



FÓRUM DE PESQUISA DA PÓS-GRADUAÇÃO
Suzana Pinheiro Machado Müller
Universidade de Brasília

Repositório digital: acesso livre à informação na Universidade Federal do Rio Grande do Sul

Caterina Groposo Pavão Correio

Universidade Federal do Rio Grande do Sul

Afonso Comba de Araújo Neto Correio

Universidade Federal do Rio Grande do Sul

Lais Freitas Caregnato Correio

Universidade Federal do Rio Grande do Sul

Janise Silva Borges da Costa Correio

Universidade Federal do Rio Grande do Sul

Zaida Horowitz Correio

Universidade Federal do Rio Grande do Sul

Zita Prates de Oliveira Correio

Universidade Federal do Rio Grande do Sul

Carla Metzler Saatkamp Correio

Universidade Federal do Rio Grande do Sul

Resumo

Este trabalho descreve a concepção e organização do projeto-piloto do LUME - Repositório Digital da Universidade Federal do Rio Grande do Sul, iniciativa conjunta da Biblioteca Central e do Centro de Processamento de Dados. Aborda os recursos de software e de hardware utilizados e os padrões de representação da informação empregados na sua implementação, tendo em vista a preservação e divulgação do seu conteúdo e sua interoperabilidade com outros repositórios digitais.

Palavras-chave: Repositórios digitais. Repositórios institucionais. Iniciativa de arquivos abertos. Metadados. *DSpace*.



1 INTRODUÇÃO

Desde 1989, quando foi implantada a automação no seu Sistema de Bibliotecas, a Universidade Federal do Rio Grande do Sul (UFRGS) vem trabalhando, prioritariamente, na coleta e registro da produção intelectual do seu corpo docente e técnico-administrativo, a fim de, dentre outros objetivos, garantirem o controle bibliográfico institucional e divulgar os documentos produzidos em seu âmbito. Em 2001, implantou sua Biblioteca Digital de Teses e Dissertações e, desde então, considerando os avanços nas áreas de tecnologia da informação e comunicação, caminha na direção dos repositórios institucionais.

A tecnologia da informação tem um papel estratégico na UFRGS, não apenas no que refere às possibilidades de expansão dos serviços bibliotecários e como instrumento fundamental para o ensino presencial e a distância, mas também como fonte de indicadores gerenciais e integração com outros repositórios de ensino e pesquisa no país. A grande quantidade de produção de conhecimento dentro das universidades faz com que seja necessário, além da sua difusão e uso, a sua preservação. É comum, depois de algum tempo, haver quebra de *link*, seja por mudança de endereço do servidor onde estava hospedado o documento ou simplesmente porque o mesmo foi excluído da página.

Os repositórios institucionais, ao utilizarem metodologia de arquivos abertos (*Open Archives Initiative*), visam a promover o acesso livre à informação científica e acadêmica e, sobretudo, possibilitam armazenar, preservar e divulgar a produção intelectual, aumentando o impacto e a visibilidade das pesquisas desenvolvidas na instituição (CAMARGO, 2006).

A gestão do conhecimento está diretamente relacionada com os projetos de desenvolvimento de repositórios institucionais nas universidades. Estes colaboram para um melhor gerenciamento e maior acesso ao saber estimulando, assim, a geração de novos conhecimentos.

Com base nesta perspectiva está sendo desenvolvido o projeto LUME, que significa luz, denominação atribuída ao Repositório Digital da Universidade Federal do Rio Grande do Sul, alvo deste trabalho.



2 LUME - REPOSITÓRIO DIGITAL DA UFRGS

2.1 Antecedentes

Com a automação das rotinas e serviços básicos das bibliotecas em estágio avançado, a Universidade procurou maximizar o uso da informação especializada disponível no acervo do Sistema de Bibliotecas da UFRGS (SBU) e daquela produzida pelo seu corpo docente e técnico-administrativo.

A biblioteca digital foi o veículo escolhido para divulgar e preservar um acervo em contínuo crescimento, por sua possibilidade de expandir os limites de acesso e uso da informação para além das barreiras de horário e espaço físico das bibliotecas tradicionais. Diferentemente de um catálogo on-line, que apresenta apenas a descrição bibliográfica dos documentos, a biblioteca digital agrega à descrição bibliográfica e temática, o conteúdo do documento, exigindo cuidados diferenciados para garantir sua integridade, preservação e, também, a observância da legislação de direito autoral.

A Biblioteca Digital de Teses e Dissertações da UFRGS (BDTD), implementada em 2001, tem demonstrado um crescimento significativo a cada ano, porém, a necessidade de ampliação da BDTD assume papel relevante na tarefa de reunir o acervo digital das demais coleções existentes no SBU e coleções dispersas em outros órgãos da Universidade, tais como, fotografias, partituras, discos sonoros, periódicos, entre outros, otimizando sua organização, gerenciamento, manutenção e compartilhamento de recursos. Aos usuários finais, possibilita a realização de buscas num único portal de informações de acesso irrestrito.

Neste contexto, em 2007, teve início a ampliação da BDTD para repositório, o qual possibilita armazenar diversos tipos de documentos em formato digital, resultado das atividades de pesquisa e docência dos professores e pesquisadores.

2.2 Objetivos

O LUME tem por objetivos reunir, preservar, divulgar e garantir o acesso confiável e permanente aos documentos acadêmicos, científicos, artísticos e administrativos gerados na Universidade, bem como às suas coleções históricas e a



outros documentos de relevância para a Instituição, que fazem parte de suas coleções, embora não produzidos por ela, maximizando o uso desses recursos.

A preservação dos documentos digitais pode ser apontada como uma das principais atribuições dos repositórios, mas também propiciam, à instituição, transparência e acessibilidade, pois disponibilizam a comunidade em geral a sua produção científica, permite maximizar o impacto das pesquisas, aumentam as possibilidades de arrecadação de fundos, obtenção de prêmios e prestígio compartilhado pelos pesquisadores e pela instituição (CAFÉ, 2003). “Por meio dos repositórios institucionais, um pesquisador tem acesso a informações que serão combinadas, reagrupadas de forma a agregar valor e acrescentar ao novo conjunto de informações um novo conhecimento explícito, gerando um novo conhecimento” (LEITE, 2006).

2.3 Tecnologias utilizadas

2.3.1 Softwares

O LUME foi concebido utilizando o *DSpace Institutional Digital Repository System* (DSPACE, 2008), resultado de um projeto desenvolvido em parceria pelo MIT (*Massachusetts Institute of Technology*, EUA) e a HP (*Hewlett-Packard Company*, EUA), orientado à criação de repositórios institucionais e à preservação digital de sua produção intelectual. Na UFRGS, o *DSpace* está instalado em um servidor Linux. O software, escrito em Java, é apoiado por um conjunto de ferramentas de código aberto:

- Banco de dados relacional PostgreSQL, onde estão armazenados os metadados (dados sobre dados);
- Servidor de aplicações Java TOMCAT Servlet Engine;
- Motor de pesquisa LUCÈNE, para indexação e recuperação dos metadados e documentos textuais;
- Apache HTTP Server;
- Bibliotecas Java (JavaBeans Activation Framework, Java Servlet e JavaMail API); e
- Compilador Ant.



O uso do *DSpace* permite estruturar a informação do repositório, conforme a estrutura organizacional da instituição, de modo a identificar setores ou grupos de pesquisa, sob a forma de comunidades, subcomunidades e coleções. A estrutura pode ser distinta para cada comunidade. Sob as coleções estão os itens que correspondem ao conteúdo digital, os quais são acessados pelos usuários. Pode haver um ou mais objetos digitais - texto, imagem, vídeo e áudio - associados a um documento digital.

Para personalizar a interface gráfica do *DSpace*, tornando-a mais amigável, é utilizado o software *Manakin* (DIGITAL INITIATIVES..., 2005), desenvolvido pela Universidade Texas A&M. É uma interface alternativa à original do *DSpace*, desenvolvida utilizando a filosofia de orientação a aspectos, o que permite a separação completa entre o desenvolvimento de funcionalidades e a customização do ambiente do sistema (temas, cores, fontes, logotipos etc.).

Para garantir o acesso permanente aos documentos digitais, independente de qualquer mudança de endereço do servidor, utiliza-se o *Handle System*, serviço fornecido pelo CNRI - *Corporation for National Research Initiatives* (CORPORATION FOR..., 2008), que consiste em atribuir identificadores persistentes para cada documento digital garantindo que, mesmo que o endereço do servidor do repositório digital mude, os recursos possam continuar sendo referenciados univocamente.

2.3.2 Hardware

O sistema e o banco de dados encontram-se em um equipamento HP Proliant ML370, com sistema operacional Linux Red Hat Enterprise.

Atualmente, o LUME contém metadados e documentos digitais oriundos da coleta automática de outros sistemas (*harvesting*), permitindo a integração de diferentes processos de depósito necessários a um sistema multidisciplinar:

- o sistema MUSEU, contém o acervo de fotos do Museu da UFRGS e o acervo histórico da Escola de Educação Física da UFRGS. Os dados encontram-se num BD SQL Server - servidor Windows, e os documentos digitais em outro servidor Windows. O *link* destes documentos digitais é feito com o uso do software *Samba*;



- o sistema SABI, Sistema de Automação de Bibliotecas da UFRGS, utiliza o software Aleph500, que gerencia as atividades e serviços oferecidos pelas bibliotecas à comunidade usuária. As informações do sistema estão armazenadas num BD Oracle, numa máquina Linux, onde se encontram, também, os documentos digitais.

2.4 Organização e interface gráfica

O *DSpace* é composto por comunidades, subcomunidades e coleções. No projeto-piloto do LUME foram definidas, até o momento, as seguintes comunidades:

- Acervo do Museu Universitário;
- Acervo histórico da Escola de Educação Física;
- Eventos;
- Teses e dissertações; e
- Trabalhos de conclusão de curso de especialização.

Cada comunidade determina sua própria política de informação, fluxo de depósito e níveis de acesso aos documentos, e pode ser dividida em subcomunidades e em coleções. Cada coleção contém itens compostos por metadados como título, autor, data, palavra-chave, resumo, tamanho em bytes, entre outros, e por um ou mais arquivos em formatos variados (pdf, jpg, avi, mp3, html, etc.) chamados de *bitstreams*, com o conteúdo digital propriamente dito. Esses dados podem ser exibidos de forma completa ou resumida. No LUME, os metadados são definidos de forma diferente, conforme as necessidades e particularidades dos documentos contidos em cada comunidade.

Para este projeto-piloto foram feitas várias alterações na interface original do *DSpace*, com o uso da ferramenta *Manakin*, onde foi possível programar alterações estruturais diretamente no código fonte da aplicação, personalizando totalmente o sistema. Além das alterações visuais, foi realizada a tradução, do inglês para o português, de diversas etiquetas de campos de itens do menu, de textos de ajuda e outros.

É possível recuperar informações a partir de filtros genéricos pré-definidos, em forma de listas, nos metadados de título, autor, palavra-chave e ano, como



também a partir de filtros específicos de uma determinada comunidade. É possível ainda fazer uma pesquisa genérica, em todo o repositório.

Os documentos no LUME estão, em sua maioria, disponíveis em texto completo, salvo algumas exceções que, por solicitação do autor, são divulgados parcialmente, por se tratarem de artigo de periódico submetido para publicação, por conterem dados sigilosos de empresas ou pacientes, entre outros.

Há, ainda, documentos digitais de acesso restrito, em função dos direitos autorais, para os quais se criou um *login* específico, viabilizando que somente os usuários UFRGS tenham direito de acesso aos mesmos. A autenticação é feita de forma segura, através de *criptografia SSL*.

O usuário tem a possibilidade de indicar o seu perfil de interesse, mediante assinatura das coleções. Desta forma, quando houver a inclusão de novos documentos nas coleções assinadas, receberá a relação dos mesmos por correio eletrônico.

Foi criado um *link* "Estatísticas" contendo informações a respeito dos itens mais consultados em cada Comunidade, bem como número de *downloads* geral ou por Comunidade e o número de usuários cadastrados UFRGS e não UFRGS.

2.5 Metadados e interoperabilidade

O LUME é composto por metadados e informações digitais. Os metadados são informações estruturadas que possuem semântica padronizada e são utilizados para representar as informações digitais de maneira bibliográfica. Seguem o padrão *Dublin Core - Dublin Core Metadata Initiative* (DUBLIN CORE..., 2008), que consiste de quinze elementos básicos usados para identificar e descrever documentos digitais, aos quais podem ser atribuídos qualificadores, de acordo com as especificidades e necessidades da instituição.

A adoção deste padrão possibilita um detalhamento mais refinado do conteúdo de documentos digitais, com características tão peculiares, que normalmente não são abrangidas num catálogo on-line, o que só vem em benefício dos usuários.

A tabela 1 relaciona os metadados adotados em cada uma das comunidades.

Tabela 1 - Metadados por comunidade do LUME

Comunidades			Conteúdo no LUME
Acervo MU ¹ Acervo ESEF ²	Eventos	TD ³ TCCE ⁴	
<i>Metadados</i>	<i>Metadados</i>	<i>Metadados</i>	
	contributor.advisor	contributor.advisor	Orientador
		contributor.advisor-co	Co-orientador
contributor.author	contributor.author	contributor.author	Autor
	contributor.event		Evento
coverage.spatial			Local de cobertura
coverage.temporal			Cobertura temporal
date.accessioned	date.accessioned	date.accessioned	Data e horário de inclusão no repositório
date.issued	date.issued	date.issued	Data de publicação
		degree.date	Data de defesa
		degree.department	Unidade universitária
		degree.grantor	Instituição
		degree.level	Grau acadêmico
		degree.local	Local da instituição de defesa
		degree.program	Programa de pós-graduação
description			Descrição
		description.abstract	Resumo/Abstract
format.extent	format.extent	format.extent	Tamanho do arquivo
format.medium			Características do original
format.mimetype	format.mimetype	format.mimetype	Tipo de arquivo
identifier.codmuseu			Número de identificação
	identifier.nrb	identifier.nrb	Número de sistema



Comunidades			Conteúdo no LUME
Acervo MU ¹ Acervo ESEF ²	Eventos	TD ³ TCCE ⁴	
<i>Metadados</i>	<i>Metadados</i>	<i>Metadados</i>	
identifier.uri	identifier.uri	identifier.uri	URI
	relation.ispartof		Contido em
subject	subject	subject	Palavra-chave
title	title	title	Título
	title alternative	title alternative	Título em outro idioma
type	type	type	Tipo de documento

¹ Acervo do Museu Universitário

² Acervo histórico da Escola de Educação Física

³ Teses e dissertações

⁴ Trabalhos de conclusão de curso de especialização

A utilização de padrões nacionais e internacionais para tratamento dos recursos de informação é uma prática há muito adotada na Universidade visando, sobretudo, à integração e compartilhamento de informações com outros sistemas.

Para garantir a interoperabilidade com outros repositórios digitais é utilizado o protocolo de coleta de metadados da iniciativa de arquivos abertos OAI-PMH - *Open Archives Initiative Protocol for Metadata Harvesting (THE OPEN...*, 2004), permitindo a transferência de dados entre repositórios digitais.

Este protocolo é construído sobre o protocolo *http* e é baseado em requisições do tipo *request-response* definidas utilizando a linguagem de descrição XML, o que lhe fornece compatibilidade praticamente universal.

3 Conclusão

O LUME está se estabelecendo como um projeto permanente na UFRGS e, como tal, é um serviço que requer continuidade. Neste sentido, já se evidencia a necessidade de definir a política institucional para inclusão dos novos tipos de documentos provenientes das unidades acadêmicas e de outros órgãos da Universidade, sejam estes documentos de acesso público ou de uso restrito.

Considerando a ampliação de sua abrangência para todo o acervo digital da Universidade planeja-se a aquisição de equipamento dedicado exclusivamente ao



sistema. Existe, também, possibilidade de propiciar uma maior autonomia aos usuários, permitindo que eles gerenciem suas informações, através do auto-arquivamento.

A experiência vivenciada até então é o ponto de partida para a consolidação do LUME como uma poderosa ferramenta auxiliar na organização, divulgação, preservação e visibilidade dos estoques de informação da Universidade.

Digital repository: free access to information at the Federal University of Rio Grande do Sul

Abstract: This paper describes the design and organization of the pilot project of LUME - Digital Repository of Federal University of Rio Grande do Sul (UFRGS), a joint initiative of the Central Library and the Computer Center of the University. It shows the resources of software and hardware and the patterns of representation concerning the information used in its implementation, aiming at the preservation and dissemination of its content and interoperability with other digital repositories.

Keywords: Digital repositories. Institutional repositories. Open archives initiative. Metadata. *DSpace*.

Referências

CAFÉ, Lígia; MELO, Bianca Amaro de; BARBOZA, Elza Maria Ferraz, et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO (INTERCOM), 26., 2003, Belo Horizonte. **Trabalhos apresentados**. 2003 Disponível em: http://dspace.ibict.br/dmdocuments/ENDOCOM_CAFE.pdf. Acesso em: 28 maio 2008.

CAMARGO, Liriane Soares de Araújo; VIDOTTI, Silvana Ap. Borseti Gregorio. Elementos de personalização em repositórios institucionais. In: Conferência Iberoamericana de Publicações Eletrônicas no Contexto da Comunicação Científica, 1., 2006, Brasília. **Trabalhos apresentados**. Disponível em:



http://dici.ibict.br/archive/00001077/01/cipecc_liriane.pdf. Acesso em: 5 jun. 2008.

CORPORATION FOR NATIONAL RESEARCH INITIATIVES. **The Handle System**. 2008. Disponível em: <http://www.handle.net/>. Acesso em: 3 jun. 2008.

DIGITAL initiatives: research and technology. **DSpace Manakin**. 2005. Disponível em: <http://di.tamu.edu/projects/xmlui>. Acesso em: 16 maio 2008.

DSPACE. **DSpace Institutional Digital Repository System**. Disponível em: <http://www.dspace.org/>. Acesso em: 20 maio 2008.

DUBLIN CORE METADATA INITIATIVE. **Using Dublin Core**. 2008. Disponível em: <http://www.dublincore.org/documents/usageguide/>. Acesso em: 28 maio 2008.

LEITE, Fernando César de Lima; COSTA, Sely. Repositórios institucionais como ferramenta de gestão do conhecimento científico no ambiente acadêmico. **Perspectivas em Ciência da Informação**, v. 11, n. 2, maio/ago. 2006. p. 206-219. Disponível em: <http://www.scielo.br/pdf/pci/v11n2/v11n2a05.pdf>. Acesso em: 12 maio 2008.

THE OPEN Archives Initiative Protocol for Metadata Harvesting. 2004. Disponível em: <http://www.openarchives.org/OAI/openarchivesprotocol.html>. Acesso em: 6 maio 2008.