

UM MODELO DE INTEGRAÇÃO ENTRE SISTEMAS DE INFORMAÇÃO NA UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL: EVENTOS E REPOSITÓRIO DIGITAL

*Manuela Klanovicz Ferreira, Afonso Comba de Araújo Neto, Zaida Horowitz, Adriana Jouris,
Caterina Groposo Pavão, Janise Silva Borges da Costa*

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

Tipo de trabalho: *Trabalho Técnico Consolidado*

Resumo: O trabalho aborda a estratégia e os detalhes referentes à integração entre o Sistema de Eventos Institucionais e o Lume, Repositório Digital da Universidade Federal do Rio Grande do Sul, tendo como “piloto” o Salão de Iniciação Científica promovido pela Universidade. Descreve aspectos relacionados à organização da informação e as etapas necessárias para a importação dos dados.

Palavras-chave: Integração entre sistemas; Eventos; Repositório institucional.

1 INTRODUÇÃO

O Centro de Processamento de Dados da UFRGS (CPD) desenvolve e mantém sistemas de informação nas várias áreas de atuação da Universidade, alguns dos quais foram concebidos para serem executados em computadores isolados e em redes locais. Os avanços proporcionados pelas tecnologias da informação e comunicação, no entanto, possibilitam a integração dos mesmos, de forma a evitar o retrabalho e duplicação de dados, promovendo sua consistência e uniformidade.

Integrar sistemas de informação distintos, mas que manipulam uma porção comum de dados, não é tarefa trivial. Por mais que os sistemas trabalhem com um conjunto comum de dados, cada sistema é desenvolvido para solucionar um problema administrativo diferente e, portanto, possui sua construção e arquitetura voltadas para uma perspectiva particular. Neste sentido, a adoção de normas e interfaces de comunicação padronizadas no desenvolvimento destes sistemas de informação é imprescindível para simplificar a sua interoperabilidade.

Neste trabalho, é apresentada a estratégia e os detalhes referentes à integração entre o Sistema de Eventos Institucionais (SEI) e o Repositório Digital da UFRGS, o Lume. Essa integração tem o objetivo de automatizar ao máximo a divulgação, no Lume, do material apresentado nos eventos geridos pelo SEI. O Salão de Iniciação Científica foi escolhido como o primeiro evento institucional a ser contemplado por essa integração por tratar-se de um dos maiores e mais representativos eventos de graduação na UFRGS. Sendo o Lume o principal portal de divulgação da produção intelectual da Universidade, o seu uso para a divulgação do material apresentado nos Salões de Iniciação Científica é auto-justificada. Porém, em tendo objetivos e modelos de dados distintos, o processo de integração de ambos os sistemas requereu a série de análises e decisões técnicas que são apresentadas neste trabalho.

Este artigo está organizado da seguinte forma: a seção 2 inclui uma breve explanação sobre o Salão de Iniciação Científica, que é o caso selecionado como foco do artigo; a seção 3 apresenta a arquitetura dos sistemas de informação envolvidos; a seção 4 apresenta a análise dos benefícios e as decisões necessárias para tornar possível a integração entre o SEI e o Lume e a seção 5 conclui o artigo.

2 SALÃO DE INICIAÇÃO CIENTÍFICA (SIC)

O Salão de Iniciação Científica (SIC) é realizado anualmente na Universidade Federal do Rio Grande do Sul (UFRGS), desde 1989, e constitui-se em um espaço para divulgação e acompanhamento dos trabalhos de iniciação científica desenvolvidos por alunos de graduação da Universidade e de outras Instituições de Ensino Superior, a fim de valorizar e dar visibilidade às pesquisas desenvolvidas no meio universitário.

Em sua primeira edição o SIC contabilizou 198 trabalhos inscritos e, em 2011, na sua XXIII edição, 2.181 trabalhos, mantendo uma média de 2.448 trabalhos apresentados nos últimos três anos, conforme apresentado na Figura 1.

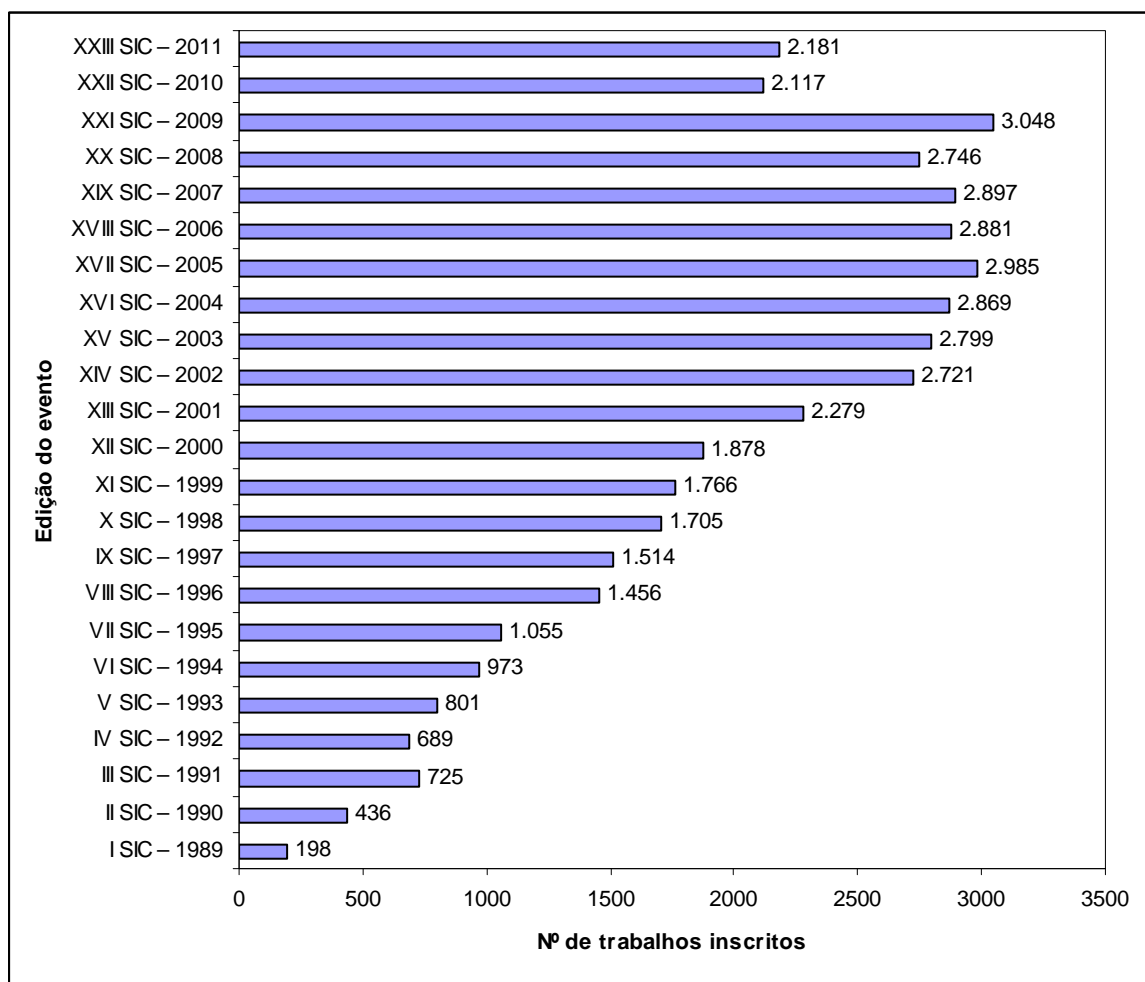


Figura 1 - Número de trabalhos inscritos no SIC, no período de 1989-2011.

Fonte: Pró-Reitoria de Pesquisa - PROPESQ.

O aumento do número de trabalhos inscritos é um indicativo relevante acerca da importância do SIC para a Universidade e, sobretudo, da visibilidade proporcionada às pesquisas desenvolvidas na UFRGS e demais Instituições de Ensino Superior do Estado, do país e do Mercosul. Atualmente, a gestão do Evento é realizada por meio do Sistema de Eventos Institucionais, desenvolvido pelo CPD, que, até então, não possui plataforma de divulgação de suas informações.

A expectativa é que a divulgação da produção científica envolvida nos salões, por meio do Lume, aumente ainda mais a visibilidade das pesquisas na Universidade e subsidie o processo de gestão da pesquisa na Instituição, contribuindo também para maximizar o uso do Lume.

3 SISTEMAS DE INFORMAÇÃO

Os sistemas de informação, objetos deste trabalho, são apresentados a seguir.

3.1 Sistema de Eventos Institucionais (SEI)

O desenvolvimento do Sistema de Eventos Institucionais teve início em 2009, por ocasião da organização do XXI Salão de Iniciação Científica, o qual serviu como “piloto” do sistema, abrangendo as etapas de submissão de trabalhos; a avaliação dos orientadores; a homologação das instituições externas; a inscrição de participantes; a organização das sessões de apresentação de trabalhos; a escolha de trabalhos premiados, entre outras. Até então, as informações referentes aos Salões de IC eram geridas por meio de um sistema específico para o Salão e que não servia aos demais eventos da Universidade.

O objetivo principal do SEI é que o mesmo seja um ambiente flexível, que possa ser usado para gestão dos diversos tipos de eventos promovidos pela Universidade como, por exemplo, conferências, seminários, congressos, salões, entre outros.

O sistema foi desenvolvido em PHP 5.2, com vasto uso de orientação a objetos, usando bibliotecas próprias do CPD para conexão e operações sobre o banco de dados, bem como a biblioteca *xajax* para simplificação de requisições assíncronas. A base de dados do SEI é o servidor SQL Server que dá suporte à base de dados institucional da Universidade, o que permite acesso direto a todas as informações de suporte que podem se fazer necessárias nos diversos tipos de eventos que o sistema irá gerir. Além dos metadados relativos aos eventos como um todo e aos trabalhos específicos de cada evento, os arquivos originais dos autores (pôsteres, resumos, etc.) também são mantidos.

Em 2010 foi iniciado o projeto de criação do Portal de Eventos Institucionais, um ambiente integrado que permite ao usuário acessar todos os serviços que estão à sua disposição. O sistema é responsável por identificar o perfil do usuário em cada evento disponível no sistema (participante, orientador, avaliador de trabalhos, organizador, gerente, etc.) e permitir o acesso aos serviços para o respectivo perfil.

A solução adotada para o SEI não é desenvolver um sistema *genérico*, no sentido de que deva prever todos os casos, e sim um sistema adaptável. Neste aspecto, o mecanismo de serviços e perfis é de extrema utilidade, permitindo a criação de novos serviços, ao invés da modificação dos já existentes, o que mantém a complexidade do sistema baixa e facilita sua adequação para novas situações. O mecanismo em que o sistema de adaptação se baseia é também bastante simples e depende apenas da execução de consultas SQL para liberar serviços, associar serviços a perfis, associar perfis a pessoas e assim por diante.(CANTO et al., 2011)

3.2 Lume - Repositório Digital da UFRGS

O Lume é o Repositório Digital da UFRGS, criado em 2008 com a finalidade de reunir o acervo digital produzido na Universidade ou que, por sua área de abrangência e/ou pelo seu caráter histórico, é de interesse da Instituição centralizar sua preservação e difusão. Utiliza o DSpace - Institutional Repository System (DSPACE, 2012), uma ferramenta aberta que disponibiliza um portal *web* combinado com um banco de dados para arquivamento e consulta dos acervos, suportando diversos formatos de arquivos e organizada por uma estrutura hierárquica de comunidades que se subdividem em coleções, o que permite uma organização lógica e intuitiva dos itens disponíveis no repositório sem prejudicar a sua diversidade.

Nos últimos anos, a divulgação e promoção da produção intelectual e dos acervos locais tornou-se um requisito fundamental para as universidades, e hoje é visto como uma necessidade para qualquer instituição que queira ter visibilidade e credibilidade por sua competência, tanto nacional quanto internacional. Essa publicização também tem um importante papel social no âmbito das universidades públicas, no sentido de que permite que as mesmas demonstrem para a sociedade de forma aberta e clara a sua importância e relevância, e assim justifiquem a sua existência.

O DSpace é um *software* livre que supre todas as propriedades que advêm destas necessidades de divulgação das universidades, incluindo de forma nativa as seguintes características:

- a) facilidade de tradução e internacionalização;
- b) facilidade de customização e adequação às necessidades internas;
- c) utilização de um padrão de armazenamento de metadados universal e extremamente difundido e aceito, o *Dublin Core*;
- d) facilidade de indexação do seu conteúdo pelos principais sites de busca (p. ex., Google, Bing, etc.);
- e) compatibilidade com absolutamente qualquer tipo de documento ou acervo que esteja em formato digital;
- f) navegação, pesquisa e consulta intuitiva, através de uma interface uniforme que não depende do usuário conhecer o conteúdo do repositório;
- g) exposição do conteúdo completo do repositório para *harvesting* automatizado, mediante um protocolo de comunicação e consulta aberto, o *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH, 2012).

4 INTEGRAÇÃO DOS SISTEMAS

A integração entre o SEI e o Lume é um projeto conjunto da Pró-Reitoria de Pesquisa, responsável pela promoção do SIC, e do Centro de Processamento de Dados da UFRGS. Foi realizada com o intuito de disponibilizar e divulgar, através do Lume, os trabalhos submetidos para o SIC por meio do SEI. A integração entre os dois sistemas foi positiva, sobretudo em três aspectos:

- 1) aproveitamento dos metadados incluídos no processo de submissão dos trabalhos, evitando que os mesmos precisem ser incluídos novamente no Lume;
- 2) facilidade na localização de um trabalho do SIC utilizando os recursos de busca do Lume, ao invés de folhear um grande arquivo em PDF;
- 3) facilidade na disponibilização dos trabalhos, evitando a necessidade de geração dos anais do SIC, feito anteriormente em PDF.

A importação dos metadados e dos arquivos dos trabalhos é realizada ao final de cada edição do SIC, a fim de disponibilizar no Repositório apenas os trabalhos efetivamente apresentados.

4.1 Organização da informação no Lume

Para abrigar os trabalhos provenientes do SIC, foi criada uma nova comunidade dentro do Lume, juntamente com uma hierarquia de subcomunidade e coleções. Foi definida a criação de uma subcomunidade específica para cada edição do evento e, dentro desta, uma coleção específica para cada grande área do conhecimento, de forma a agrupar os trabalhos de cada edição, conforme apresentado abaixo:

- Eventos UFRGS (comunidade)
 - Salão de Iniciação Científica (subcomunidade)
 - Salão de IC 2009 (subcomunidade)
 - Ciências Agrárias (coleção)
 - Ciências Biológicas (coleção)
 - Ciências da Saúde (coleção)

- Ciências Exatas e da Terra (coleção)
- Ciências Humanas (coleção)
- Ciências Sociais e Aplicadas (coleção)
- Engenharias (coleção)
- Linguística, Letras e Artes (coleção)
- Salão de IC 2010 (subcomunidade)
- ...
- Salão de IC 2011 (subcomunidade)
- ...

Também foi necessária a definição, no Lume, dos metadados que irão conter as informações relativas a cada um dos trabalhos do SIC. O mapeamento entre metadados e respectivos conteúdos é apresentado no Quadro 1. Os metadados são baseados no padrão *Dublin Core - Dublin Core Metada Initiative* (DUBLIN CORE, 2011), que é o utilizado pelo DSpace.

Quadro 1 - Mapeamento entre metadados e informações dos trabalhos do SIC

Nome do metadado	Conteúdo
dc.title	Título do trabalho
dc.contributor.author	Autor do trabalho
dc.contributor.advisor	Orientador do trabalho
dc.subject.session	Sessão de apresentação
dc.subject.theme	Temática do trabalho
dc.subject.cnpq	Grande área do conhecimento
dc.subject.macro	Macro temática do trabalho
dc.type.presentation	Modalidade de apresentação
dc.description.number	Número do trabalho
dc.identifier.sic	Identificador único do trabalho no SE
dc.date.issued	Ano do trabalho
dc.date.accessioned	Data de publicação do trabalho no Lume
dc.type	Tipo de trabalho
dc.format.mimetype	Mimetype dos arquivos do trabalho
dc.format.extent	Extensão dos arquivos do trabalho
dc.relation.ispartof	Evento (i.e., edição do SIC) ao qual o trabalho está relacionado
dc.language.iso	Idioma no padrão de três letras

4.2 Importando dados

Para executar a importação dos dados do SEI, que utiliza um banco de dados SQL Server, para o Lume, que utiliza um banco de dados PostgreSQL independente, existiam duas alternativas. A primeira era utilizar os mecanismos de importação disponibilizados pelo próprio DSpace. Para utilizar esta forma, seria necessário exportar os dados e os arquivos dos trabalhos dos salões para uma estrutura XML específica. A segunda, seria criar um programa que escrevesse diretamente no banco de dados do DSpace e esta foi a escolhida, por apresentar diversas vantagens, dentre as quais destacam-se: proporcionar maior flexibilidade na adaptação e formatação dos metadados, permitir a atualização dos itens já importados (o DSpace nativamente permite apenas importação e exclusão) e propiciar a redução no espaço em disco utilizado.

A implementação da importação foi feita através de um script PHP que consulta as informações diretamente no SEI e, para cada trabalho, executa os seguintes passos:

- 1) cria um *log* do processo de importação que permite detectar possíveis erros ocorridos durante o mesmo;
- 2) consulta a localização do arquivo PDF do texto e do pôster do trabalho no banco de dados do SEI, verifica a existência desses arquivos e os copia para uma área de armazenamento permanente acessível pelo Lume;
- 3) cria *links* simbólicos (i.e. atalhos) no sistema de arquivos do DSpace que apontam para o local de armazenamento permanente onde os trabalhos foram copiados no passo 2. Estes atalhos têm o objetivo de permitir que estes sejam acessíveis pelo DSpace atendendo às suas restrições mas sem a necessidade de duplicá-los;
- 4) consulta diretamente no banco de dados do SEI os metadados dos trabalhos, listados na Quadro 1. Antes de serem incluídos, estes dados são transformados, a fim de padronizá-los para o formato de exibição do Lume.

Neste ponto são executados três comandos fornecidos pela ferramenta DSpace e localizados dentro da pasta *bin* do diretório de instalação desta ferramenta.

- 5) executa o comando *cleanup* do DSpace para limpar registros antigos;
- 6) executa o comando *filter-media* do DSpace para criar as miniaturas (*thumbnails*) das imagens e PDFs que irão aparecer na busca de registros;
- 7) executa o comando *index-update* do DSpace para atualizar os índices de pesquisa para os novos itens.

A representação gráfica dos passos acima descritos é apresentada na Figura 2.

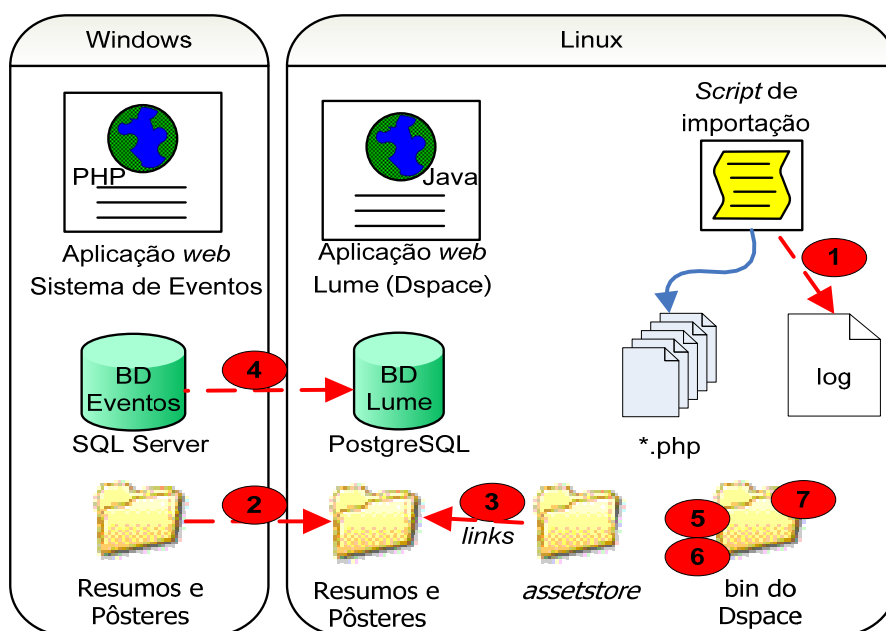


Figura 2 - Importação de dados do Sistema de Eventos Institucionais para o Lume.

5 CONCLUSÃO

A solução adotada para a importação dos trabalhos do SIC diretamente do SEI para o Lume vem reforçar a recomendação explícita no Plano de Desenvolvimento de Tecnologia de Informação (PDTI), elaborado pela Universidade para o período de 2011 a 2015, sobre a necessidade da adoção de soluções institucionais que permitam a disponibilização das informações geradas pelos diversos órgãos gestores da UFRGS de forma integrada, a fim de evitar a redundância de dados e a proliferação de informações sem a garantia de sua procedência, confiabilidade e responsabilidade (UNIVERSIDADE..., 2011).

A disponibilização dos trabalhos no Lume, por sua vez, amplia a visibilidade aos projetos de pesquisa desenvolvidos na Universidade e em outras Instituições de Ensino Superior, consoante

com os objetivos do próprio SIC, ao reunir os trabalhos em um único portal e, sobretudo, possibilitar a recuperação da informação de forma mais amigável e abrangente, por meio de ferramentas de busca, além de garantir a preservação dos trabalhos, tornando desnecessária a geração dos anais do SIC em PDF.

A integração e interoperabilidade de sistemas atualmente é, sobretudo, uma necessidade, diante da quantidade de informações e sistemas, imprescindíveis para a gestão e visibilidade das Instituições de Ensino Superior.

REFERÊNCIAS

DSPACE. **Dspace Institutional Digital Repository System**. 2012. Disponível em: <<http://www.dspace.org>>. Acesso em 23 jan. 2012.

DUBLIN CORE. **Dublin Core Metadata Initiative**. 2011. Disponível em: <<http://www.dublincore.org/documents/usageguide>>. Acesso em: 23 jan. 2012.

CANTO, F. H. et al. **Sistema de Eventos Institucionais**. 2011. Trabalho apresentado no V Workshop de Tecnologia da Informação das Instituições Federais de Ensino Superior, Florianópolis, 2011.

OAI-PMH. **Open Archives Initiative Protocol for Metadata Harvesting**. 2012. Disponível em: <<http://www.openarchives.org/pmh/>>. Acesso em: 24 fev. 2012.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL. **Plano de Desenvolvimento de Tecnologia de Informação (PDTI): relatório final 2011-2015**. [2011]. 127 f. Documento em tramitação no Conselho Universitário.