repositório, desta forma, foi necessária a inserção de outros campos.

Quanto à representação de assunto de cada obra, identificou-se que não há um número limite de descritores para isso. Desta forma, os documentos são indexados de acordo com o que cada indexador acha necessário. Outro fator citado pela respondente diz respeito a UnB ainda não ter possibilitado a implementação de uma política de autoarquivamento, desta forma, quem realiza o registro e controle dos metadados na submissão é a Biblioteca.

Observou-se que todos os repositórios estudados seguem o padrão MARC (para documentos provenientes de seus sistemas de gestão de acervo), e os 15 campos do Dublin Core (DSpace), sendo que no caso do repositório Lume, houve a necessidade de qualificar alguns dos campos Dublin Core para atender à necessidade de determinada tipologia documentária.

A interoperabilidade de dados é permitida pelos sistemas Sophia, Aleph 500 e Pergamum utilizados pelas bibliotecas das instituições nos repositórios. Cabe mencionar que não há possibilidade de manter um padrão único/estático de metadados, tendo em vista as características de cada um dos documentos existentes. Assim, como não há autodepósito, o controle de qualidade dos itens indexados ainda é realizado pela própria biblioteca, responsável tanto pela atribuição quanto pelo controle dos metadados, sendo que, para isto, são utilizados vocabulários controlados e a linguagem natural para indexar os descritores que serão definidos pelo profissional bibliotecário.

7.1.4 Interoperabilidade Humana e Semântica

No que tange ao aspecto de recuperação da informação, observou-se que o DSpace UFPR permite a recuperação da informação em sua página inicial por meio dos mecanismos denominados: busca, buscar no DSpace, pesquisa

avançada e percorrer por comunidades e coleções, por data, por autor, por título e por assunto.

- a) *Busca*: este formato permite realizar uma busca por um termo simples, ou seja, uma única palavra; em sua segunda tela, a busca pode ser realizada em todo o repositório ou em suas comunidades, definindo assim a quantidade de resultados por página, permitindo que retornem por relevância, data de submissão, título e data de submissão.
- b) *Buscar no DSpace: pesquisa avançada*: permite buscar em todo o DSpace, ou na seleção por suas comunidades. Possibilita combinar termos de busca utilizando os operadores “e”, “ou” e “não”. Neste caso, a busca pode ser por: palavra-chave, autor, título, assunto, resumo, série, patrocinador, identificador, idioma. Assim como na Busca, a recuperação da informação apresenta os dados organizados da seguinte forma: data de publicação, título e autores.
- c) *Comunidades e coleções*: permite a busca por palavra, ou pela seleção por assunto, título autor e data. A seleção da comunidade-coleção também é permitida.
- d) *Data*: permite que seja escolhido um ponto no índice (por ano e por mês), ou que seja feita a opção por ano. Pode-se determinar a classificação por data de criação, data de submissão e título ou optar por apresentar os resultados em ordem ascendente, descendente ou por número de resultados por página.
- e) *Autor*: a pesquisa por autor apresenta o alfabeto para seleção de uma letra específica e também a digitação das primeiras letras do sobrenome do autor a ser pesquisado, possibilitando a ordenação ascendente ou descendente, e a quantidade de resultado por página. Ainda apresenta um cabeçalho em ordem alfabética (sobrenome, nome) dos autores, ao selecionar o autor, o repositório remete para os trabalhos relacionados.
- f) *Título*: permite a seleção por letras iniciais, sendo possível ordenar por data de criação, data de submissão e título, de forma ascendente ou descendente, bem como pelo número de resultados por página.
- g) *Assunto*: apresenta em forma de glossário vários termos, que ao serem selecionados, remetem para os registros do repositório. Há também a opção pela seleção da letra inicial ou letras iniciais do assunto.

A respeito da interoperabilidade semântica do repositório são apontados pela respondente quando menciona a existência da utilização de tesouros e outros instrumentos, que combinados com a utilização de termos livres possibilitam a representação do documento, viabilizando sua recuperação por meio dos mecanismos de busca.

Uma outra questão de extrema relevância para a manutenção de um repositório institucional diz respeito à interoperabilidade humana, ou seja, à sensibilização dos envolvidos (pesquisadores, docentes, servidores, etc.) e às questões relativas ao acesso aberto e aos repositórios institucionais. Assim, o resultado obtido nesta entrevista indica que a avaliação da interoperabilidade humana na UFPR, no que tange à manutenção do repositório institucional, de acordo com a respondente *“é muito boa, há envolvimento das equipes e diálogo entre elas, e sempre recebemos pronto atendimento às solicitações enviadas às unidades e/ou pessoas externas ao SiBi”*.

A recuperação da informação no Repositório Lume é realizada por meio da pesquisa simples, pesquisa avançada, pesquisa por todo o repositório, ou por suas comunidades (ou seja: neste repositório, por tipologia documentária) e por coleções (área de conhecimento), bem como por autor, título e ano.

- a) *Pesquisa simples*: busca no repositório geral ou nas suas comunidades e coleções, permite ordenar os resultados de acordo com o número de ocorrências por página, por relevância, data de envio, data de submissão e por título, apresentando-os em ordem ascendente e descendente.
- b) *Pesquisa avançada*: ocorre da mesma forma que na pesquisa simples, no entanto, amplia-se a possibilidade de combinar termos por meio dos operadores *booleanos*, para a busca composta a um determinado assunto.
- c) *Busca por comunidades e coleções*: possibilita a escolha exata da tipologia documentária que mantenha a informação desejada.
- d) *Percorrer por título*: permite a seleção alfabética localizada na parte superior da tela, bem como pela digitação das primeiras letras da palavra a ser buscada.
- e) *Percorrer por autor*: possibilita a seleção da inicial do sobrenome (utilizando-se do alfabeto) ou pela digitação das primeiras letras do sobrenome. O resultado se apresenta em ordem alfabética (sobrenome, nome). Ao clicar no nome do autor abre uma tela que remete para os títulos relacionados.

- f) *Percorrer por palavra-chave*: ocorre da mesma forma que os itens anteriores, ou seja, em ordem alfabética ou pela digitação das primeiras letras da palavra que se quer buscar.
- g) *Percorrer por ano*: oferece a opção de seleção de mês e ano da publicação e também um campo para digitação de ano (se necessário).

Cabe mencionar que, tanto as coleções, quanto os resultados de pesquisa do Lume, organizam-se por entrada recente. Infere-se que esta opção permite ao usuário do repositório atualização constante acerca dos temas ali depositados.

De acordo com a especificidade de cada comunidade, o repositório permite organização diferente; ou seja: no caso dos Artigos de periódicos, a organização dos itens ocorre por áreas do conhecimento e recupera por mais recente e periódico nacional e estrangeiro e por ano; na comunidade Teses e Dissertações as coleções são ordenadas por área do conhecimento, por tipo: mestrado acadêmico, profissional ou tese e por ano; a comunidade TCC de especialização ordena por entrada recente e por ano; a comunidade TCC de graduação ordena por curso, ano e entrada recente; a comunidade Trabalho em eventos ordena suas coleções por tipo de trabalho e entrada recente.

Quanto à sensibilização da instituição, no concernente à iniciativa dos arquivos abertos, especificamente o caso dos repositórios institucionais, constata-se que a UFRGS tem avançado, mas, segundo a entrevistada “*temos muito a fazer visto que o repositório é relativamente recente*”. De acordo com ela, a administração superior propicia a participação da equipe gestora nos fóruns dos coordenadores de pós-graduação e graduação para a divulgação do mesmo. Além disso, é necessário um trabalho com as Coordenações de curso, com as Secretarias Acadêmicas, e demais órgãos envolvidos.

Em relação às bibliotecas da UFRGS, realizam-se reuniões e seminários nos quais são mostradas as novas implementações, ou seja, o repositório tem sido amplamente divulgado, até para que possam divulgá-lo no âmbito de suas unidades.

Um repositório precisa envolver as pessoas. A certeza é que, para que haja maior participação, é necessário o envolvimento dos indivíduos. “*A administração cria um ambiente favorável, tem interesse no repositório, mas temos ainda que avançar, melhorar o fluxo, divulgar, precisamos dispor de mais tempo*”

para envolver mais pessoas". A respondente informa que gradativamente tem sido possível mobilizar os docentes, os bibliotecários e a administração central.

Para que este envolvimento pudesse ocorrer de forma institucionalizada, faltava à UFRGS a política de informação do Lume. Acredita-se que após ser aprovada, os mecanismos de divulgação do repositório tornar-se-ão ainda mais efetivos. *"Pretende-se que o Reitor realize um evento de lançamento/divulgação do repositório, tendo em vista que, com a política de informação aprovada, as pessoas passem a conhecer melhor o que pode ser disponibilizado"*.

O RIUnB permite diversas formas de busca e recuperação da informação nele depositada, dentre as quais se pode citar:

- a) *Pesquisa rápida*: permite buscar o termo em todo o repositório, ou que a pesquisa ocorra por comunidade. É possível, neste tipo de pesquisa, utilizar-se de palavras-chave, autor, ou palavras do título para busca à informação. A ordenação dos resultados pode ser por número de resultados por página, os itens podem ser ordenados por relevância, por título, por data de publicação e submissão e ainda, que os resultados se apresentem em ordem crescente ou decrescente.
- b) *Pesquisa avançada*: Este tipo de pesquisa, assim como a simples, permite pesquisar em todo o repositório, ou que a pesquisa ocorra por comunidade. Também é possível neste tipo de pesquisa utilizar palavras-chave, autor, ou palavras do título para busca à informação. No entanto, as opções de seleção ampliam-se à medida em que se acrescentam as seguintes opções: autor, título, resumo, assunto, séries, contribuidor, identificador e língua, sendo possível utilizar dos operadores "e", "ou" e "não".
- c) *Busca por autor e orientador*: A busca por autor e orientador pode ocorrer ao optar pela escolha alfabética ou pela digitação de um termo de busca. Apresenta um índice de sobrenome, nome de autores, levando em consideração as primeiras letras da palavra/nome digitada. Pode optar para que o resultado retorne em ordem crescente ou decrescente, bem como restringir ou expandir o número de resultados por página. Ao clicar no nome do autor e/ou orientador, remete para o título, bem como para o documento em texto completo.

- d) *Data de publicação*: A busca por data de publicação possui as seguintes opções de seleção: “pular para um ponto no índice” – ao escolher um ano ou um mês para a busca. Possui também um campo para a digitação do ano. Há opção de ordenar por título, data de submissão e data de publicação.
- e) *Busca por título*: segue a mesma sequência da busca por autor/orientador, onde é possível pesquisar por ordem alfabética, selecionando a letra ou digitando um termo de busca. A opção de ordenação pode ocorrer por título, data de publicação ou submissão, ordenando-o de forma crescente ou decrescente.
- f) *Índice de assunto*: retorna um glossário de termos que remetem aos documentos existentes.

Ao buscar a obtenção de dados a respeito da interoperabilidade semântica e humana do RIUnB, observou-se que este repositório utiliza vocabulário controlado para indexação de seus conteúdos, de acordo com a respondente, *“utilizamos o vocabulário controlado do pergamum como apoio, porém não é seguido com rigidez. Combinamos a indexação com uma linguagem natural, com palavras-chave do próprio autor e dos bibliotecários”*.

A respeito da interoperabilidade humana, a avaliação da gestora indica que os docentes da UnB têm se mostrado bastante receptivos quanto ao repositório. *“Eles têm muito interesse em divulgar suas obras. Até o momento não houve grandes resistências pelo fato de estarem disponibilizando documentos em um repositório de acesso aberto. A aceitação é muito boa, até mesmo por parte dos editores, que têm sempre autorizado essa divulgação”*.

A grande resistência que encontramos é o que diz respeito ao autoarquivamento, pois os docentes não querem ter o trabalho de depositar seus documentos. Quanto aos funcionários da BCE não há resistência e o trabalho flui bem”.

Quem está implantando um repositório precisa ter em mente que este é um trabalho contínuo e que demanda um grupo multidisciplinar de indivíduos trabalhando para que funcione de forma correta. Sob este aspecto, a gestora menciona que *“Criar um repositório não é apenas instalar um software e saber lidar com ele, mas envolve toda uma teoria e toda uma política de gerenciamento que não é simples. Porém, é um trabalho gratificante e que tem tido cada vez mais aceitação por parte da comunidade acadêmica”*.

Finalizando a análise da quarta categoria, interoperabilidade semântica e humana, observou-se que todos os repositórios utilizam linguagem controlada combinada com a linguagem natural, sendo que a primeira possibilita o controle da qualidade e a segunda a expansão e alcance dos termos usualmente utilizados. Quanto aos mecanismos de busca, identificou-se que os três repositórios utilizam basicamente as mesmas opções: pesquisa simples, pesquisa avançada, pesquisa por todo o repositório, ou por suas comunidades e por coleções bem como, por autor, título e ano.

No que tange à questão da interoperabilidade humana, as três instituições avaliam positivamente a cooperação e interesse da comunidade acadêmica para a disponibilização de seus documentos em acesso aberto. O suporte institucional é também fator de extrema relevância para estes repositórios, caso não houvesse apoio e abertura para que esta ferramenta fosse implantada, certamente a visibilidade da instituição e de seus pesquisadores não estaria surtindo tantos resultados quanto estão neste momento. Cabe mencionar que embora a maioria destes repositórios ainda não possua uma política formalizada, conseguiram obter respaldo institucional, sobretudo pelo conhecimento dos profissionais que os gerenciam, e tornaram os repositórios institucionais pioneiros no Brasil, fator que confere aos mesmos um *status* que os diferenciam dos demais que o seguiram.

Acredita-se que muitos projetos de implantação de repositórios buscaram em experiências pioneiras um norte para a definição de suas estruturas. Este estudo buscou, por meio da análise dos resultados, uma oportunidade de apresentar as ações necessárias para a implementação de repositórios institucionais, bem como propor uma estrutura de organização da informação para a UTFPR.

Com o objetivo de estruturar as ações necessárias para a definição do modelo de organização da informação para a UTFPR, apresentam-se os elementos essenciais a serem considerados na implantação de um repositório. Entretanto, a ordem em que os elementos são apresentados não significa sua relevância, principalmente porque muitas das ações elencadas ocorrem simultaneamente, ou seja, uma ação depende da outra.

Após obter dados de cada uma das categorias de análise, apresenta-se a sistematização das principais ações para implantação de um repositório institucional:

Identificar a necessidade de gestão da informação institucional
Buscar informações a respeito de repositórios já implantados e se possível trocar experiências
Buscar mecanismos de apoio à implantação (editais, consórcios, parcerias)
Identificar parcerias para a execução dos trabalhos (comitê gestor)
Sensibilizar o alto escalão administrativo e a comunidade interna de forma contínua
Envolver a equipe de bibliotecários e informáticos em suas atividades
Prever custos (financeiros e de recursos humanos)
Institucionalizar a proposta do repositório
Definir o <i>software</i> a ser utilizado e customizá-lo
Elaborar e explicitar uma política de informação institucional
Identificar o fluxo de informação institucional e definir quem realizará a coleta dos itens ou definir se haverá autoarquivamento
Definir uma política de funcionamento para o repositório (objetivos do repositório, estrutura de organização da informação, tipologia documentária a ser inserida, definição de metadados, uso de linguagem natural ou controlada, fluxo de submissão dos itens, responsáveis pelas comunidades, os aspectos legais e de direitos autorais e outros)
Identificar se há interconexão entre o repositório e o sistema gerenciador da informação da biblioteca
Criar mecanismos que garantam a permanência, sustentabilidade e ampliação do repositório
Elaborar mecanismos eficientes de busca à informação depositada no repositório
Registrar o repositório em um diretório visando ampliar o seu alcance e visibilidade
Garantir o acesso aberto e a interoperabilidade dos dados

Quadro 7 - Ações para implementação de repositórios institucionais

Fonte: Elaborado pela autora.

7.2 PROPOSTA DE ESTRUTURA DE ORGANIZAÇÃO DA INFORMAÇÃO PARA O REPOSITÓRIO DA UTFPR

A sistematização das categorias de análise proposta no item anterior possibilitou que a UTFPR seguisse algumas diretrizes apresentando neste capítulo uma proposta de estrutura de organização da informação para seu repositório institucional.

De posse do *kit* tecnológico fornecido pelo IBICT, coube à UTFPR direcionar a customização do *software*, definir a estrutura de organização das comunidades e coleções, identificar os metadados que representarão cada uma das tipologias documentárias e definir uma política de informação para respaldar institucionalmente o repositório e garantir sua permanência.

A política de informação da UTFPR baseou-se no modelo proposto por Kuramoto e em outras políticas registradas no ROARMAP⁸. A elaboração da política de informação da UTFPR permitiu a integração de representantes de cada um dos onze *campi* da universidade, para que fosse possível elaborar um documento pautado no exercício democrático de estabelecimento de ações que pudessem contemplar, na maior parte, as necessidades existentes em cada uma das regiões.

Acredita-se que a política de informação da UTFPR tenha sido estabelecida de forma participativa e com intuito de atender às exigências explícitas no edital do IBICT, bem como, com a intenção de promover o acesso aberto à produção intelectual desta comunidade.

Cabe mencionar que nesta mesma política define-se a criação de um portal de informação institucional, onde estarão o Repositório Institucional (depósito de documentos revisados por pares) que comporá o catálogo nacional (por meio do IBICT), o Repositório de Outras Publicações Abertas (imagens, materiais didáticos, vídeos, etc) e o SEER (Sistema de Editoração Eletrônica de Revistas). Esta definição amplia sobremaneira a questão do acesso à informação produzida pela comunidade UTFPR, bem como o acesso às demais publicações disponíveis em acesso aberto.

Por tratar-se de um documento de tamanha complexidade, a definição desta política mobilizou diversos segmentos da instituição, sobretudo as Pró-Reitorias de Graduação e Pós-graduação, o Sistema de Bibliotecas e a Diretoria de Gestão de Tecnologia da Informação. Atualmente, esta política tramita pelos órgãos competentes, para que, após aprovada, seja cumprida.

As etapas de customização do *software* Dspace e a definição da estrutura de organização do repositório institucional estão a cargo da Diretoria de Gestão de Tecnologia da Informação e do Sistema de Bibliotecas. No entanto, por ser interesse deste estudo aplicar os resultados obtidos, a equipe gestora da UTFPR permitiu que a estrutura inicial pudesse ser proposta inicialmente, neste estudo.

Uma instituição que pretenda possibilitar o acesso e visibilidade de sua produção deve estruturá-la no intuito de refletir essencialmente os níveis hierárquicos nela existentes, desde que atenda às necessidades daquele que busca pelas informações ali depositadas.

⁸ Disponível em: <<http://www.eprints.org/openaccess/policysignup>>.

Assim, atendendo ao objetivo do Repositório Institucional da UTFPR de “reunir, armazenar, preservar e permitir acesso às produções científicas da UTFPR” surge como proposta inicial para sua organização a seguinte estrutura de alta generalização (Comunidades): por categorias de ensino: Cursos de Educação Profissional Técnica de Nível Médio, Cursos de Graduação, Programas de Pós-Graduação *Lato-Sensu* (Especializações), Programas de Pós-Graduação *Stricto-sensu* (Mestrados e Doutorados) e Sistema de Bibliotecas (abará as produções intelectuais dos membros da comunidade: docentes e técnicos administrativos realizadas fora do âmbito da UTFPR). Conforme demonstra a figura 12:

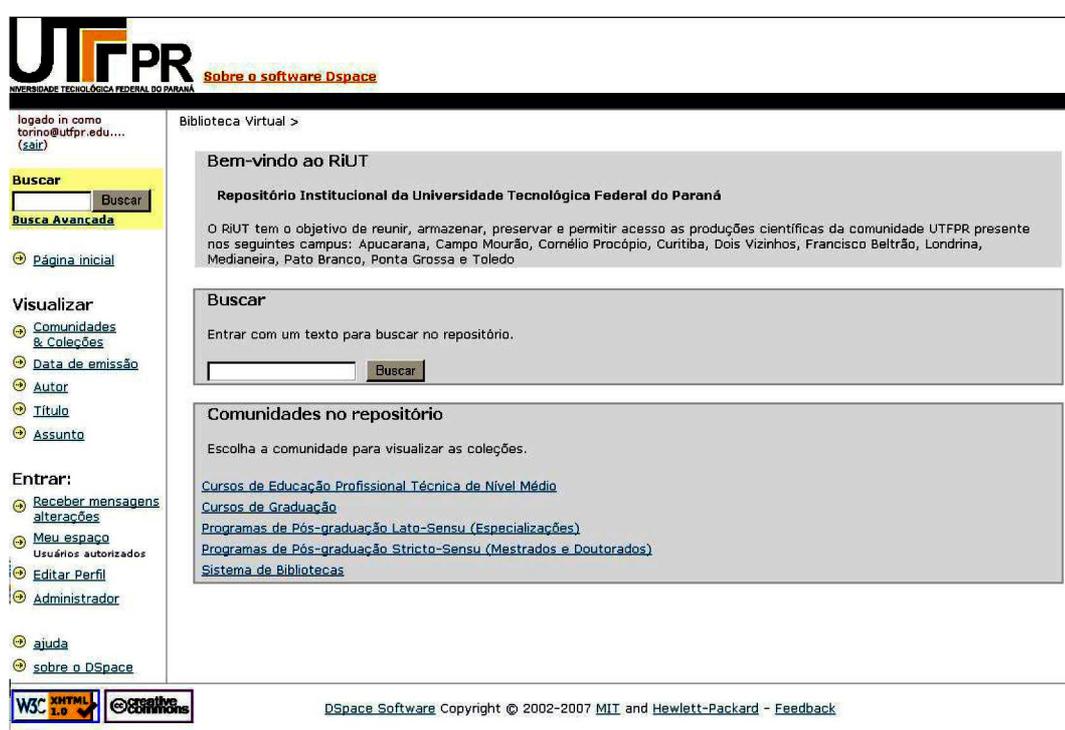


Figura 12 - Estrutura de organização do Repositório da UTFPR

Fonte: Elaborado pela autora.

Acredita-se que esta estrutura represente o objetivo maior da instituição que é ofertar educação formal em todos os segmentos acima mencionados. Cabe ressaltar que, a opção pela não categorização por *campi*, deu-se como forma de propor uma estrutura mais unificada, buscou-se englobar a UTFPR de forma ampla, ou seja, em uma visão “sistêmica”.

De acordo com a definição da política institucional da UTFPR, comporão o repositório institucional apenas as publicações validadas por pares. Inicialmente, a proposta é que as Comunidades sejam organizadas por categoria de ensino; e as Coleções, organizadas de acordo com as tipologias documentárias

passíveis de revisão por pares: livros e capítulos de livros, artigos de periódicos, trabalhos em eventos, Monografias, Teses e Dissertações.

A fim de minimizar esforços nesta etapa, optou-se por iniciar a submissão pelos documentos que já se encontram em formato digital, ou seja, os itens depositados na BDTD da instituição, sobretudo porque estes já possuem a autorização do autor. Posteriormente, julga-se necessário definir um projeto piloto, buscando algum programa ou coordenação com interesse no repositório.

Como todas as coleções serão basicamente estruturadas com as mesmas tipologias documentárias, optou-se por utilizar siglas para diferenciá-las. A sigla EPT refere-se às coleções da comunidade Educação Profissional Técnica; a GRAD pertence aos cursos de graduação; a ESP pertence à comunidade Programas de Pós-Graduação (especialização); as siglas MEST e DOUT pertencem ao Programas de Pós-Graduação (mestrado e doutorado) e a sigla SIBI pertence ao Sistema de Bibliotecas da UTFPR, conforme demonstra a figura abaixo:

UTFPR Sobre o software DSpace
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

logado in como torino@utfpr.edu... (sair)

Biblioteca Virtual >

Comunidades e Coleções

Apresentado abaixo a lista de Comunidades, Sub-Comunidades e Coleções. Clique no termo desejado para visualizar a página da Comunidade ou Coleção.

- **Cursos de Educação Profissional Técnica de Nível Médio**
 - o [EPT - Artigos de Periódicos](#)
 - o [EPT - Livros e Capítulos de livros](#)
 - o [EPT - Trabalhos em Eventos](#)
- **Cursos de Graduação**
 - o [GRAD - Artigos de Periódicos](#)
 - o [GRAD - Livros e Capítulos de livros](#)
 - o [GRAD - Trabalhos de Conclusão de Curso](#)
 - o [GRAD - Trabalhos em Eventos](#)
- **Programas de Pós-graduação Lato-Sensu (Especializações)**
 - o [ESP - Artigos de Periódicos](#)
 - o [ESP - Livros e Capítulos de livros](#)
 - o [ESP - Trabalho de Conclusão de Curso](#)
 - o [ESP - Trabalhos em Eventos](#)
- **Programas de Pós-graduação Stricto-Sensu (Mestrados e Doutorados)**
 - o [DOUT - Artigos de Periódicos](#)
 - o [DOUT - Livros e Capítulos de Livros](#)
 - o [DOUT - Teses](#)
 - o [DOUT - Trabalho em Eventos](#)
 - o [MEST - Artigos de Periódicos](#)
 - o [MEST - Dissertações](#)
 - o [MEST - Livros e Capítulos de Livros](#)
 - o [MEST - Trabalho em Eventos](#)
- **Sistema de Bibliotecas**
 - o [SIBI - Artigos de periódicos](#)
 - o [SIBI - Dissertações](#)
 - o [SIBI - Livros e capitulos de livros](#)
 - o [SIBI- Teses](#)
 - o [SIBI - Trabalhos de conclusão de curso](#)
 - o [SIBI - Trabalhos publicados em eventos](#)

W3C XHTML 1.0 creative commons
DSpace Software Copyright © 2002-2007 MIT and Hewlett-Packard - Feedback

Figura 13 - Comunidades e coleções do Repositório da UTFPR
Fonte: Elaborado pela autora.

Sugere-se para a organização da informação a utilização dos seguintes instrumentos: AACR2, CDD, Vocabulário controlado do Pergamum e *Library Congress* que combinados com a linguagem natural serão utilizados para representação das informações da UTFPR. No que tange ao aspecto relacionado à definição dos metadados, o Sistema de Bibliotecas da UTFPR, possivelmente partirá do uso do padrão MARC, por tratar-se do mesmo padrão utilizado no sistema Pergamum, compatibilizando-o com o padrão Dublin Core presente no DSpace.

Acredita-se que utilizar os metadados já existentes no Pergamum facilitaria o trabalho dos bibliotecários do SIBI UTFPR, possibilitando maior qualidade na indexação, tendo em vista que se trata de um padrão já utilizado por todos. Recomenda-se que os nomes dos cursos e programas, bem como o *campus* onde o documento fora produzido, sejam informados nos campos de metadados indicados para esta finalidade para que seja possível recuperá-los.

Conforme mencionado anteriormente, os metadados deverão ser definidos conforme necessidade de cada uma das tipologias documentárias. Utilizando-se do padrão MARC, encontrado no sistema Pergamum, é possível mencionar quais os elementos deverão ser utilizados para representar cada uma das tipologias documentárias existentes no repositório.

Os metadados utilizados para a tipologia “Livros” são: Entradas secundárias/autor, título principal, local de Publicação, descrição física, ISBN e Assuntos. Para descrever “capítulos de livros” os campos são: autor principal, título principal, descrição física e assuntos. Para os “Trabalhos em eventos”: Autor principal, entradas secundárias/autor, título principal, outros títulos, local de publicação, descrição física e assuntos. Para descrever as “Teses e Dissertações”: Autor principal, entradas secundárias/autor, título principal, local de publicação, descrição física, notas gerais, notas de dissertações/teses, notas de bibliografia, notas gerais, resumo e assunto. Os “Artigos de periódicos” podem ser representados pelos metadados de: Autor do artigo, entradas secundárias/autor, título do artigo, analítica, notas, assuntos e os “Trabalhos de conclusão de curso” pelas entradas: Autor principal, entradas secundárias/autor, título principal, local de publicação, descrição física, notas gerais, notas de dissertações/teses, notas de bibliografia, notas de resumo e assuntos.

Acredita-se que a indexação é uma etapa de valiosa importância. Desta forma, propõe-se que a instituição adote uma política para garantir a

qualidade dos metadados utilizados como forma de garantir a satisfação do usuário no momento da busca à informação, ao qual se recomenda a utilização de vocabulários controlados, que, em conjunto com a linguagem natural, possam possibilitar maior satisfação do usuário.

No que tange ao acesso à informação, sugere-se que a equipe de tecnologia da informação da UTFPR elabore filtros de busca que permitam cruzar dados como: produções depositadas por campus, por curso, *downloads*, entre outros. Alguns destes mecanismos já estão em teste, dentre eles pode-se citar: busca por todo o repositório, busca por comunidade, por autor, título, palavra-chave, assunto e resumo, inclusive com a possibilidade de utilizar os operadores *booleanos*.

Quanto ao fator interoperabilidade, a questão semântica possivelmente pode ser resolvida com a utilização de formulários de padronização (destinado a bibliotecários) e pelos mecanismos de busca. Já a interoperabilidade humana, acredita-se que seja este o grande desafio para esta instituição, sobretudo para a equipe gestora do repositório. O trabalho de esclarecimento, divulgação e permanência do repositório depende muito das pessoas e do grau de envolvimento que elas dedicam às atividades.

Assim, o primeiro passo é “institucionalizar” a proposta do repositório, ou seja, demonstrar o quanto ele é viável e necessário para a instituição. Acredita-se que esta sensibilização deve ocorrer, inicialmente, no sistema de bibliotecas. Desta forma, a equipe de bibliotecários estará apta a multiplicar as informações acerca desta ferramenta junto à comunidade acadêmica do campus ao qual pertence e também estará apta a contribuir com a divulgação do repositório e com o fluxo de produção intelectual institucional, coletando, registrando e permitindo a disseminação das informações obtidas.

Concomitantemente, deve ocorrer a sensibilização junto à administração superior, traçando estratégias que agreguem docentes e profissionais com interesse em compartilhar o conhecimento por meio dos repositórios institucionais, favorecendo, assim, a atualização e manutenção do repositório institucional.

CONSIDERAÇÕES FINAIS

Apresentar uma proposta de organização para uma instituição com característica tão peculiar quanto à UTFPR é uma tarefa bastante desafiadora, sobretudo pela condição que esta instituição assume ainda em fase de ampliação (tanto em termo de cursos, quanto em quantidade de *campi*), e de estruturação administrativa, a qual em meio a tantas definições a serem tomadas, apóia as iniciativas de acesso aberto, possibilitando ainda mais a obtenção de informações e a geração de conhecimento de seus acadêmicos e servidores.

Ao finalizar a análise de cada uma das estruturas encontradas nas instituições em estudo, os questionamentos a respeito de qual seria a melhor estrutura para a UTFPR persistiram, sobretudo porque percebe-se que a melhor estrutura de organização é aquela que permite representar tanto a estrutura da instituição, quanto a necessidade daquele que busca pelo conhecimento ali produzido. Esta estrutura deve estar aberta a modificações, tendo em vista o processo dinâmico de uma instituição de ensino. Ainda, as informações devem estar organizadas visando ao acesso aberto, não somente o acesso à comunidade interna.

Embora possa parecer bastante simples, a estruturação de um sistema de gestão da informação como um repositório exige trabalho criterioso, no qual as definições institucionais devem ocorrer respaldadas em decisões da equipe gestora, possibilitando a integração entre diversas áreas/setores e permitindo que o trabalho e o conhecimento de cada uma das pessoas envolvidas agreguem ao repositório ainda mais qualidade. O repositório deve ser “institucionalizado”; não pode ser implementado de forma “isolada” ou seja, ele não é responsabilidade de um setor apenas, mas de toda a instituição.

Este estudo, ao apresentar teoricamente os conteúdos que combinados com a prática, permite destacar a importância de cada um dos elementos anteriormente discutidos, ou seja: é necessário que um repositório institucional seja respaldado por uma política aprovada institucionalmente, é necessária a mobilização da comunidade universitária, é necessário definir critérios de organização da informação, bem como os fluxos da informação na instituição. E estas tarefas são bastante complexas.

O projeto de instalação de um repositório na UTFPR, embora tenha sido planejado inicialmente por uma equipe pequena composta por bibliotecárias, um docente e sua orientanda (vinculados ao Programa de Mestrado em Tecnologia da UTFPR) ampliou-se na instituição após ter sido aprovado pelo edital do IBICT. Os processos de elaboração da política de informação institucional, que atualmente encontra-se em trâmite para aprovação pela Pró-Reitoria de Graduação (PROGRAD) e Pró-Reitoria de Pesquisa e Pós-graduação (PROPPG), bem como a customização do *software*, permitiram aos envolvidos no projeto, além de maior integração e possibilidade de novos conhecimentos, a oportunidade de demonstrar à comunidade UTFPR a relevância da organização para a gestão da informação institucional.

Em meados de 2010 Comitê Gestor da UTFPR é articulado, reunindo diferentes profissionais com o objetivo de apoiar a implantação e definição de políticas para o repositório.

Atualmente a Reitoria da UTFPR, por meio da Assessoria de Comunicação, das Diretorias de Graduação e Pós-Graduação, do Sistema de Bibliotecas e da Assessoria em Tecnologia de Informação, apóiam a consolidação deste projeto, bem como as iniciativas de acesso aberto junto à comunidade acadêmica.

A etapa de implementação é a de maior complexidade em termos de definições das ações a serem realizadas, entretanto, a responsabilidade social da instituição na manutenção, divulgação e continuidade do projeto é uma etapa bastante relevante a ser discutida pela equipe gestora daqui em diante.

Aliar as questões teóricas à realidade profissional pode ser tarefa árdua e desafiadora, entretanto, ao visualizar os resultados obtidos tanto para a instituição, colaborando com a gestão da informação produzida por ela, quanto para os profissionais envolvidos nas etapas de estudo e execução de um projeto como este, possibilitando a transformação do meio onde se encontra inserido, é algo muito valioso. Sobretudo porque “é necessário que as ciências humanas sejam auto-reflexivas, ou seja, que sejam capazes de se pesquisar como pesquisam os outros. É desafiador sim, mas necessário se queremos transformar o mundo, mesmo que só um pouquinho” (MERKLE, 2010).

Por tratar-se de um assunto complexo e permeado por inúmeras facetas, este estudo conseguiu obter dados relevantes à determinação de uma

estrutura de organização da informação para a UTFPR, tendo em vista principalmente o contato permitido pelos gestores das instituições que já haviam passado pela experiência de implantação de seu repositório institucional.

No entanto, acredita-se que novos estudos poderão ampliar esta discussão, ao observar os critérios adotados pelas instituições para manter em constante atualização seus repositórios.

REFERÊNCIAS

- ALVARENGA, L. Organização da informação nas bibliotecas digitais. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006. Cap. 6, p. 76-98.
- _____. A teoria do conceito revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. *DataGramaZero: Revista de Ciência da Informação*, Rio de Janeiro, v. 2, n. 6, dez. 2001. Disponível em: <<http://dici.ibict.br/archive/00000309>>. Acesso em: 2 ago. 2009.
- ALVES, V. B. A. Open archives: via verde ou via dourada? *PontodeAcesso*, Salvador, v. 2, n. 2, ago./set. 2008. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/1780/2172>>. Acesso em: 8 jan. 2009.
- ANDRADE, M. E. A. Um olhar sobre representações no universo do conhecimento: o caso das micro e pequenas empresas. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006. Cap. 4, p. 46-61.
- ANZOLIN, H. H. Atualizações em AACR2. In: ENCONTRO NACIONAL DOS USUÁRIOS DA REDE PERGAMUM, 9., 2007, Curitiba. [*Anais...*]. Curitiba: PUCPR, 2007. 1 CD-ROM.
- ARANALDE, M. M. Reflexões sobre os sistemas categoriais de Aristóteles, Kant e Ranganathan. *Ciência da Informação*, Brasília, v. 38, n. 1, jan./abr. 2009. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewArticle/1056>>. Acesso em: 17 abr. 2010.
- ARELLANO, M. A. M.; CAREGNATO, S. E.; FERREIRA, S. M. S. P. Editoração eletrônica de revistas científicas com suporte do protocolo OAI. In: FERREIRA, S. M. S. P.; TARGINO, M. G. (Org.). *Preparação de revistas científicas: teoria e prática*. São Paulo: Reichmann e Autores, 2005. p. 195-229.
- AUN, M. P. Políticas públicas de informação e desenvolvimento. In: PAIM, I. (org.). *A gestão da informação e do conhecimento*. Belo Horizonte: ECI/UFMG, 2003. p. 55-90.
- BARBALHO, R. C. S. Periódico científico: parâmetros para avaliação de qualidade. In: FERREIRA, S. M. S. P.; TARGINO, M. G. (Org.). *Preparação de revistas científicas: teoria e prática*. São Paulo: Reichmann & Autores, 2005. Cap. 5, p. 123-158.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual em bibliotecología y documentación. In: CARRARA, K. (Org.). *Educação, universidade e pesquisa*. Marília: Unesp Publicações, 2001. p. 35-60.
- BARROS, A. J. S.; LEHFELD, N. A. S. *Fundamentos de metodologia científica: um guia para a iniciação científica*. 2. ed. São Paulo: Pearson, 2006.

BARTON, M. R.; WALKER, J. H. Building a business plan for DSpace, MIT Libraries Digital Institutional Repository. *Journal of Digital Information*, Austin, v. 4, n. 2, maio 2003. Disponível em: < <http://journals.tdl.org/jodi/article/view/99/98>>. Acesso em: 15 jul. 2010.

BELLOTTO, H. L. *Como fazer análise diplomática e análise tipológica de documento de arquivo*. São Paulo: Arquivo do Estado, 2002.

BRASIL. Ministério da Educação. Resolução CES nº. 2, de 7 de abril de 1998. Estabelece indicadores para comprovar a produção intelectual institucionalizada, para fins de credenciamento, nos termos do Art. 46 do Art. 52, inciso I, da Lei 9.394/96 de 20 de dezembro de 1996. *Diário Oficial [da] República Federativa do Brasil*, Brasília, 14 abr. 1998. Seção 1, p. 32.

CAFÉ, L.; MELO, B. A.; BARBOSA, E. M. F.; NUNES, E. M. A. N.; MÁRDERO ARELLANO, M. A. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: ENCONTRO NACIONAL DE CENTROS DE INFORMAÇÃO E BIBLIOTECAS DA ÁREA DE COMUNICAÇÃO, 13., 2003, Belo Horizonte. *Anais...* Belo Horizonte, 2003. Disponível em: < <http://www.uepg.br/bicen/repositorio.htm>>. Acesso em: 11 abr. 2009.

CAMPOS, M. L. A. A problemática da compatibilização terminológica e a integração de ontologias: o papel das definições conceituais. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 8., 28-31 out. 2007. Salvador. *Anais...* Salvador, 2008. Disponível em: <www.uff.br/ppgci/editais/marialuizaleituras.pdf>. Acesso em: 28 jan. 2009.

_____. Indexação e descrição em arquivos: a questão da representação e recuperação de informações. *Arquivo e Administração*, Rio de Janeiro, v. 5, p. 17-31, 2006.

_____. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004.

CARUSO, A. T. R.; GUILHEM, C. B.; MATSUMOTO, L. A.; TAVARES, H.; TORINO, L. P. *O sistema de bibliotecas da UTFPR: a primeira universidade tecnológica federal do Brasil*. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 22-27 out. 2006, Salvador. *Anais...* Salvador: UFBA, 2006. Disponível em: <http://www.pergamum.pucpr.br/redepergamum/trabs/SNBU2006_UTFPR.pdf>. Acesso em: 15 jul. 2010.

CAVALCANTI, C. R. *Indexação e tesouro: metodologia e técnicas*. Brasília: Associação dos Bibliotecários do Distrito Federal, 1978.

CERVANTES, B. M. N. *Organização e representação da informação e do conhecimento: repositórios institucionais*. 2009. Londrina. 21 mar. 2009. 18 slides em Power Point. Disciplina Seminários avançados I. Programa de Pós-graduação em Gestão da Informação, Universidade Estadual de Londrina, Londrina, 2009.

CERVANTES, B. M. N.; FUJITA, M. S. L.; RUBI, M. P. Terminologias em políticas de indexação. In: GARCIA MARCO, F. J. (Org.). *Avances y perspectivas en sistemas de*

información y documentación. Zaragoza: Universidad de Zaragoza, 2008, v. 1, p. 211-222.

CHIZZOTTI, A. *Pesquisa em ciências humanas e sociais*. 6. ed. São Paulo: Cortez, 2003.

CINTRA, A. M. M.; TÁLAMO, M. F. G. M; LARA, M. G. L.; KOBASHI, N. Y. *Para entender as linguagens documentárias*. São Paulo: Polis, 2002.

COSTA, S. M. S. Abordagens, estratégias e ferramentas para o acesso aberto via periódicos e repositórios institucionais em instituições acadêmicas brasileiras. *Liinc em Revista*, Rio de Janeiro, v. 4, n. 2, set. 2008. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/view/281>>. Acesso em: 21 maio 2009.

_____. Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica. *Ciência da Informação*, Brasília, v. 35, n. 2, maio/ago. 2006. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/827/669>>. Acesso em: 20 jun. 2009.

COSTA, S.; LEITE, F. C. L. Repositórios institucionais: potencial para maximizar o acesso e impacto da pesquisa em universidades. In: CONFERÊNCIA IBEROAMERICANA DE PUBLICAÇÕES ELETRÔNICAS NO CONTEXTO DA COMUNICAÇÃO CIENTÍFICA, 1., 2006, Brasília. *Anais...* Brasília: Universidade de Brasília, 2006. p. 1-10.

CRESWELL, J. W. *Projeto de pesquisa*. 2. ed. Porto Alegre: Artmed, 2007.

CURRÁS, E. *Ontologias, taxonomia y tesauros: manual de construcción y uso*. Gijón: Ediciones Trea, 2005.

CURTY, R. G.; ARAUJO, N. C.; CERVANTES, B. M. N.; GIRALDES, M. J. C. A organização e a descrição de estruturas de informação em repositórios institucionais acadêmicos: uma análise com base em elementos de arquitetura de informação. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 23., 2009, Bonito. *Anais...* Bonito: FEBAB, 2009. p. 1-12.

DIAS, E. W. Organização do conhecimento no contexto de bibliotecas tradicionais e digitais. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006. Cap. 5, p. 62-75.

DIAS, E. W.; NAVES, M. M. L. *Análise de assunto: teoria e prática*. Brasília: Thesaurus, 2007. (Estudos Avançados em Ciência da Informação, v. 3).

DUARTE, J., BARROS, A. *Métodos e técnicas de pesquisa em comunicação*. São Paulo: Atlas, 2009.

FEITOSA, A. *Organização da informação na web: das tags à web semântica*. Brasília: Thesaurus, 2006. (Estudos Avançados em Ciência da Informação, v. 2).

FOSKETT, A. C. *A abordagem temática da informação*. São Paulo: Polígono, 1973.

FRIEDMAN, T. *O mundo é plano*. São Paulo: Objetiva, 2007.

FUJITA, M. S. L. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. *Tendências da Pesquisa Brasileira em Ciência da Informação*, João Pessoa, v. 1, n. 1, 2008. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/4/13>>. Acesso em: 12 ago. 2009.

FUJITA, M. S. L.; RUBI, M. P. O ensino de procedimentos de política de indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário. *Perspectiva em Ciência da Informação*, Belo Horizonte, v. 11, n. 1, jan./abr. 2006. Disponível em: <<http://www.eci.ufmg.br/pcionline/index.php/pci/article/view/445>>. Acesso em: 24 ago. 2009.

GALVÃO, M C.B. A análise, a síntese, a representação da informação e a gestão do conhecimento nas empresas. In: RODRIGUES, G. M.; LOPES, I. L. (Org.). *Organização e representação do conhecimento na perspectiva da Ciência da Informação*. Brasília: Thesaurus, 2003. p. 230-239. (Estudos avançados em ciência da informação, v. 2).

GIL, A. C. *Como elaborar projetos de pesquisa*. São Paulo: Atlas, 2002.

_____. *Como elaborar projetos de pesquisa*. São Paulo: Atlas, 2007.

_____. *Métodos e técnicas de pesquisa social*. 6. ed. São Paulo: Atlas, 2008.

GIL URDICIAIN, B. *Manual de lenguajes documentales*. 2. ed. Gijón: Trea, 2004.

GUIMARÃES, J. A. C. A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais. In: RODRIGUES, G. M.; LOPES, I. L. (Org.). *Organização e representação do conhecimento na perspectiva da Ciência da Informação*. Brasília: Thesaurus, 2003. p. 100-117. (Estudos Avançados em Ciência da Informação, v. 2).

HOROWITZ, Z.; ARAÚJO NETO, A. C.; SAATKAMP, C. M.; PAVÃO, C. M. G.; COSTA, J. S. B.; CAREGNATO, L. F.; OLIVEIRA, Z. C. P. Repositório digital da Universidade Federal do Rio Grande do Sul. In: WORKSHOP DE TECNOLOGIA DA INFORMAÇÃO DAS IFES, 2., 2008, Gramado. *Resumos...* Gramado: UFRGS, 2008. Disponível em: <<http://www.Lume.ufrgs.br/handle/10183/12730>>. Acesso em: 6 jan. 2009.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. *Acesso livre à informação científica impulsiona desenvolvimento do País*. 2009. Disponível em: <<http://www.ibict.br/noticia.php?id=596>>. Acesso em: 6 jan. 2009.

_____. *Ibict lança manifesto pelo acesso livre à informação científica*. 2005. Disponível em: <<http://www.ibict.br/noticia.php?page=23&id=142>>. Acesso em: 6 jan. 2009.

INTERNATIONAL RESEARCH AND CULTURAL HERITAGE ORGANIZATIONS. *Berlin declaration on open access to knowledge in the sciences and humanities*. 2003. Disponível em: <<http://oa.mpg.de/index.html>>. Acesso em: 6 jan. 2009.

JORENTE, M. J. V.; SANTOS, P. L. V. A. C. Mídias de pós-vanguarda, direito de autor, cultura livre e produtos de criação contemporânea. In: GUIMARÃES, J. A. C.; MOLINA, J. C. F. (Org.). *Aspectos jurídicos e éticos da informação digital*. Marília: FUNDEPE, 2008, p. 113-131.

KNELLER, G. F. *A ciência como atividade humana*. Rio de Janeiro: Zahar, 1980.

KURAMOTO, H. Acesso livre: maximizando a visibilidade da pesquisa e dos pesquisadores. Palestra Magna apresentada no III Seminário em Ciência da Informação: tendências para a organização e compartilhamento da informação, 23-25 nov. 2009. Disponível em: <<http://kuramoto.blog.br/>>. Acesso em: 28 nov. 2009.

_____. Acesso livre: um caso de soberania nacional?. In: COLÓQUIO MEDIAÇÕES E USOS DE SABERES E INFORMAÇÃO, 1., 2008, Rio de Janeiro. *Anais...* Rio de Janeiro: MUSSI, 2008. p. 167-181.

_____. Biblioteca digital brasileira: integrando a ICT brasileira. In: MARCONDES, C. H. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. Brasília: IBICT, 2006a. p. 287-303.

_____. *IBICT estimula adoção dos open archives no Brasil*. 2005. Disponível em: <<http://www.ibict.br/noticia.php?id=148> >. Acesso em: 10 jan. 2009.

_____. Sintagmas nominais: uma nova abordagem no processo de indexação. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006b. Cap. 8, p. 117-137.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos, 2004.

LANGRIDGE, D. *Classificação: abordagem para estudantes de biblioteconomia*. Rio de Janeiro: Interciência, 2006.

LARA, M. L. G. Termos e conceitos da área de comunicação e produção científica. In: POBLACION, D. A.; WITTER, G. P.; SILVA, J. F. M. *Comunicação & produção científica: contexto, indicadores e avaliação*. São Paulo: Angellara, 2006. p. 387-414.

LEITE, F. C. L. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Brasília: IBICT, 2009. Disponível em: <<http://www.ibict.br/noticia.php?id=665>>. Acesso em: 6 nov. 2009.

LEMOS, R. *Creative Commons: o que é e modo de usar*. 2005. Disponível em: <http://www.creativecommons.org.br/index.php?option=com_content&task=view&id=42&Itemid=80>. Acesso em: 17 fev. 2010.

LIMA, G. A. B. Organização da informação para sistemas de hipertextos. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006. Cap. 7, p. 98-116.

LOPES, I. L. Uso das linguagens controlada e natural em base de dados: revisão da literatura. *Ciência da informação*, Brasília, v. 31, n. 1, jan. 2002. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-19652002000100005&script=sci_arttext>. Acesso em: 20 ago. 2009.

LYNCH, C. A. Institutional repositories: essential infrastructure for scholarship in the digital age. *ARL Bimonthly Report*, Washington, n. 226, 2003. Disponível em: <<http://www.arl.org/resources/pubs/br/br226/br226ir~prints.html>>. Acesso em: 17 dez. 2008.

MACHADO, M. M. *Open archives: panorama dos repositórios*. 2006. 101 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Santa Catarina, Florianópolis, 2006. Disponível em: <www.cin.ufsc.br/pgcin/Murilo.pdf>. Acesso em: 22 jan. 2010.

MARQUES, A. M. N.; MAIO, S. R. S. *Repositórios institucionais*. 2006. Disponível em: <<http://repositoriosdigitais.web.simplesnet.pt/PDF%27S/Artigo%20%20Repositorios%20Institucionais.pdf>>. Acesso em: 5 jun. 2009.

MARTINS, G. A. *Estudo de caso: uma estratégia de pesquisa*. São Paulo: Atlas, 2006.

MCGEE, J.; PRUSAK, L. *Gerenciamento estratégico da informação*. Rio de Janeiro: Campus, 1995.

MEADOWS, A. J. *A comunicação científica*. Brasília: Briquet de Lemos, 1999.

MEDEIROS, M. B. Documento. In: CID-UNB. *Glossário geral de Ciência da Informação*. Disponível em: <http://www.cid.unb.br/M452/M4522012.ASP?txtID_PRINCIPAL=2>. Acesso em: 12 fev. 2009.

MÉNDEZ RODRÍGUEZ, E. M. *Metadatos y recuperación de información: estándares, problemas y aplicabilidad em bibliotecas digitales*. Gijón: Trea, 2002.

MERKLE, L. E. *Estrutura de informação do repositório da UTFPR* [mensagem pessoal]. Mensagem recebida por <lptorino@gmail.com> em 13 jun. 2010.

_____. O interagir humano-computacional: mapeando relações heterodisciplinares. *DatagramaZero: Revista de Ciência da Informação*, Rio de Janeiro, v. 1, n. 2, abr. 2000. Disponível em: <http://dgz.org.br/abr00/Art_02.htm>. Acesso em: 22 out. 2009.

MIRANDA, M. L. C. *Organização e representação do conhecimento: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes digitais*. 2005. 353 f. Tese (Doutorado em Ciência da Informação) – Instituto Brasileiro em Informação em Ciência e Tecnologia, Rio de Janeiro. Disponível em: <<http://biblioteca.ibict.br/phl8/anexos/miranda2006.pdf>>. Acesso em: 7 fev. 2008.

MONTEIRO, F. S. *Organização da informação em repositórios digitais institucionais com ênfase na descrição física e descrição temática*. 2008. 199 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília. Disponível em: <<http://repositorio.bce.UnB.br/handle/10482/1096>>. Acesso em: 3 maio 2009.

MORENO, F. P.; LEITE, F. L.; ARELLANO, M. A. Acesso livre a publicações e repositórios digitais em ciência da informação no Brasil. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 11, n. 1, jan./abr. 2006. Disponível em: <<http://www.eci.ufmg.br/pcionline/index.php/pci/article/view/447/258>>. Acesso em: 1 jun. 2009.

MORI, A.; CARVALHO, C. L. *Metadados no contexto da web semântica: relatório técnico*. Goiânia: UFG, 2004.

MUELLER, S. P. M. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. (Org.). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: UFMG, 2000. p. 21-34.

_____. A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, Brasília, v. 35, n. 2, maio/ago. 2006. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewPDFInterstitial/826/668>>. Acesso em: 2 maio 2009.

NAVES, M. M. L. A importância de Ranganathan para a organização do conhecimento. In: NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos, 2006. Cap. 3, p. 36-45.

NONATO, R. S.; LIMA, G. A. B. O. Determinação de links hipertextuais: uma abordagem da Ciência da Informação. *Informação & Sociedade*, João Pessoa, v. 18, n. 1, jan./abr. 2008. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/1726/1648>>. Acesso em: 8 set. 2008.

NÚCLEO DE INFORMAÇÃO, PESQUISAS E ESTUDOS EM EDUCAÇÃO BIBLIOTECÁRIA. *Tipologia de documentos*. 2009. Disponível em: <<http://notes.ufsc.br/aplic/bibced.nsf/tipodoc?OpenNavigator>>. Acesso em: 3 jan. 2009.

OPEN SOCIETY INSTITUTE. *Budapest open access initiative*. 2002. Disponível em: <<http://www.soros.org/openaccess/index.shtml>>. Acesso em: 6 jan. 2009.

PIEIDADE, M. A. R. *Introdução à teoria da classificação*. Rio de Janeiro: Interciência, 1983.

PULCINELLI, E. O. *O que é lingüística*. São Paulo: Brasiliense, 2005.

RAMALHO, R. A. S.; VIDOTTI, S. A. B. G.; FUJITA, M. S. L. Bibliotecas digitais na era da web semântica: reflexões no âmbito da gestão de conteúdos informacionais. In: FUJITA, M. S. L. et al. *A dimensão social da biblioteca digital na organização e*

acesso ao conhecimento: aspectos teóricos e aplicados: projetos, produtos e relatos de experiência. São Paulo: Sistema de Bibliotecas da USP, 2005. v. 2.

RAUPP, M.; BEUREN, I. M. Metodologia da pesquisa aplicável às ciências sociais. In: LONGARAY, A. A. et.al. *Como elaborar trabalhos monográficos em contabilidade: teoria e prática.* São Paulo: Atlas, 2003. p. 81-93.

ROBREDO, J. A indexação automática de textos: o presente já entrou no futuro. In: MACHADO, U. D. (Org.). *Estudos avançados em Ciência da Informação.* Brasília: Associação dos Bibliotecários do Distrito Federal, 1982. v. 1, p. 235-274.

ROLLEMBERG, R. *Projeto de Lei 1.120, de 21 de maio de 2007.* Dispõe sobre o processo de disseminação da produção técnico-científica pelas instituições de ensino superior no Brasil e dá outras providências. Disponível em: <<http://www2.camara.gov.br/internet/proposicoes>>. Acesso em: 6 jan. 2009.

ROSENFELD, L.; MORVILLE, P. *Information architecture: for the world wide web.* United States of America: O'Reilly, 1988.

SANTOS, J.; TEIXEIRA, C.; PINTO, J. S. *eABC: um repositório institucional virtual.* 2005. In: CONFERÊNCIA IADIS IBERO-AMERICANA WWW/INTERNET. Disponível em: <<http://veqpar.fe.up.pt/xata2005/papersfinal/31.pdf>>. Acesso em: 21 dez. 2008.

SAYÃO, L. F. Preservação digital no contexto das bibliotecas digitais: uma breve introdução. In: MARCONDES, C. H.; KURAMOTO, H.; TOUTAIN, L. B. (Org.). *Bibliotecas digitais: saberes e práticas.* 2. ed. Brasília: IBICT, 2006. p. 113-143.

SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, Campinas, v. 20, n. 2, maio/ago. 2008. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=215>>. Acesso em: 10 jul. 2009.

SCHAFF, A. *Linguagem e conhecimento.* Coimbra: Almedina, 1964.

SCHALLENBERGER, E. Uma universidade tecnológica *multicampi*. In: LIMA FILHO, D. L.; TAVARES, A. G. (Org.). *Universidade tecnológica: concepções, limite e possibilidades.* Curitiba: SINDOCEFET, 2006. p.89-100.

SILVA, N. S. Universidade tecnológica: uma alternativa? In: LIMA FILHO, D. L.; TAVARES, A. G. (Org.). *Universidade tecnológica: concepções, limite e possibilidades.* Curitiba: SINDOCEFET, 2006. p.61-80.

SILVA, T. E. Política de informação na pós-modernidade: reflexões sobre o caso do Brasil. *Informação & Sociedade*, João Pessoa, v. 1, n. 1, p. 47-53, jan./dez. 1991.

SILVA, T. E.; TOMAÉL, M. I. Política de informação: tendências internacionais. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 10., 2009, João Pessoa. *Anais...* João Pessoa: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, 2009. Disponível em: <<http://dci2.ccsa.ufpb.br:8080/jspui/handle/123456789/436>>. Acesso em: 12 abr. 2010.

_____. Repositórios institucionais e o modelo open. In: TOMAÉL, M. I. (Org.). *Fontes de informação na internet*. Londrina: EDUEL, 2008. p. 123-149.

SIMON, I. Universidade diante de novas tecnologias de informação e comunicação. In: SIMPÓSIO INTERNACIONAL UNIVERSIDADE E NOVAS TECNOLOGIAS: IMPACTOS E IMPLICAÇÕES, 1999, São Paulo. *Anais...* São Paulo: USP, 1999. p. 47-49.

SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. *Ciência da Informação*, Brasília, v. 29, n. 1, jan./abr. 2000. Disponível em: <<http://www.scielo.br/pdf/ci/v29n1/v29n1a10.pdf>>. Acesso em: 9 nov. 2009.

STRAIOTO, F. A. *Arquitetura da informação para world wide web: um estudo exploratório*. 2002. 120 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

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

TARGINO, M. G. Comunicação científica: uma revisão de seus elementos básicos. *Informação e Sociedade*, João Pessoa, v. 10, n. 2, 2000. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/326/248>>. Acesso em: 2 maio 2009.

TERRA, J. C. C. *Gestão do conhecimento: o grande desafio empresarial*. Rio de Janeiro: Elsevier, 2005.

TOMAÉL, M. I., SILVA, T. E. *Repositórios institucionais: diretrizes para políticas de informação*. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 8., 28-31 out. 2007, Salvador. *Anais...* Salvador, 2008. Disponível em: <www.enancib.ppgci.ufba.br/artigos/GT5--142.pdf>. Acesso em: 18 mar. 2009.

TOUTAIN, L. M. B. B. Biblioteca digital: definição de termos. In: MARCONDES, C. H.; KURAMOTO, H.; TOUTAIN, L. B. (Org.). *Bibliotecas digitais: saberes e práticas*. 2. ed. Brasília: IBICT, 2006. p. 15-24.

TRISKA, R.; CAFÉ, L. Arquivos abertos: subprojeto da biblioteca digital brasileira. *Ciência da Informação*, Brasília, v. 30, n. 3, set./dez. 2001. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/207/184>>. Acesso em: 18 jun. 2009.

TRISTÃO, A. M. D.; FACHIN, G. R. B.; ALARCON, O. E. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. *Ciência da informação*, Brasília, v. 33, n. 2, maio/jun. 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652004000200017>. Acesso em: 11 abr. 2010.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. *Plano de desenvolvimento institucional da UTFPR - PDI 2009-2013*. 2009. Disponível em: <<http://www3.utfpr.edu.br/comissoes/plano-de-desenvolvimento-institucional>>. Acesso em: 21 abr. 2010.

VENTURA, M. M. O estudo de caso como modalidade de pesquisa. *Revista SOCERJ*, Rio de Janeiro, v. 20, n. 5 p. 383-386, set./out. 2007. Disponível em: <http://sociedades.cardiol.br/socerj/revista/2007_05/a2007_v20_n05_art10.pdf>. Acesso em: 29 maio 2010.

VIANA, C. L. M.; MARDÉRO ARELLANO, M. A.; SHINTAKU, M. *Repositórios institucionais e tecnologia: uma experiência de customização do DSPACE*. 2005. Disponível em: <<http://www.comiteitajai.org.br/Dspace/handle/123456789/143>>. Acesso em: 12 fev. 2010.

VICENTINI, L. A. Gestão em bibliotecas digitais. In: MARCONDES, C. H.; KURAMOTO, H.; TOUTAIN, L. B. (Org.). *Bibliotecas digitais: saberes e práticas*. 2. ed. Brasília: IBICT, 2006. p. 243-262.

VICKERY, B. C. *Classificação e indexação nas ciências*. Rio de Janeiro: BNG/Brasilart União Editorial, 1980.

VIDOTTI, S. A. B. G.; SANT'ANNA, R. G. Infra-estrutura tecnológica de uma biblioteca digital: elementos básicos. In: MARCONDES, C. H. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. Brasília: IBICT, 2006. p. 77-93.

WEITZEL, S. R.; FERREIRA, S. M. S. P. Arena científica: um repositório da área das ciências da comunicação promovendo o acesso livre e o desenvolvimento científico. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS – SIBD, 3., 2005, São Paulo. *Anais...* São Paulo: USP, 2005.

YIN, R. K. *Estudo de caso: planejamento e métodos*. 3. ed. Porto Alegre: Bookman, 2005.

ZHAODONG, L.; XAOCHU, W. Estudo preliminar sobre a política nacional de informação científica e tecnológica da China: uma opinião pessoal. *Ciência da Informação*, Brasília, v. 16, n. 1, jan./jun. 1987. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/download/1599/1212>>. Acesso em: 12 jan. 2009.

APÊNDICES

APÊNDICE A

**FORMULÁRIO PARA ANÁLISE DOS ELEMENTOS DE ORGANIZAÇÃO DA
INFORMAÇÃO EM REPOSITÓRIOS INSTITUCIONAIS****I INFORMAÇÕES GERAIS**

Nome do Repositório: _____

Url: _____

Instituição responsável: _____

Data de implantação do repositório: _____

Total de itens disponíveis: _____

Data da análise: _____

1- Qual a tipologia documental existente no repositório? quantidade

- () Livros _____
- () Capítulos de livros _____
- () Artigos _____
- () Teses _____
- () Dissertações _____
- () Trabalho de Conclusão de Curso (Graduação)
- () Trabalho de Conclusão de Curso (Especialização)
- () Trabalho de Conclusão de Curso (Aperfeiçoamento)
- () Documentos institucionais _____
- () Filmes / vídeos _____
- () Fotografias _____
- () Recursos didáticos _____
- () Trabalho em eventos _____
- () Outros – Quais?

2- Há política adotada para a inserção de conteúdos no repositório?

II ESTRUTURA DE ORGANIZAÇÃO DA INFORMAÇÃO

1- Qual a estrutura de informação do repositório?

- () Comunidades
- () Subcomunidades
- () Coleções

2- Como está estabelecida a estrutura de informação (alta generalização) do repositório?

- () áreas do conhecimento (CNPq)
- () áreas administrativas
- () Tipologia documental

3- Quais são as comunidades (categoria)

4- Quais as sub-comunidades (subcategoria)

5- Quais as coleções existentes (classes)

6- Como estão ordenadas as coleções?

- título
- assunto
- autor
- data de publicação

III METADADOS:

1- Qual o padrão de metadados utilizados?

2- Quais são os campos utilizados para descrição dos documentos?

3- A instituição utiliza um formulário padrão para a entrada dos documentos?

IV INTEROPERABILIDADE SEMÂNTICA

1 Como ocorre a recuperação da informação no Repositório:

- pesquisa rápida por palavras
- pesquisa avançada (operadores booleanos)
- a busca é permitida para em todo o repositório?
- a busca deve ser feita por tipo de material
- a busca deve ser feita por comunidade/subcomunidade
- por autor
- data de publicação
- orientador
- título
- índice de assunto (glossário)

APÊNDICE B

ROTEIRO DE ENTREVISTA

Prezado(a) Senhor(a):

Este instrumento de pesquisa visa contemplar os objetivos propostos no estudo: **ORGANIZAÇÃO DA PRODUÇÃO CIENTÍFICA EM REPOSITÓRIOS INSTITUCIONAIS**: um parâmetro para a UTFPR, vinculado ao Mestrado Profissional em Gestão da Informação da Universidade Estadual de Londrina. O objetivo é identificar as etapas que procederam a implantação do repositório nesta instituição, analisando a estrutura de organização utilizada de forma a propor um modelo adequado às necessidades da UTFPR, instituição a qual a pesquisadora encontra-se vinculada.

I INFORMAÇÕES GERAIS

INSTITUIÇÃO:

NOME DO REPOSITÓRIO:

NÍVEL DE PARTICIPAÇÃO NO REPOSITÓRIO:

- Coordenador
- Assessor
- Equipe Técnica
- Outra responsabilidade. Qual: _____

DATA DE IMPLEMENTAÇÃO DO REPOSITÓRIO:

QUAL O OBJETIVO DO REPOSITÓRIO:

NOME DO RESPONDENTE:

1- Comente as etapas que permitiram a criação do repositório nesta instituição (de quem partiu a iniciativa, quais foram os processos necessários, como se deu a escolha hardware e software, etc)

2- A instituição definiu um Comitê Gestor para o repositório?

- () formalmente, por quais meios: _____
- () informalmente, por quais meios: _____
- () não possui comitê gestor _____

3- Se a resposta anterior foi positiva, a composição deste comitê deu-se desde a implantação repositório? Ou após sua implantação?

4- Quais profissionais compõem o Comitê Gestor?

- () docentes
- () técnicos de informática
- () bibliotecários
- () outros profissionais. Quais: _____

5- Mencione as principais atribuições deste Comitê na instituição.

6- Aos bibliotecários da instituição competem quais atividades no desenvolvimento do repositório? _____

7- Há uma política de informação institucional explícita?

8- Há uma política institucional para a inserção de conteúdos no repositório que contemple: Qual tipologia pode ser inserida? Se irá conter somente a produção institucional ou se o repositório aceita disponibilizar outras informações de interesse institucional. Um servidor que produz fora da instituição, como fica definido o depósito deste tipo de documento?

9- Qual o sistema gerenciador da biblioteca?

10- Este sistema permite a interconexão de dados com o repositório? (campo 856)

11- Todos os itens depositados no repositório estão disponíveis em acesso aberto? Ou possui itens restritos a comunidade externa e se positivo, por quais motivos ocorre a restrição de acesso?

II ESTRUTURA DE ORGANIZAÇÃO DA INFORMAÇÃO

1- Qual a estrutura de organização da informação estabelecida para o repositório?

2- Quais são os critérios para a criação de uma comunidade, sub-comunidade ou coleção?

3- Quais instrumentos são utilizados a fim de representar os documentos depositados no repositório no que tange à representação descritiva e temática?

4- Mencione se há existência de políticas indexação (visando a qualidade e instrumentos utilizados para indexação dos conteúdos) para o repositório institucional.

5- Em sua opinião quais são os pontos positivos e passíveis de melhoria no que tange à estrutura adotada pela instituição para organizar sua produção intelectual em repositórios institucionais.

III METADADOS

1- De que forma a instituição definiu seu esquema de metadados?

2- Utilizou-se dos metadados pré-estabelecidos no DSpace? Ou houve a necessidade de inserir novos campos?

3- Há uma definição quanto ao número de descritores para representação de assunto de cada obra?

4- Qual padrão de metadados utilizados?

MARC

Dublin Core

Outro: _____

5- Quem realiza o registro e controle dos metadados na submissão? A biblioteca ou autor?

IV INTEROPERABILIDADE SEMANTICA E HUMANA

1- Utiliza vocabulário controlado para a indexação dos conteúdos?

Não Justifique

Sim especifique:

tesouros

taxonomias

ontologias

2- A interoperabilidade humana diz respeito à sensibilização dos envolvidos: (pesquisadores, docentes, servidores, e outros) no que diz respeito ao acesso aberto/repositórios institucionais. Desse modo, como você avalia a interoperabilidade humana da instituição, no que tange à manutenção do repositório institucional?