



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

GT 2 - Organização e Representação do Conhecimento
Modalidade de apresentação: comunicação oral

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

Lígia Patrícia Torino
Universidade Tecnológica Federal do Paraná
Brigida Maria Nogueira Cervantes
Universidade Estadual de Londrina

Resumo: Analisa os elementos de organização da informação em repositórios institucionais, com o intuito de elaborar uma proposta de organização da informação para a Universidade Tecnológica Federal do Paraná (UTFPR), utilizando como método o estudo de caso e a pesquisa descritiva, com enfoque na análise qualitativa. Norteados pelo objetivo geral de propor uma estrutura de organização da informação em repositórios institucionais como um parâmetro para a UTFPR, este estudo buscou identificar nos diretórios *OpenDoar* e *Roar* os repositórios de instituições de ensino superior registrados; analisar as estruturas de organização da produção intelectual existentes; comparar suas estruturas de organização e selecionar as mais adequadas. Na etapa de análise e apresentação dos resultados apresentam-se os elementos necessários para auxiliar a elaboração de uma estrutura de organização da informação para a UTFPR, contemplando as necessidades que vão desde a elaboração da política de informação, as etapas de definição de padrões de metadados e a interoperabilidade humana e semântica necessária à instituição que pretende implantar um repositório institucional. Concluiu-se que 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 ela representa quanto a facilidade de acesso às informações nela depositadas. Partindo dos resultados obtidos, foi possível propor uma estrutura de organização da informação para o repositório da UTFPR.

Palavras-chave: Organização da Informação. Produção científica. Repositórios institucionais.



1 INTRODUÇÃO

Este estudo está inserido 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 apresentar um modelo de organização da informação aplicável à implantação do repositório institucional da Universidade Tecnológica Federal do Paraná (UTFPR).

Em âmbito acadêmico, um repositório institucional atuará como um importante mecanismo de gestão da informação institucional, podendo ser caracterizado como “um conjunto de serviços que a universidade oferece aos membros de sua comunidade para a gestão e disseminação de objetos digitais criados pela universidade e membros de sua comunidade” (LYNCH, 2003, p. 2).

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 vão sendo ampliados, as bibliotecas e/ou unidades de informação passam a aderir a este movimento e o apóiam principalmente no que tange aos critérios de organização das informações.

Viana, Márdeno Arellano e Shintaku (2005, p. 8) constatam 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 do usuário ao documento; a preservação dos documentos depositados e a garantia dos direitos autorais.

No intuito de identificar os fatores que motivaram a implantação de um repositório para a UTFPR, destaca-se que, embora recente como universidade, a instituição 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, 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.

A relevância deste estudo está na oportunidade de apresentar teoricamente os conteúdos científicos sobre o tema e propor à instituição mencionada sua aplicação. Isto se deu, sobretudo, pela iniciativa da pesquisadora que é servidora da instituição. O projeto de instalação de um repositório na UTFPR, embora tenha sido planejado por uma equipe pequena composta por três bibliotecários, um docente e uma mestranda do Programa de Mestrado em Tecnologia da UTFPR, ampliou-se na instituição após ter sido aprovado pelo edital FINEP/PCAL/XBDB n.01/2009 que apoiava a implantação de repositórios institucionais no Brasil. O referido edital possibilitou à instituição receber do Instituto Brasileiro de Informação Ciência e Tecnologia (IBICT), um *kit* tecnológico contendo o *hardware* e o *software* Dspace para implantação de seu repositório.

Partindo deste pressuposto, identificou-se a necessidade de um estudo que pudesse analisar as estruturas de organização da informação existentes em repositórios institucionais de universidades brasileiras que viessem a nortear um modelo propício de estruturação da informação condizente às necessidades da UTFPR. Supõe-se que experiências de outras instituições de ensino sejam fundamentais para que novas ações possam ser propostas partindo da análise de modelos já existentes.

Desta forma, este estudo teve como objetivo propor uma estrutura de organização da informação em repositórios institucionais como um parâmetro para a UTFPR, por meio da análise das estruturas de organização da produção intelectual existentes em repositórios de instituições acadêmicas de ensino superior. Partindo dos resultados obtidos, foi possível analisar e comparar os critérios de organização utilizados nos repositórios das instituições em estudo, identificando elementos aplicáveis ao caso UTFPR.



2 REPOSITÓRIOS INSTITUCIONAIS: os objetivos, critérios de organização e a definição de políticas para sua implementação

Os repositórios digitais emergem do movimento de acesso aberto à informação apresentando um novo modelo de gestão da produção intelectual de uma instituição. Esta ferramenta possibilita reunir, organizar, preservar e ampliar a visibilidade das produções, disponibilizando-as em texto completo para livre acesso na internet. 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 opção por pesquisar a utilização de repositórios em instituições acadêmicas surgiu porque, nas universidades, a produtividade científica é um dos critérios de análise para a avaliação e atribuição de conceito aos programas de pós-graduação, 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. Desta forma, 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).

Além de ampliar a visibilidade e o fator de impacto, tanto para o programa quanto para o pesquisador, As universidades estão atendendo a esta demanda e identificam a necessidade de ampliar a variedade de tipologias documentárias a serem armazenadas, fator que tem sido possibilitado por meio dos repositórios digitais.

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”.

A implementação de um repositório institucional depende sobretudo, da definição das políticas que garantirão sua sustentabilidade. Considerando o posicionamento das autoras Tomaél e Silva (2007, p.4-5), as políticas 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.

3 ELEMENTOS DE ORGANIZAÇÃO DA INFORMAÇÃO EM REPOSITÓRIOS INSTITUCIONAIS

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; salientando 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).

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 de descrever os documentos, além de estabelecer estruturas físicas ou não, que possibilitem armazenar os documentos e os seus conteúdos.

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 a qual definirá o formato em que os dados serão apresentados na página onde estão disponíveis, 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), tais autores citam que os esquemas de organização da informação são ordenados de acordo com elementos exatos e ambíguos. Os elementos exatos categorizam a informação de forma clara, caracterizado como o esquema de organização que atende às necessidades daquele usuário que sabe o que quer, enquanto o ambíguo trabalha com categorias subjetivas, indicado àqueles que não sabem exatamente o que precisam.



Os esquemas híbridos são considerados um sistema capaz 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”.

Ao abordar as técnicas utilizadas e seus instrumentos de forma que fosse possível atingir aos objetivos propostos neste estudo, fez-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) esta linha de pesquisa consiste em representar a informação de forma que haja possibilidade de explicitar os dados que remetam tanto à autoria de um documento quanto ao local onde fora produzido, bem como aos conteúdos nele contemplados.

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); trata-se de um 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 (quando aborda o aspecto físico do documento), ou catalogação por assunto



(quando diz respeito ao conteúdo). Essa técnica requer tratamento realizado por especialistas, os quais possibilitarão que o usuário localize a 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”.

De acordo com Monteiro (2008, p. 51) “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 formato 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*. 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. 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).

4 ESTRATÉGIAS METODOLÓGICAS

A metodologia da pesquisa caracteriza-se como um estudo de caso, descritivo, com enfoque na análise qualitativa. A extensão da amostra é representada pelo total de instituições acadêmicas que possuem repositórios institucionais registrados nos diretórios *OpenDoar* e *ROAR*. 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. Esta seleção buscou obter dados de instituições com características similares à UTFPR, e que utilizassem o mesmo software oferecido pelo IBICT.

Neste estudo, foram definidos como instrumentos para coleta de dados: o formulário e a entrevista. 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, buscando identificar junto ao comitê gestor dos repositórios, os elementos não explícitos nos repositórios analisados.

A opção por combinar o instrumento formulário à técnica de entrevista deu-se com a finalidade de permitir que possíveis dúvidas pudessem ser esclarecidas junto aos gestores dos repositórios selecionados. 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).

Após selecionar os 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, possibilitando a elaboração de uma pesquisa de melhor qualidade.

5 ANÁLISE DOS RESULTADOS

A análise dos resultados enfatiza os elementos necessários para auxiliar a elaboração de uma estrutura de organização da informação para repositórios institucionais. A análise foi sistematizada em quatro categorias: Informações gerais sobre os repositórios; Estruturas de organização da informação; Metadados; e Interoperabilidade (semântica e humana).

5.1 INFORMAÇÕES GERAIS SOBRE OS REPOSITÓRIOS ESTUDADOS

A obtenção de informações gerais a respeito das etapas que procederam a implantação do repositório em cada uma das instituições oportunizou maior aproximação com o tema, possibilitando avaliar o percurso desde a etapa de seleção do *software*, definição da equipe operacional do repositório, a sensibilização junto à comunidade acadêmica, bem como as decisões a respeito da definição da política de informação que norteie as ações de uma ferramenta de gestão, como o repositório. Constata-se que a definição de uma política de informação deve ser uma das primeiras ações da instituição que pretende 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 tempo 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 em várias instâncias, percebe-se que, se por um lado os processos administrativos diminuem o ritmo e desenvolvimento do repositório, por outro, envolvem mais instâncias da universidade, tornando-se instrumento democrático de discussão. 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 ímpar.

Nas três instituições estudadas, bibliotecários estão envolvidos com o repositório, cujas atribuições perpassam as etapas de divulgação da ferramenta junto às instâncias administrativas, abrangendo a 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.

5.2 ESTRUTURAS DE ORGANIZAÇÃO DA INFORMAÇÃO

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. Pode-se mencionar que 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 ele quer buscar a informação: especialmente, para a comunidade externa à instituição. As coleções se apresentam 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 sua página inicial, há indicação de que o autor é o titular dos direitos autorais dos itens submetidos, 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, fator que minimiza esforços daquele 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 esclarece-se que, o DSpace UFPR estrutura seu repositório de acordo com as solicitações que lhes são feitas, ou seja, são criados de acordo com a demanda, fator que ocorre também no repositório Lume, este diferencia-se no sentido que cada uma das solicitações devem ser aprovadas pelo Comitê Gestor, que irá avaliar a relevância e o comprometimento do solicitante com a continuidade de submissão de documentos, visando manter em constante atualização o repositório institucional. Quanto ao RiUnB, esta definição é da própria universidade, visto que o repositório reflete sua estrutura administrativa e acadêmica. Por conseguinte, se surgirem novas coordenações, departamentos, cursos, conseqüentemente surgirão novas comunidades, sub-comunidades e coleções no RiUnB.



De acordo com as informações obtidas, verificou-se que comum aos três repositórios é o fato de o maior número de itens depositados estar atrelado à tipologia Teses e Dissertações, o que nos faz refletir que grande parte dos repositórios concentra os resultados de pesquisas científicas geradas em universidades. Este resultado indica que, no Brasil, uma grande parcela das universidades está trabalhando na implantação de seus repositórios digitais. 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. Sob este aspecto, infere-se que muitas instituições ampliam assim as suas bibliotecas digitais de teses e dissertações, acrescentando a elas novas tipologias documentárias, novos formatos de documentos, sobretudo aqueles validados por pares.

Constata-se que a maioria dos metadados são cooperados de sistemas gerenciadores de bibliotecas que nos casos aqui analisados, utilizam o padrão de metadados MARC, o Código de Catalogação Anglo Americano AACR2, como tabelas de classificação a CDD ou CDU e vocabulários controlados da LC (*Library Congress*) ou do sistema Pergamum, sempre combinados com a linguagem natural. Pondera-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 e melhoria na qualidade do que é produzido. Entretanto, para que o repositório mantenha-se em constante crescimento, é necessário a sua institucionalização, ou seja, ser conhecido e reconhecido por todas as instâncias da universidade. Para isso, necessita de grande divulgação e controle dos fluxos dos documentos.



5.3 METADADOS UTILIZADOS

Observou-se que todos os repositórios estudados seguem o padrão MARC (para documentos provenientes de seus sistemas de gestão de acervo) compatibilizando-o para os campos do Dublin Core (DSpace), sendo que, de acordo com 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 auto-depósito, o controle de qualidade dos itens indexados é realizado pelas próprias bibliotecas, que são responsáveis tanto pela atribuição quanto pelo controle dos metadados, sendo que, para isto, utilizam-se de vocabulários controlados e da linguagem natural para indexar os descritores que serão definidos pelo profissional bibliotecário.

5.4 INTEROPERABILIDADE HUMANA E SEMÂNTICA

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 apoio institucional é também fator de extrema relevância para estes repositórios, se não houvesse apoio e abertura para que esta ferramenta fosse implantada, certamente a visibilidade da instituição e a 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 dos projetos de implantação de repositórios buscaram em experiências pioneiras um norte para a definição de suas estruturas. Este estudo buscou, na 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 ocorrem simultaneamente, visto que uma ação depende da outra.

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

Quadro 1: Ações para implementação de repositórios institucionais

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

Fonte: A pesquisa.

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

De acordo com nossos estudos e com o objetivo proposto para o repositório institucional da UTFPR que consiste em “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 (produções intelectuais dos membros da comunidade realizadas fora do âmbito da UTFPR) Conforme demonstra a figura 1:

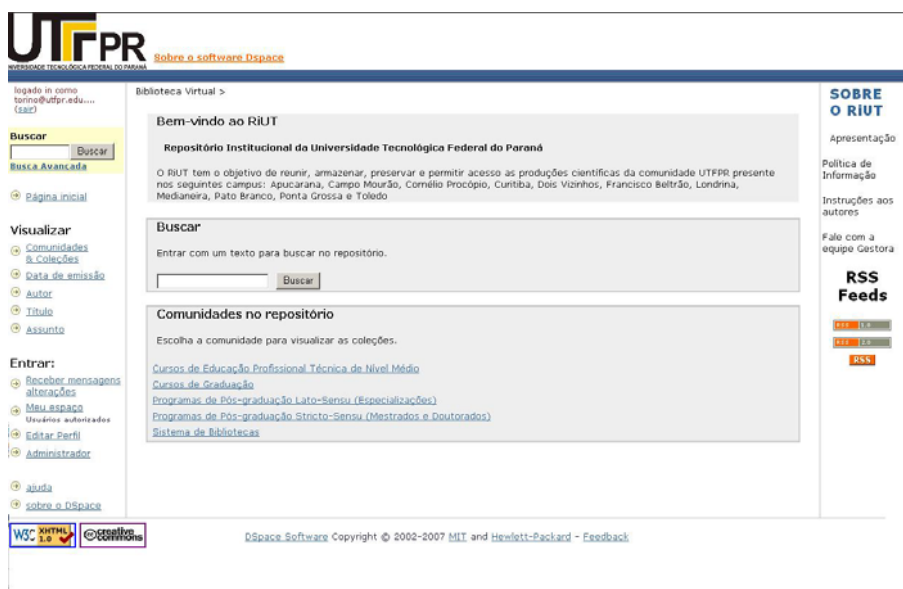


Figura 1: Estrutura de organização do Repositório da UTFPR

Fonte: A pesquisa

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, embora sua estrutura seja *multicampi*, a opção pela não categorização por *campusi*, deu-se como forma de propor uma estrutura mais unificada, buscou englobar a UTFPR de forma ampla, ou seja, em uma visão “sistêmica”. Pela definição da política institucional da UTFPR, ficou determinado compor 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.

A organização da informação se dará por meio 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, utilizará o 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 inicialmente o trabalho dos bibliotecários do SIBI UTFPR que farão as submissões inicialmente, 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 no metadado indicado 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 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.

A indexação é uma etapa de valiosa importância. Destarte, 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.

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



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

institucionais devem ocorrer respaldadas em decisões da equipe gestora, possibilitando a integração entre diversas áreas/setores e permitindo que o trabalho e 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 enriquece-se na medida em que apresenta teoricamente os conteúdos que combinados com a prática permitem 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 são tarefas são bastante complexas.

O projeto de instalação de um repositório na UTFPR, embora tenha sido planejado inicialmente por uma equipe pequena composta por bibliotecários, um docente e uma mestrande do Programa de Mestrado em Tecnologia da UTFPR, ampliou-se na instituição após ter sido aprovado pelo edital do IBICT. Os processos de aprovação da política de informação institucional, que atualmente encontra-se em trâmite, para aprovação, pela PROGRAD e 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. 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 muito importante e precisa ser discutida pela equipe gestora.

Aliar as questões teóricas à realidade profissional é um fator bastante complexo e desafiador, porém a recompensa é observada nos 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. 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).



Embora se trate 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 com gestores das instituições que já haviam passado pela experiência de implantação de seu repositório institucional. Contudo, acredita-se que novos estudos poderão ampliar esta discussão, ao observar os critérios adotados pelas instituições para manter em constante atualização seus repositórios.

ABSTRACT: This paper analyzes the elements of organizing information in institutional repositories, in order to develop a proposal for organizing information for the Federal Technological University of Paraná (UTFPR) using as a method the case study and the descriptive research, focusing on qualitative analysis. Guided by the general goal of proposing a structure for organizing information in institutional repositories as a parameter for the UTFPR, this study sought to identify in the directories OpenDoar and Roar the registered repositories of High Education Institutions; analyze the existing organizational structures of intellectual production; compare their organizational structures and select the most appropriate ones. In the stage of analysis and presentation of results, the elements necessary to assist the elaboration of a structure for organizing information for the UTFPR are presented, addressing needs ranging from the preparation of information policy, the steps of defining metadata standards and human and semantic interoperability necessary for the institution which aims to deploy an institutional repository. It was concluded that the analyzed repositories comply with the intended objectives and present a framework of systematic organization taking into consideration both the institution it represents and the ease of access to the information deposited in it. According to the results obtained, it was possible to propose a structure for organizing information for the repository of the UTFPR.

Keywords: Information Organization. Scientific production. Institutional repositories

REFERÊNCIAS:

ANZOLIN, H. H. Atualizações em AACR2. In: ENCONTRO NACIONAL DOS USUÁRIOS DA REDE PERGAMUM, 9., 2007, Curitiba. [Anais]. Curitiba: PUCPR, 2007. 1 CD-ROM.

BARROS, A. J. S.; LEHFELD, N. A. S. *Fundamentos de metodologia científica: um guia para a iniciação científica*. 2. ed. São Paulo: Pearson, 2006.

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



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

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).

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).

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

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.

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.

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

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 <iptorino@gmail.com>, 13 jun. 2010.

MIRANDA, M. L. C. *Organização e representação do conhecimento: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes digitais*. 2005. 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. 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.

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



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

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://vecpar.fe.up.pt/xata2005/papersfinal/31.pdf>>. Acesso em: 21 dez. 2008.

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.

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.

TOMAÉL, M. I., SILVA, T. E. *Repositórios institucionais: diretrizes para políticas de informação*. In: ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação, 8, 28 a 31 out. 2007. Salvador. *Anais...* Salvador, 2008. Disponível em: <www.enancib.ppgci.ufba.br/artigos/GT5--142.pdf>. Acesso em: 18 mar. 2009.

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.