



A MIGRAÇÃO DA BIBLIOTECA DIGITAL DE TESES E DISSERTAÇÕES DA UNIVERSIDADE FEDERAL DE GOIÁS (BDTD/UFG): TÓPICOS A SEREM CONSIDERADOS

Carla Lopes Ferreira¹, Cláudia Oliveira de Moura Bueno², Maria Silvério da Silveira³

¹Bibliotecária, bolsista de pesquisa na Gerência de Informação Digital e Inovação (GIDI) da Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil. E-mail: carlaferreira66@gmail.com

¹Bibliotecária na Gerência de Informação Digital e Inovação (GIDI) da Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil. Doutoranda da Universidad de Murcia – ES. E-mail: claudia_bueno@ufg.br

³Bibliotecária, Diretora do Sistema de Bibliotecas da Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil. E-mail: msilverioufg@gmail.com

RESUMO

Este trabalho apresenta dois aspectos da Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade Federal de Goiás (UFG). O primeiro, descrevendo sua criação e as dificuldades impostas na formalização do processo de depósito na instituição, assim como as estratégias para solucioná-las. Os resultados demonstram que a BDTD/UFG possui mais de 2 mil Teses e Dissertações Eletrônicas (TDEs) no *software* TEDE, que são de grande importância para os Programas de Pós-Graduação e para a comunidade universitária como um todo, assim como na promoção do Acesso Aberto à informação técnico-científica brasileira. São demonstrados ainda, dados estatísticos durante o período de funcionamento do TEDE. No segundo aspecto, apresenta os problemas técnicos ocasionados com o TEDE, devido à desatualização e instabilidade do *software*, gerando a obsolescência da BDTD/UFG. Diante da situação, deu-se início aos estudos para escolha do *software* a ser utilizado para a migração e configuração dos dados, cujo resultado foi o DSpace, sistema utilizado no Repositório Institucional da UFG (RI/UFG). Foi definida a metodologia utilizada para a correção dos metadados importados do TEDE e a criação de novos *plugins*, para permitir a interoperabilidade dos dois sistemas. As vantagens da preparação para a migração e seus resultados alcançados até o momento são descritos ao final.

Palavras-chave: 1. Biblioteca digital - Migração. 2. Biblioteca digital – UFG. 3. Biblioteca digital - DSpace. 4. Biblioteca digital - TEDE. 5. Interoperabilidade.



ABSTRACT

This work presents two aspects of the Digital Library of Theses and Dissertations (BDTD) of Federal University of Goiás (UFG). The first aspect, describing its creation and the questions imposed on formalizing the guarding process of collections of these material, as well as strategies to solve them. The results demonstrate that BDTD/UFG has more than 2 thousands Electronic Theses and Dissertations (TDEs) in TEDE software, which are of great importance for Graduate Programs and the university community as a whole as well as in promoting Open Access brazilian technical and scientific information. Statistics are also demonstrated during the TEDE's run time. In the second aspect, presents caused technical problems with the TEDE due to outdated software and instability, which led to the inoperability of BDTD/UFG. Given the situation, has begun studies to choose the software to be used for migration and configuration data, which resulted in the DSpace system used in the UFG Institutional Repository (RI/UFG). The methodology used for the correction of imported TEDE metadata and creating new plugins to enable interoperability between the two systems was defined. The advantages of the preparation for migration and the results achieved to date are described at the end.

Keywords: 1. Digital library - Data migration. 2. Digital library - UFG. 3. Digital library - Dspace. 4. Digital library - TEDE. 5. Interoperability.

INTRODUÇÃO

A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) é um projeto do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) que “integra os sistemas de informação de teses e dissertações existentes nas instituições de ensino e pesquisa brasileiras e estimula o registro de publicação em meio eletrônico” (IBICT, 2008). Em 2005 a Biblioteca Central (BC), unidade do Sistema de Bibliotecas da Universidade Federal de Goiás (SIBI/UFG), e a Pró-Reitoria de Pesquisa e Pós-Graduação¹ (PRPPG), por meio da Câmara de Pesquisa e Pós-Graduação da UFG, iniciaram os estudos para a implantação da Biblioteca Digital de Teses e Dissertações da instituição. A ferramenta utilizada para possibilitar o funcionamento da BDTD foi o Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE), desenvolvido pelo próprio IBICT em duas versões: modular e simplificado.

¹Atualmente desmembrada em Pró-Reitoria de Pós-Graduação (PRPG) e Pró-Reitoria de Pesquisa e Inovação (PRPI).



Desde sua implantação até meados de 2012, o TEDE esteve funcionando normalmente, quando, devido à sua vulnerabilidade ocasionada pela desatualização do mesmo, o sistema foi alvo do ataque de *crackers*², tornando-se obsoleto a partir disso.

Assim, a Gerência de Informação Digital e Inovação (GIDI/BC), responsável pela gestão da BDTD na instituição, iniciou os estudos sobre a escolha do sistema que possibilitasse a migração das teses e dissertações eletrônicas (TDEs) depositadas no TEDE, para outro sistema que utilizasse o protocolo *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) possibilitando a coleta e disseminação dos metadados para repositórios. Primeiramente o módulo Biblioteca Digital, do software SophiA (Prima³) foi estudado para verificação se atendia às necessidades de um gerenciamento de teses e dissertações tal como era feito no TEDE. Após a constatação de que o SophiA não atendia as necessidades, o software DSpace (Duraspace, MIT e HP⁴) foi estudado e definido como satisfatório. Como a UFG possui o Repositório Institucional (RI) e este é gerenciado pelo DSpace, iniciou-se o processo de migração das TDEs que estavam no TEDE para o RI.

Os Repositórios Institucionais são sistemas de informação que possibilitam o acesso à produção da instituição, sem que haja obstáculos de custo para tal, além de oferecem benefícios ao pesquisador, para os administradores acadêmicos, para as instituições e para a comunidade científica (UNIVERSIDADE DE MANCHESTER, 200?, apud LEITE, 2009, p. 23). Eles disseminam a produção intelectual de determinada instituição, possibilitando a disponibilização dos tipos de documentos que forem definidos em sua Política de Informação, como: artigos científicos, trabalhos publicados em eventos, relatórios de pesquisa, teses, dissertações, monografias, dentre outros.

Atualmente a BDTD está em processo de migração para o RI, em fase de finalização e espera-se que até o final de 2014 todas as TDEs sejam disponibilizadas na página do RI e novos registros comecem a ser inseridos pela equipe. Este estudo aborda o processo de preparo e início da migração do banco de dados do TEDE para o RI, as dificuldades encontradas e as alterações realizadas para melhorar o sistema e atender às especificações das teses e dissertações.

² Termo designado a indivíduos que quebram um sistema de segurança de forma ilegal ou antiético.

³ www.prima.com.br

⁴ www.dspace.org



CRIAÇÃO E IMPLANTAÇÃO DA BIBLIOTECA DE TESES E DISSERTAÇÕES DA UFG

Em 2006 o IBICT lançou edital para fomentar a instalação de novas bibliotecas digitais em Instituições Federais de Ensino Superior de todo país. A UFG, dentre várias outras instituições, concorreu e foi selecionada para participar do projeto. A Biblioteca Digital Brasileira de Teses e Dissertações da Universidade Federal de Goiás (BDTD-UFG) foi criada em 29 de junho de 2006 e teve como objetivo disponibilizar via internet as dissertações e teses produzidas pelos programas de pós-graduação da UFG. A versão utilizada do TEDE foi a simplificada.

Para dar início à implantação da BDTD da UFG três unidades acadêmicas foram definidas como piloto. Desta forma os técnicos envolvidos nesta ação teriam maior familiaridade com as tecnologias e os procedimentos necessários à realização da mesma. No piloto foram observados os seguintes problemas: falta de resumos em português e em outra língua; resumos fora das normas; falta de palavras-chave em português e em outra língua; ausência de título em outra língua; ausência da catalogação na fonte ou catalogação fora das normas; formato digital diferente do formato impresso; ausência do termo de autorização; e arquivo protegido.

Além de todas estas questões, houve grande resistência por parte dos autores e de seus orientadores na disponibilização do CPF, necessário para acesso ao TEDE. Observou-se que eram necessárias melhorias na relação entre a biblioteca e os programas de pós-graduação, pois estes não verificavam se os documentos enviados pelos autores estavam condizentes com as normas de publicação da ABNT.

Evolução e estatísticas do TEDE

O objetivo do Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE) é proporcionar a implantação de bibliotecas digitais de teses e dissertações nas instituições de ensino e pesquisa e sua integração a BDTD nacional; gerando bibliotecas digitais institucionais. O *software* é gratuito, distribuído pelo IBICT em um pacote com o Sistema TEDE já implementado, o MTD-BR e a camada do Protocolo OAI-PMH, metodologia de implantação, os manuais operacionais e de usuário, documentação e treinamento. (BDTD, 2006).

A alimentação do TEDE na UFG teve início em 2007 depois dos treinamentos e preparação do material. Até 2009 apenas 81 documentos haviam sido inseridos no sistema, já que a exigência do CPF dos autores ocasionou dificuldades e atraso na alimentação do banco de dados.

Em 2009 a Gerência de Informação Digital e Inovação (GIDI) assumiu a coordenação da BDTD e para otimizar o funcionamento da biblioteca digital, utilizou a seguinte metodologia de trabalho: a Biblioteca Central, em conjunto com a Pró-Reitoria de Pesquisa e Pós-Graduação, trabalhou uma normativa interna descrevendo quais eram as informações essenciais para a disponibilização do documento. Em seguida foram coletadas, por meio de uma tabela separada por programa, todas as pendências referentes a cada dissertação e tese encaminhadas para depósito nas bibliotecas da UFG. De posse deste material foram adotadas duas estratégias para solucionar as pendências: 1- enviar para o coordenador ou secretário de cada programa de pós-graduação, via *e-mail*, a listagem dos trabalhos e respectivos problemas encontrados em cada um juntamente com uma carta explicativa sobre a BDTD; 2- contatar por telefone cada coordenador ou secretário para esclarecer o assunto e solicitar apoio na solução das pendências.

Após um determinado período a equipe envolvida avaliou que as estratégias não apresentaram resultados satisfatórios. Decidiu-se então agendar visitas a cada um dos programas de pós-graduação, o que apresentou resultado muito positivo. Também foi decidido que, para resolver a questão do CPF, o respectivo campo seria retirado do Termo de Autorização e Ciência do autor e a coleta do mesmo seria feita utilizando os dados fornecidos pela própria UFG. Para coletar essa informação, foram traçadas duas estratégias: 1 - usar o Cadastro de Pessoas da UFG; 2 - solicitar, à Pró-Reitoria de Pesquisa e Pós-Graduação da UFG, autorização para consultar o Sistema de Cadastro da Pós-Graduação (SisPG). Como o uso do Cadastro de Pessoas da UFG não foi eficaz, a utilização do SisPG resolveu a questão do CPF.

Devido às pendências encontradas na maioria dos documentos enviados para as bibliotecas da UFG, até março de 2009 constavam apenas 81 documentos depositados na BDTD. Com a retirada do CPF do formulário e a instalação do sistema de consulta da pós-graduação da UFG, em agosto de 2009 notou-se um ganho em quantidade de documentos disponibilizados na BDTD da UFG, assim como nos anos seguintes. O quadro abaixo apresenta o total de documentos inseridos no sistema por ano, de 2008 a 2013.

Quadro 1 - Teses e dissertações inseridas no TEDE de 2008 a 2013

2008	2009	2010	2011	2012	2013	TOTAL
81	417	430	455	574	469*	2326

Fonte: Relatórios GIDI

* Inseridas no TEDE até junho de 2013



Como consequência da desatualização do TEDE⁵ e de problemas técnicos, o banco de dados tornou-se vulnerável e foi alvo da ação de *crackers* em outubro de 2012. O acesso ao público foi removido, ficando somente a inserção de novos documentos. Em setembro de 2013 verificou-se que o TEDE não oferecia diversas funções, tornando-se inoperante. Desta forma, não era possível realizar buscas no TEDE e nem inserir novos documentos.

Insatisfeitos com a situação, estavam os autores, pela não disponibilização dos seus documentos, assim como os servidores que viam parte do seu trabalho acumular, sem poder dar prosseguimento no serviço.

Preparação para migração para o DSpace

Diante da situação, a equipe decide fazer a migração dos dados para outro *software*, sendo necessário, portanto, um estudo que apresentasse qual sistema melhor atenderia às necessidades de uma biblioteca digital de teses e dissertações. Os critérios para escolha do software era ser um sistema capaz de submeter, gerenciar, editar e disponibilizar teses e dissertações e possuir o protocolo OAI-PMH⁶ para fazer a coleta e disseminação dos metadados para outros repositórios de dados, permitindo assim, a disseminação dos documentos em âmbito nacional e internacional. Assim, foram analisados os sistemas SophiA (utilizado no Sistema de Bibliotecas da UFG) e o DSpace (utilizado no Repositório Institucional).

Em outubro de 2013 a equipe da GIDI realizou um treinamento *online* sobre o novo Módulo de Biblioteca Digital do sistema SophiA, que poderia ser utilizado para gerenciar as teses e dissertações. A partir desse treinamento, percebeu-se que o sistema não era satisfatório, pois os recursos necessários para alimentação, gestão e disseminação dos dados eram insuficientes à instituição, como o esquema de metadados para teses e dissertações, que descreviam apenas de forma básica esses documentos.

⁵ Os desenvolvedores do TEDE não atualizaram o sistema desde 2010.

⁶ OAI-PMH (*Protocol for Metadata Harvesting*) é um protocolo que permite a interoperabilidade dos conteúdos armazenados em arquivos abertos.



A próxima ação foi verificar se o DSpace atendia tais necessidades e se seria possível migrar o banco do TEDE para esse sistema. O estudo, realizado em fevereiro, demonstrou que o *software* era capaz de gerenciar teses e dissertações, através do processo de *workflow* (submissão, revisão, edição e depósito) de forma prática e segura, além de utilizar opção de embargo e restrição para documentos específicos, como havia no TEDE. O sistema também utiliza o esquema de metadados Dublin Core⁷, que é abrangente e pode ser modificado para melhor atender às necessidades de cada instituição. Além dessas funcionalidades, o DSpace é o *software* do Repositório Institucional da UFG desde 2009, permitindo o depósito dos documentos técnico-científicos institucionais.

Um repositório institucional (RI) “visa ao melhoramento da comunicação científica. Para isso, provêm os mecanismos que aumentam tanto a eficácia da preservação da produção intelectual de pesquisadores e instituições acadêmicos quanto a visibilidade de ambos.” (LEITE; COSTA, 2006, p. 213). Como definição de RI, Lynch afirma: “é um conjunto de serviços que a universidade oferece aos membros de sua comunidade, visando o gerenciamento e disseminação dos materiais digitais criados pela instituição e pelos membros de sua comunidade” (2003, apud LEITE; COSTA, 2006).

Esses serviços têm como característica permitir o armazenamento e a disseminação de tipos de documentos diversificados, tais como: artigos científicos, livros eletrônicos, capítulos de livros, relatórios técnicos, trabalhos publicados em eventos, trabalhos de conclusão de curso, objetos educativos, teses, dissertações e outros.

Entremeio aos estudos a cerca de uma possível migração do TEDE para o DSpace, a equipe é informada que o IBICT estava trabalhando na construção de *plugins* capazes de facilitar e permitir uma migração mais eficaz entre os dois sistemas; o que foi determinante para que a equipe finalmente escolhesse migrar os dados para o repositório institucional.

A partir de então, a equipe se mobilizou no levantamento de estratégias para tornar a migração entre os dois sistemas mais eficaz, assim como facilitar a inserção de novos documentos. Dentre elas:

⁷ É um esquema de metadados para descrição de objetos digitais, tais como, vídeos, sons, imagens, textos e sites na web, composto de 15 elementos básicos, dentre eles (criador, data, assunto, descrição, tipo, formato etc).

- Modificação e especificação dos metadados no repositório institucional, que eram mais básicos do que os existentes no TEDE – novos campos foram adicionados, como: campos para identificação de participantes de banca, orientador e co-orientador, CPF ou matrícula do autor, país, unidades/departamentos, nome do programa de pós-graduação e outros.
- Padronização dos campos de metadados conforme as diretrizes OpenAIRE 2.0⁸, DRIVER 2.0⁹ e do livro ‘Boas práticas para Repositórios Institucionais’¹⁰ (IBICT, 2012) – essas orientações são direcionadas a provedores de conteúdo que visam melhorar a qualidade dos metadados fornecidos durante a catalogação dos objetos digitais, em iniciativas de arquivos abertos. (SHINTAKU; ALMEIDA; MELO, 2011). Nessa etapa, fez-se um estudo acerca da situação do Repositório Institucional em relação às diretrizes acima citadas, que resultou na inserção do campo ‘direitos autorais’ (dc.rights), na adoção da padronização dos títulos dos arquivos conforme o título do documento, correção do campo ‘tipos de documentos’ e no diagnóstico da digitalização para verificação da qualidade dos arquivos digitalizados. Algumas recomendações necessitam ainda serem estudadas e implementadas.
- Identificação dos problemas técnicos de infraestrutura do servidor onde opera o DSpace.

A preparação para o processo permitiu que alguns erros fossem percebidos e solucionados antes da migração, tais como: problema de tradução que havia na versão 3.2 do DSpace, na etapa de preenchimento dos metadados e correção do serviço de *harvesting*¹¹, que não estava funcionando adequadamente.

O IBICT auxiliou no procedimento de customização e migração do TEDE para o DSpace, sendo que essa cooperação foi de extrema importância no sucesso da migração e importação, visto que para que esse ideal ocorra, é necessário conhecimento aprofundado no sistema e um aplicativo que permita a comunicação entre os dois *softwares*.

⁸ <http://www.openaire.eu/>

⁹ <http://www.driver-support.eu/>

¹⁰ <http://livroaberto.ibict.br/handle/1/703>

¹¹ Coleta dos metadados para disseminação.



MATERIAIS E MÉTODOS

A preparação para o processo de migração das teses e dissertações envolveu esforços por parte da equipe da GIDI (SIBI/UFG), do Cercomp (Centro de Recursos Computacionais) e do IBICT. O IBICT desenvolveu e customizou um *plugin* capaz de migrar todos os documentos que estavam no TEDE para o DSpace, versão 4.0. A migração foi realizada em ambiente de testes e depois no servidor do repositório, sendo este armazenado no *data center* da UFG.

Foram realizadas três etapas para que o procedimento obtivesse êxito:

1. Preparação para a migração: nesta etapa foram estudados os dois sistemas e o IBICT desenvolveu e customizou o *plugin* que permitisse a comunicação entre os dois *softwares*. Foram realizados testes em março, que demonstraram sucesso, porém algumas melhorias ainda eram necessárias no aplicativo.
2. Processo de migração: após a finalização do *plugin*, novos testes foram realizados, demonstrando que havia uma maior compatibilidade entre os sistemas e os metadados eram mais específicos para teses e dissertações.
3. Processo de importação e finalização dos documentos: todos os documentos migrados do TEDE encontravam-se na opção de Importação, pelas ferramentas do administrador do repositório (DSpace). Sendo assim, após a finalização da migração, ainda era necessário que as coleções fossem criadas no DSpace, de forma que o Repositório de Teses e Dissertações da UFG tivesse a mesma estrutura que o Repositório Institucional, sendo:

comunidades – regionais das cidades;

subcomunidades – unidades acadêmicas;

subcomunidades – programas de pós-graduação e

coleções – tipos de documentos (mestrado em...
doutorado em...)

Após a criação da estrutura, todos os documentos aglomerados na coleção própria da migração foram importadas e movidos para suas coleções respectivas.



CONSIDERAÇÕES FINAIS

Com a invasão do TEDE confirmou-se sua vulnerabilidade e serviu de alerta para o estudo de um sistema que melhor atendesse a demanda tanto da biblioteca com dos usuários, porque havia muitas críticas em relação a estratégia de busca e inserção de dados no sistema. A definição pelo DSpace foi natural depois que o software gerenciador SophiA não atendeu às expectativas da equipe. Foi gratificante para a UFG ser piloto junto ao IBICT na migração do TEDE para o DSpace. A execução da migração obteve uma taxa de acertos de 96%.

REFERÊNCIAS

BUENO, Cláudia Oliveira de Moura; SILVA, Odete Jacomini da; FERREIRA, Liliene Juvência Azevedo. Biblioteca digital de teses e dissertações e os programas de pós-graduação: uma interação necessária. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS - SNBU, 16., 2010, Rio de Janeiro. *Anais eletrônicos* do XVI Seminário Nacional de Bibliotecas Universitárias, Rio de Janeiro: UFRJ, 2010a. n.p. Disponível em: <<http://www.bc.ufg.br/sophia/bc/publicacoes/bdtd.pdf>>. Acesso em: 16 jul. 2014.

_____ et al. O Portal de Periódicos da UFG. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS - SNBU, 16., 2010, Rio de Janeiro. *Anais eletrônicos* do XVI Seminário Nacional de Bibliotecas Universitárias, Rio de Janeiro: UFRJ, 2010b. n.p. Disponível em: <http://www.gapcongressos.com.br/eventos/z0070/trabalhos_pesquisa.asp?pag=1>. Acesso em: 20 jun. 2014.

CUNHA, Murilo Bastos. Construindo o futuro: a biblioteca universitária brasileira em 2010. *Ciência da Informação*, Brasília, v. 29, n.1, p.71-89, jan./abr. 2000.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Biblioteca Digital Brasileira de Teses e Dissertações. Brasília: IBICT, 2010. Disponível em: <<http://www.ibict.br>>. Acesso em: 25 jun. 2014.

KURAMOTO, Hélio. *BDTD: uma questão de interoperabilidade humana?* [Online]. Disponível em: <<http://kuramoto.blog.br/2006/04/14/bdtd-uma-questao-de-interoperabilidade-humana>>. Acesso em: 20 ago. 2014.

LEITE, Fernando César Lima. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Brasília, DF: IBICT, 2009. Disponível em: <http://www.ibict.br/anexos_noticias/repositorios.institucionais.F.Leite_atualizado.pdf>. Acesso em: 15 jul. 2014.



_____; COSTA, Sely. Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico. *Perspect. ciênc. inf.*, Belo Horizonte, v.11 n.2, p. 206 -219, mai./ago. 2006.

MARCONDES, Carlos H.; SAYÃO, Luis F. Interação e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: proposta da Biblioteca Digital Brasileira. *Ciência da Informação*, Brasília, v. 30, n.3, p. 24-33, set./dez. 2001.

_____; _____. À guisa de introdução: repositórios institucionais e livre acesso. In: SAYÃO, L. F. et al. *Implantação e gestão de repositórios institucionais : políticas, memória, livre acesso e preservação*. Salvador: EDUFBA, 2009. p. 9-19.

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

SILVA, Odete Jacomini da. *Projeto para implantação da biblioteca digital de teses e dissertações na Universidade Federal de Goiás – TEDE-UFG*. Goiânia: UFG, 2006.