

Repositório Digital da Universidade Federal do Rio Grande do Sul

Zaida Horowitz, Afonso Comba de Araújo Neto, Carla Metzler Saatkamp, Caterina Groposo Pavão, Janise Silva Borges da Costa, Lais Freitas Caregnato, Zita Prates de Oliveira

Universidade Federal do Rio Grande do Sul - UFRGS
Centro de Processamento de Dados
Rua Ramiro Barcelos, 2574
90035-003 Porto Alegre, RS
Tel: +55 (51) 3308-5029/3308-5071
Fax: +55 (51) 3308-5962
E-mail: comissao@cpd.ufrgs.br

Resumo: A Tecnologia da Informação constitui-se, cada vez mais, numa importante ferramenta para o manejo do conhecimento humano. A Universidade Federal do Rio Grande do Sul - UFRGS, através de uma iniciativa conjunta da Biblioteca Central e do Centro de Processamento de Dados está integrada a esta concepção. O presente trabalho descreve, sucintamente, o projeto-piloto do LUME - Repositório Digital da UFRGS. São abordados o histórico, os objetivos, a arquitetura, a tecnologia empregada e a implementação do mesmo na Universidade. Por fim, são comentadas algumas possibilidades futuras.

1 INTRODUÇÃO

As mudanças tecnológicas têm influenciado, de forma importante, a comunicação entre os pesquisadores dos meios acadêmicos. Apesar de coexistirem o meio escrito e o meio digital como formas de representar o conhecimento, a combinação da informática com as telecomunicações tem contribuído para acelerar a disponibilidade aos conteúdos. Os usuários demandam, cada vez mais, o acesso on-line às fontes de informação, em um contexto dinâmico que supere as restrições relativas a espaço geográfico, tempo, tamanho e extensão que têm caracterizado o acesso aos produtos e serviços de informação operados nos limites físicos das bibliotecas e centros de documentação tradicionais.

A Tecnologia da Informação tem um papel estratégico na Universidade Federal do Rio Grande do Sul - 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 disseminação e o uso do conhecimento gerado, a preservação desse conhecimento.

A gestão do conhecimento está estreitamente relacionada com os projetos de desenvolvimento de repositórios institucionais nas universidades. Estes colaboram para um melhor gerenciamento e maior acesso ao saber, além de aprimorar e potencializar a transferência deste, 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, apresentado neste 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 potencializar 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, ela agrega à descrição bibliográfica e temática, o conteúdo do documento, exigindo cuidados diferenciados para garantir sua integridade e, também, a observância da legislação de direito autoral.

A Biblioteca Digital de Teses e Dissertações da UFRGS - BDTD foi implementada em 2001, numa ação conjunta da Pró-Reitoria de Pós-Graduação, da Biblioteca Central e do Centro de Processamento de Dados. A necessidade de ampliação da BDTD, porém, torna-se relevante na tarefa de reunir o acervo digital das demais coleções existentes no SBU e dispersas em outros órgãos da Universidade, tais como, fotografias, partituras, discos sonoros, periódicos, entre outros, otimizando a organização, o gerenciamento, a manutenção e o compartilhamento de recursos. Aos usuários finais, possibilita a realização de buscas num único portal de informações.

Neste contexto, em 2007, teve início a ampliação da BDTD para repositório digital, 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, ampliando o uso desses recursos.

Os documentos digitais que integram as coleções podem conter texto, imagem, vídeo e áudio e são, em sua maioria, de acesso livre. Em alguns casos, o acesso é restrito à comunidade da UFRGS.

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 à comunidade em geral a sua produção científica, permitem maximizar o impacto das pesquisas, aumentam as possibilidades de arrecadação de fundos, obtenção de prêmios e prestígio compartilhado pelos pesquisadores e pela instituição.

2.3 Arquitetura e tecnologia

2.3.1 Softwares utilizados

O LUME foi desenvolvido com o uso do *DSpace - Institutional Digital Repository System* (DSpace, 2008), resultado de um projeto desenvolvido em parceria pelo MIT (*Massachusetts Institute of Technology*) e a HP (*Hewlett-Packard Company, USA*), orientado à criação de repositórios institucionais e à preservação digital de sua produção intelectual. Na UFRGS, o *DSpace* está instalado num 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.

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 uma 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.

Para o compartilhamento seguro de arquivos que estão em servidores Windows em rede é usado o software *Samba*.

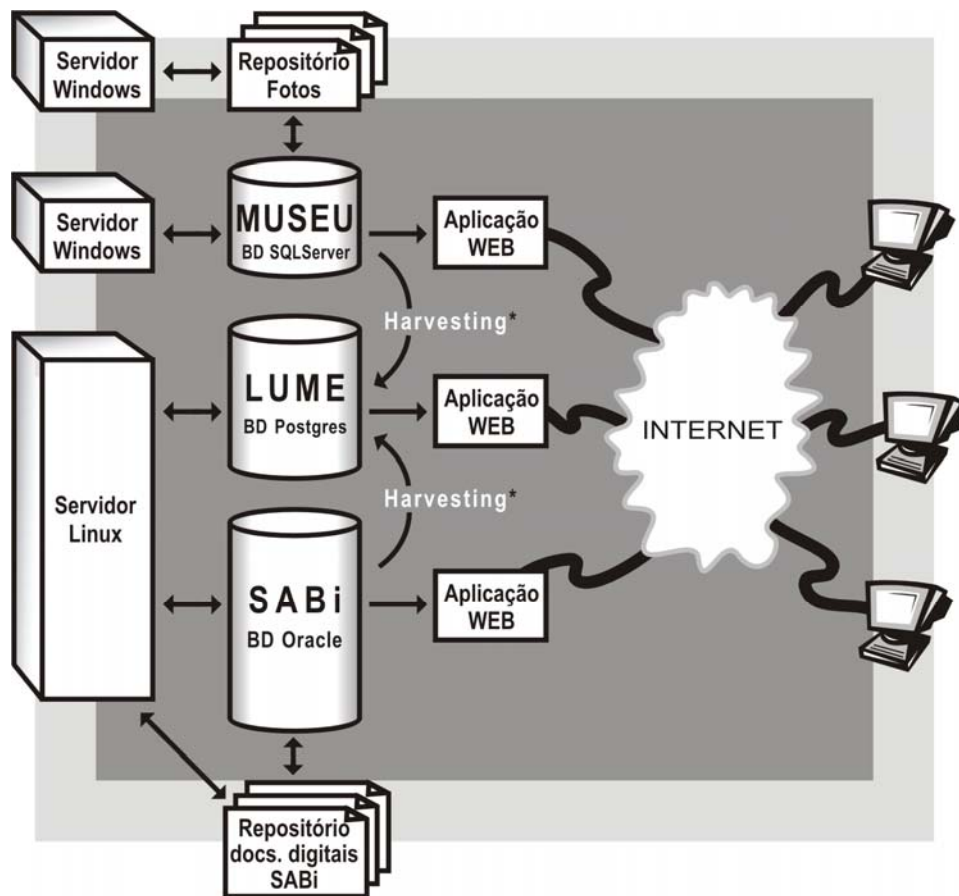
2.3.2 Ambiente

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, que 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 comercial 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 encontram-se, também, os documentos digitais.

A figura 1 apresenta a estrutura do LUME:



* *harvesting incremental*: coleta automática de metadados a cada 24 horas.

Figura 1 - Arquitetura do LUME.

2.3.3 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, palavras-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 propriamente dito. Esses dados podem ser exibidos de forma completa ou resumida. No LUME os metadados são definidos de forma diferente, conforme as necessidades de 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 Autor, Título, Assunto, Ano, como também a partir de filtros específicos de uma determinada Comunidade. É possível ainda fazer uma pesquisa genérica, em todo o repositório.

Para os documentos digitais de acesso restrito, em função dos direitos autorais, criou-se um “Login” específico, viabilizando que somente os usuários UFRGS tenham direito de acesso aos mesmos. A autenticação é feita através do software SSL.

O usuário pode criar seu perfil de interesse, mediante assinatura das coleções. Desta forma, receberá a relação dos novos itens adicionados por correio eletrônico.

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

As figuras 2, 3 e 4 ilustram algumas telas do LUME:



Figura 2 - Tela inicial do LUME.



Figura 3 - Tela de consulta à Comunidade Acervo do Museu Universitário.

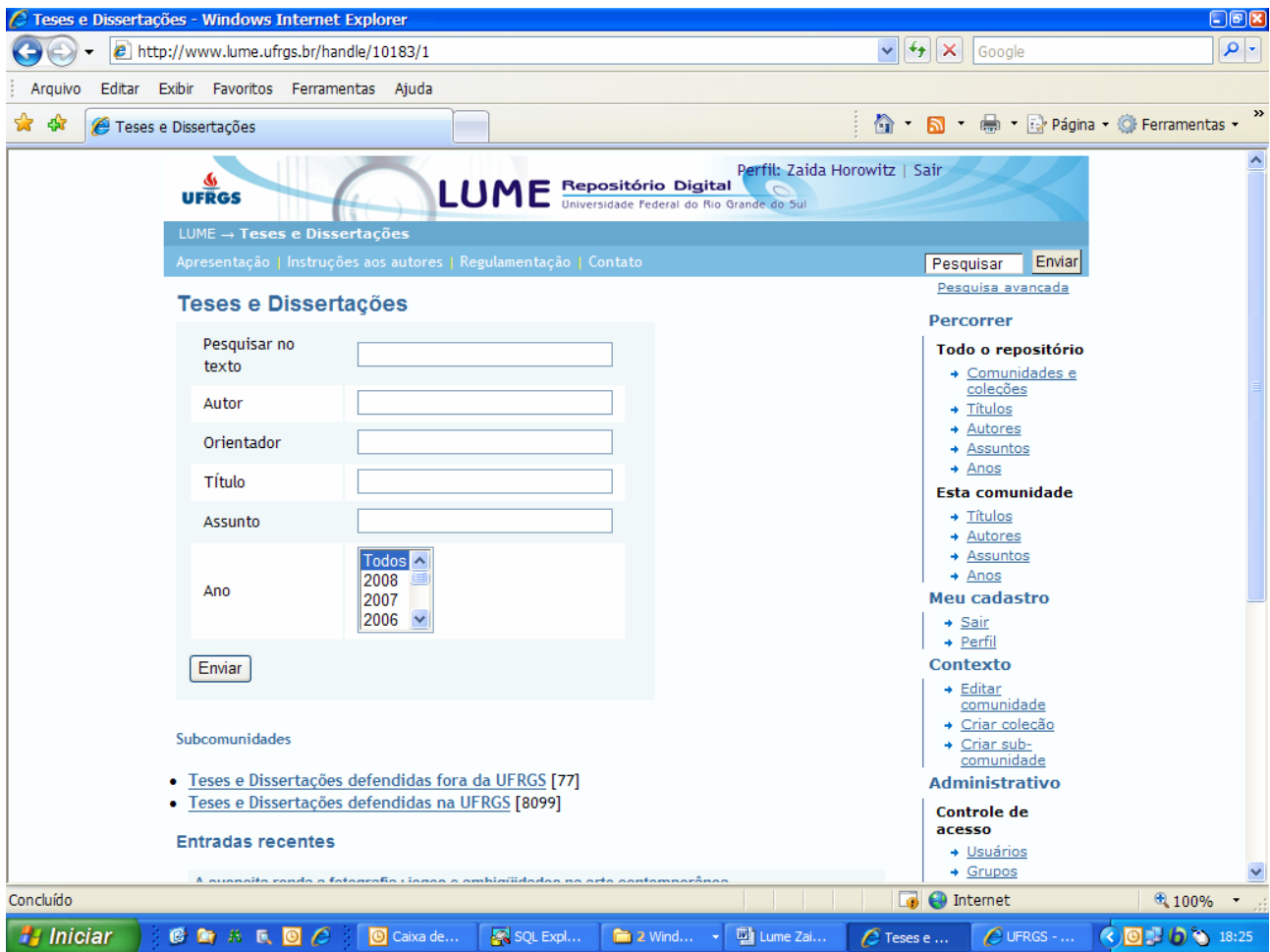


Figura 4 - Tela de consulta à Comunidade Teses e Dissertações.

2.3.4 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).

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

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 faz uso de transações *http* muito simples, baseadas em “request-response”, em formato XML.

3 Perspectivas futuras

O LUME é ainda um projeto-piloto sendo, portanto, de caráter experimental. Está se consolidando, no entanto, como um projeto permanente, devido à sua aceitação por parte da Administração Central da Universidade.

Como todo repositório digital é um serviço que requer continuidade. Neste sentido, já se evidencia a necessidade de criação de um Comitê Gestor, com atribuições 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, bem como estabelecer e acompanhar a implantação das prioridades propostas.

Tendo em vista a ampliação de sua abrangência para todo o acervo digital da Universidade planeja-se, igualmente, a aquisição de equipamento dedicado exclusivamente ao sistema. Existe também possibilidade de propiciar uma maior autonomia aos usuários, permitindo que eles gerenciem suas informações.

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.

4 Referências bibliográficas

CORPORATION FOR NATIONAL RESEARCH INITIATIVES. *The Handle System*. 2008. Disponível em: <<http://www.handle.net/>>. Acesso em: 03 mar. 2008.

DUBLIN CORE METADATA INITIATIVE. *Using Dublin Core*. 2008. Disponível em: <<http://www.dublincore.org/documents/usageguide/>>. Acesso em: 05 mar. 2008.

DSPACE. *DSpace Institutional Digital Repository System*. Disponível em: <<http://www.dspace.org/>>. Acesso em: 18 mar. 2008.

DIGITAL initiatives: research and technology. *DSpace Manakin*. 2005. Disponível em: <<http://di.tamu.edu/projects/xmlui>>. Acesso em 20 mar. 2008.

THE OPEN Archives Initiative Protocol for Metadata Harvesting. 2004. Disponível em: <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>. Acesso em: 17 mar. 2008.