RepositóriUM – implementação do DSpace em português: Lições para o futuro e linhas de investigação

ELOY RODRIGUES[1]; ANA ALICE BAPTISTA[2]; ISABEL RAMOS[2]; MARIA FERNANDA SARMENTO E SOUZA [3]
[1] Serviços de Documentação da Universidade do Minho, Portugal. eloy@sdum.uminho.pt
[2] Departamento de Sistemas de Informação (DSI), Universidade do Minho, Portugal. analice@dsi.uminho.pt, iramos@dsi.uminho.pt
[3] Estudante de Doutoramento DSI, Universidade do Minho, Portugal; bolsista do Governo Brasileiro – CAPES – Brasil; Universidade Estadual Paulista, Brasil. fsarmento@dsi.uminho.pt

RESUMO

Neste trabalho descrevemos a implementação do Repositório Institucional (RepositóriUM) da Universidade do Minho (UMinho), Portugal. A UMinho traduziu e está a utilizar a primeira versão em língua portuguesa do DSpace, um programa em código aberto desenvolvido numa parceria entre o Massachusetts Institute of Technology (MIT) e a Hewlett-Packard (HP). As cinco fases definidas para o plano de implementação são descritas neste trabalho (1. instalação, tradução e configuração; 2. carregamento de teses e dissertações; 3. instituição das comunidades piloto; 4. abertura do repositório ao público, e 5. abertura para toda comunidade da UMinho), bem como as razões que conduziram à escolha das seis comunidades piloto de diferentes áreas do conhecimento para participarem nesse período de implementação (Sistemas de Informação; Engenharia de Polímeros; Engenharia Biológica; Gestão, e uma da área de Ciências Humanas e outra da área de Ciências Sociais). Neste artigo pretendemos partilhar as lições aprendidas com o processo, e os caminhos de investigação que o projecto abriu, em particular para investigação e comparação de comportamentos de comunidades com diferentes culturas, de acordo com algumas questões e observações que surgiram durante os primeiros dez meses de experiência. Tendo estas informações como base, novas investigações começam a ser desenvolvidas no contexto de projectos de mestrado ou doutoramento na UMinho.

Palavras-chave: Acesso livre; comunicação científica; repositório institucional; DSpace.

INTRODUÇÃO

A criação do RepositóriUM foi decidida no âmbito do projecto Campus Virtual (e-UM) da Universidade do Minho, submetido à iniciativa da Universidade Eletrónica (e-U), estabelecida pelo Governo Português. O projecto foi executado pelos Serviços de Documentação (SDUM), unidade organizacional responsável pelas bibliotecas da UMinho e autores da proposta do projecto.

A primeira etapa foi um estudo do estado da arte sobre repositórios institucionais, e a avaliação das plataformas em código aberto[i] que permitem a criação desse tipo de sistema.

Em Abril de 2003 decidiu-se utilizar o DSpace (http://www.dspace.org) - um sistema em código aberto desenvolvido numa parceria entre o MIT e a HP – para a criação do Repositório Institucional da Universidade do Minho. O DSpace foi lançado em 4 de Novembro de 2002 e é definido por seus criadores como "um sistema inovador de bibliotecas digitais para recolha, armazenamento, indexação, preservação e redistribuição, em formato digital, da produção intelectual de comunidades universitárias"(1). Entre outras características, o DSpace é um sistema em código aberto 1) com uma arquitetura de *software* simples mas eficaz, 2) utiliza tecnologia recente, 3) desenvolvido para ser implementado em repositórios institucionais e 4) muito direccionado para o acesso aberto à publicação académica. Adicionalmente, foi criada uma Federação DSpace, composta por algumas das mais importantes universidades Norte

Americanas, e este facto por si só é considerado como uma prova da importância e confiança de que o sistema será mantido e desenvolvido nos próximos anos.
A UMinho foi a primeira universidade da comunidade mundial de língua portuguesa (de aproximadamente 200 milhões de pessoas) e uma das primeiras na Europa a traduzir e implementar o DSpace.
O RepositóriUM (http://repositorium.sdum.uminho.pt) está a ser utilizado desde Julho de 2003. Neste artigo desejamos compartilhar as lições aprendidas nesses dez meses de experiência e apontar algumas linhas de investigação, em particular inter-cultural.
Não é objectivo deste artigo descrever as particularidades tecnológicas do DSpace, nem as razões da sua utilização na Universidade do Minho, que são aspectos tratados em outros documentos de relevo para esse trabalho (1,2,3).
Na próxima secção descreveremos a implementação do DSpace na UMinho, o que resultou no repositório institucional (RepositóriUM). A seguir apresentaremos uma análise da adesão ao sistema pelas comunidades escolhidas; apontaremos alguns problemas encontrados durante a implementação, e delinearemos algumas linhas de investigação. Após as conclusões falaremos sobre alguns trabalhos que já estão a ser desenvolvidos na Universidade do Minho.

IMPLEMENTAÇÃO DO DSPACE

No âmbito das actividades do Campus Virtual (e-UM), os Serviços de Documentação da UMinho (SDUM), iniciaram as actividades do DSpace em Maio de 2003, com a sua tradução. O plano de implementação consistiu de 5 fases principais:

1. Instalação, tradução e configuração – Esta fase inclui todos os passos realizados desde que o *software* foi instalado nos nossos computadores até ele estar pronto para o amplo uso da comunidade.

2. Carregamento de teses e dissertações – alimentação do sistema com teses e dissertações desenvolvidas na UMinho. Ex-estudantes de mestrado e doutoramento da UMinho foram convidados a enviar para os SDUM uma cópia digital da sua tese, dissertação ou artigos, bem como, a autorização para publicação na Internet. Estes foram carregados no sistema e os seus metadados foram editados e validados de acordo com regras previamente definidas. Esta fase teve início em Julho de 2003.

3. Estabelecimento de comunidades piloto – seis comunidades piloto foram convidadas pelos SDUM a participar do RepositóriUM com os seus recursos de carácter científico, como artigos submetidos a conferências ou periódicos, literatura cinzenta, etc. Quatro dessas seis comunidades aceitaram participar e enviaram recursos aos SDUM para serem incluídos no sistema. Duas comunidades não aceitaram fazer parte do grupo inicial e enviar documentos para o RepositóriUM.

4. Divulgação do repositório – em 20 de Novembro de 2003, e apenas com uma pequena alteração no calendário previamente planeado, o RepositóriUM tornou-se público. Foram organizadas uma cerimónia e entrevista colectiva com a imprensa e o Reitor da UMinho.

5. Abertura formal para toda comunidade da UMinho – O RepositóriUM foi aberto para todas as comunidades da UMinho desde o comunicado da sua abertura pública. Mas, devido ao período de férias, somente em Janeiro de 2004 uma chamada geral para participação foi divulgada. O estabelecimento de comunidades

piloto (fase 2) foi a chave para esta fase, uma vez que algumas lições aprendidas naquela fase são úteis nesta.

A fase 1 foi praticamente técnica, correu tranquilamente e dentro do prazo. As principais dificuldades surgiram de pequenas adaptações resultantes do contexto organizacional da UMinho que difere do existente no MIT. Algumas traduções também resultaram em modificações no código. Todas as alterações realizadas foram cuidadosamente documentadas. Além disso, foi decidido desde o início que não seriam realizadas grandes alterações no código do DSpace, de forma que o sistema pudesse ser facilmente actualizado.
Do ponto de vista tecnológico tudo correu como esperado. No entanto, as questões sociais levantadas não são fáceis de solucionar. As outras fases do plano resultaram na identificação de alguns problemas, alguns dos quais foram resolvidos, enquanto outros necessitarão de estudos mais aprofundados e soluções de carácter estratégico.

ANÁLISE DA ADESÃO DAS COMUNIDADES AO SISTEMA

Foram efectuados convites para o depósito de teses e dissertações no RepositóriUM. As primeiras foram enviadas através da lista de correio electrónico da Universidade (UM-Net) no início de Julho de 2003, resultando no recebimento de teses e dissertações, sendo que alguns desses documentos foram incluídos nos últimos dias do mês de Julho, após o tratamento dos mesmos pelos SDUM. Em Agosto, não ocorreram convites e somente dois documentos foram submetidos ao repositório. O convite seguinte ocorreu em 18 de Setembro durante uma reunião realizada nos SDUM. A partir de 29 de Setembro, foi incluído um aviso na página inicial dos SDUM, sobre o depósito de teses e dissertações.
Observámos que as pessoas somente costumavam enviar os seus documentos nos dias próximos dos convites dos SDUM. Alguns dias depois ocorria uma grande diminuição no envio de documentos, aumentando novamente após outra solicitação, duas ou três semanas mais tarde, e assim sucessivamente, até Setembro. As Figuras 1 e 2 mostram a relação entre as chamadas dos SDUM, a recepção dos documentos e a colocação dos mesmos no repositório.

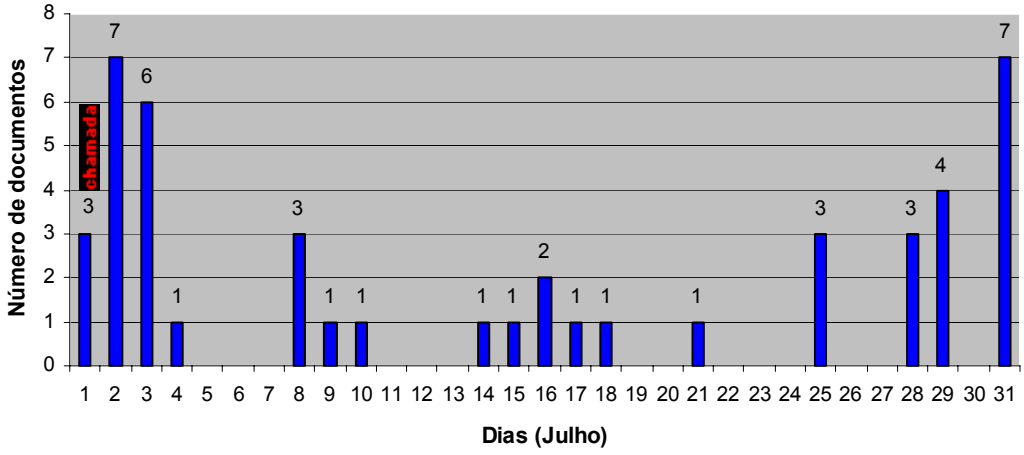

FIGURA 1 – Número de documentos submetidos ao RepositóriUM depois de 1 de Julho de 2003, quando ocorreu o primeiro convite à comunidade da UMinho.

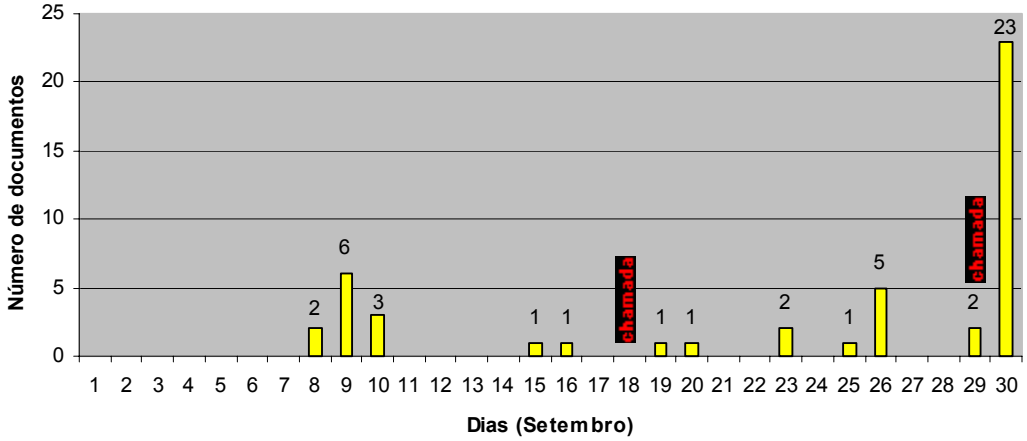

FIGURA 2 – Número de documentos submetidos ao RepositóriUM depois da reunião com os directores das comunidades piloto da UMinho, em 18 de Setembro e o aviso na *homepage* em 29 de Setembro de 2003.

TABELA 1 – Número Total de documentos submetidos ao RepositóriUM até a sua abertura ao público.

Mês	Número de documentos	%
Julho	46	16,43
Agosto	2	0,71
Setembro	48	17,14
Outubro	69	24,64
Novembro	115	41,08
Total	**280**	**100**

Do número total de trabalhos incluídos no RepositóriUM, uma maior percentagem foi observada entre Outubro e Novembro, com um aumento substancial em Novembro.

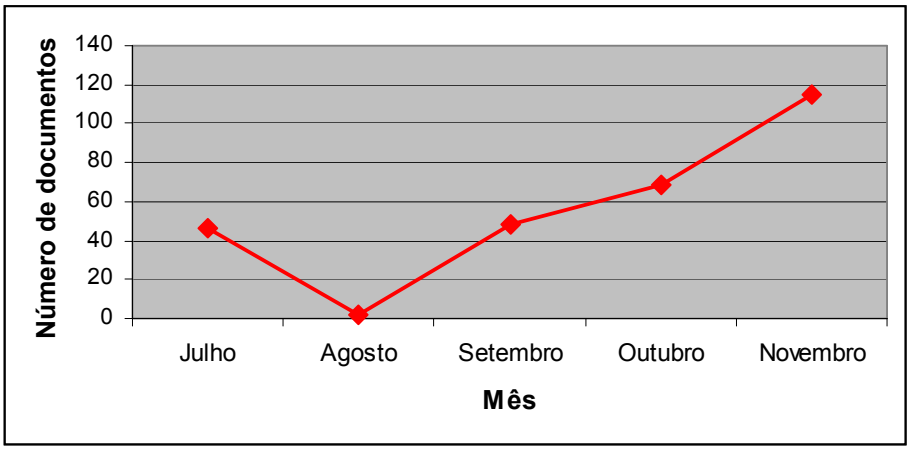

FIGURA 3 – Curva da tendência de crescimento em relação à inclusão de documentos no RepositóriUM.

Em Setembro de 2003, simultaneamente com o carregamento de teses e dissertações as seguintes seis comunidades foram convidadas a utilizar o RepositóriUM como comunidades piloto: Sistemas de Informação; Engenharia de Polímeros; Engenharia Biológica; Gestão, e outras duas, uma da área de Ciências Humanas e outra da área de Ciências Sociais. Essas comunidades foram escolhidas por uma ou várias das seguintes razões: (a) intensa actividade de investigação (verificado pelo número de publicações e excelente avaliação externa); (b) desejo de ter comunidades de diferentes áreas, localização geográfica e afiliação institucional (diferentes Escolas/Institutos); (c) contactos pessoais nessas comunidades que as tornariam, em princípio, mais activas na adopção da tecnologia.

Como o processo de depósito (*workflow*) de teses e dissertações foi diferente para as colecções e ocorreram várias submissões de autores que não eram afiliados às comunidades piloto, foi criada a comunidade da biblioteca da UMinho, para armazenar todas as teses e dissertações.

As primeiras quatro comunidades responderam adequadamente ao desafio, enquanto as outras duas não o fizeram declaradamente, mostrando a sua indisponibilidade em participar naquele momento. Embora não tenha sido possível realizar um estudo detalhado, podemos apontar algumas razões que acreditamos tenham contribuído para o facto de estas duas comunidades não se terem envolvido:

- O repositório foi disponibilizado como um projecto da UMinho, não resultando de uma necessidade expressa ou não expressa dos investigadores. Neste caso, há a necessidade do líder do projecto se esforçar consideravelmente para defender a ideia.

- Os investigadores não estão cientes do grande potencial da tecnologia que lhes foi disponibilizada. Não foram facilmente compreendidos os benefícios em armazenar no repositório os trabalhos já publicados em periódicos bem quotados.

- Os investigadores daqueles departamentos estão normalmente sujeitos a uma sobrecarga de trabalho, tentando conciliar investigação, ensino e responsabilidades de gestão. Eles encararam a tarefa de reunir publicações anteriores e enviá-las para o repositório como uma carga suplementar de trabalho.

- Algumas das comunidades não se sentem confortáveis e não estavam habituadas a partilhar os resultados da sua investigação mais recente. O esforço parece ser maior em partilhar conhecimento com colegas do que torná-lo visível internacionalmente.

- Investigadores menos confortáveis em utilizar a ampla quantidade de tecnologias de informação, agora disponíveis, sentiram grande dificuldade em utilizar a interface do DSpace. Alguns investigadores perceberam a perda de tempo considerável para carregar apenas um artigo e desistiram da tarefa.

Os recursos disponibilizados no RepositóriUM tiveram uma grande variabilidade de acordo com as comunidades científicas. É possível verificar na Tabela 2 que as comunidades que apresentaram maior resposta são as de Polímeros e Sistemas de Informação.
Neste caso não consideramos o número de documentos da comunidade SDUM, uma vez que ele reúne teses e dissertações de todas as comunidades da UMinho, incluindo as envolvidas neste estudo.

TABELA 2 – Distribuição de documentos por comunidades.

Comunidades	Número de documentos até Novembro/2003	Número de documentos até Abril/2004
Engenharia Biológica	22	22
Engenharia de Polímeros	115	115
Sistemas de Informação	37	63
Gestão	15	15
TOTAL	**189**	**215**

A tabela mostra o número de documentos enviados ao RepositóriUM até Novembro/2003 e Abril/2004. Estes valores indicam que três das quatro comunidades não sofreram alteração enquanto uma delas aumentou significativamente a quantidade de documentos enviados. Não deixa de ser interessante o facto de apenas uma destas comunidades ter continuado a arquivar os seus documentos depois da fase intensa de carregamento inicial.

PRINCIPAIS PROBLEMAS ENCONTRADOS

Tal como em outros repositórios institucionais, o principal desafio foi alimentar o sistema. A resposta aos convites para teses e dissertações foi limitada: 111 documentos, a maioria deles dos últimos 3 anos (2000 a 2003). Em relação aos conteúdos das comunidades piloto a maioria deles foi recolhida por um ou vários representantes de cada comunidade, e submetidas em lote.
O número de documentos submetidos por auto-arquivo, antes e depois da abertura formal do RepositóriUM é muito pequeno (menos de 10 documentos). Existem várias explicações para este facto, mas a maioria delas está relacionada com a questão das instituições não estarem

habituadas a fazer o auto-arquivo dos seus próprios conteúdos em repositórios institucionais ou outros serviços externos.
O auto-arquivo é uma tarefa nova, muitas vezes apreendida como uma carga adicional em agendas demasiado cheias e que ainda não faz parte dos hábitos e rotinas dos investigadores e membros do corpo académico. Adicionalmente, muitos dos investigadores não se lembram (ou desconhecem) das licenças de direitos de autor que assinaram quando publicaram em periódicos, o que é um problema importante neste contexto. Para o auto-arquivo, os autores devem conhecer os seus direitos: o que podem arquivar (*preprints, postprints*, etc.) e quais os direitos de acesso que pretendem atribuir ao documento (acesso mundial, restrito à sua organização, etc.).

LINHAS DE INVESTIGAÇÃO

A criação do repositório institucional da UMinho e o progresso do projecto descrito nas secções anteriores têm dado suporte à identificação de algumas linhas de investigação que começam a ser desenvolvidas no contexto de trabalhos de mestrado ou doutoramento na UMinho. Em particular, quatro investigadores estão activamente a trabalhar com o intuito de contribuir para esclarecer parte dos seguintes aspectos:

1. Como reagem os investigadores (em ambos os papéis de produtor e consumidor da informação) a essa tecnologia? O que está envolvido nesta reacção? O que a motiva?

2. Como se faz a alteração das reacções de resistência (e como essa alteração é expressa) nas diferentes culturas e áreas disciplinares?

3. Quais são os factores importantes para que a tecnologia possa ser melhor aceite? O que pode motivar a sua utilização pelos investigadores? E como podem eles ser motivados? Que iniciativas de gestão devem ser implementadas para promover a melhor comunicação e partilha dos resultados de investigação por meio do uso intensivo do repositório institucional da UMinho?

4. Qual a aplicabilidade do DSpace a outros tipos de materiais? O que deve ser mantido e o que deve ser mudado? E com outros utilizadores?

5. Que outras funcionalidades podem ser adicionadas ao DSpace a fim de promover o melhor relacionamento entre investigadores?

Apesar do uso do DSpace ainda estar no seu início, já se podem retirar algumas conclusões. O projecto começou como uma iniciativa dos Serviços de Documentação da Universidade do Minho. Essa iniciativa resultou do entendimento da importância do RepositóriUM em disseminar a investigação da Universidade, tanto internamente como externamente. Enquanto muitos dos departamentos envolvidos entendem essa importância, muitos investigadores expressaram menos entusiasmo por razões apresentadas na secção 3. As poucas iniciativas em divulgar e afirmar o repositório levaram ao aumento no envio de artigos científicos do Departamento de Sistemas de Informação (Tabela 2). Essa experiência restrita motiva-nos a planear iniciativas semelhantes noutros departamentos.
Além disso, há alguma evidência de que a disseminação das estatísticas de acesso ao sistema, mostrando que os artigos armazenados estão a ser acedidos nacional e internacionalmente, pode levar a um aumento do interesse dos investigadores pelo repositório. No entanto, isto

pode não ser verdade para todas as comunidades científicas da Universidade. Naquelas áreas científicas onde há uma cultura de partilha de investigação num estágio inicial, os investigadores provavelmente sentir-se-ão motivados a enviar os seus artigos mais recentes. Nas áreas científicas onde essa cultura não existe, o acesso fácil fornecido pelo repositório pode ser um factor chave a sustentar a resistência à utilização do repositório.
A Tabela 2 indica que os departamentos de Engenharia Biológica e Gestão apresentam uma fraca adesão ao repositório. Uma das razões pode ser o facto de que os investigadores nesses departamentos têm menos experiência em utilizar regularmente sistemas de informação. Esse facto pode resultar numa maior tensão para realizar as tarefas requeridas no envio de artigos.

CONCLUSÕES

A Universidade do Minho foi a primeira instituição da comunidade mundial de língua portuguesa (de aproximadamente 200 milhões de pessoas) e uma das primeiras na Europa a traduzir e implementar o DSpace.
A primeira fase do plano de implementação do RepositóriUM (a nossa instanciação do DSpace) foi dividida em cinco passos principais, iniciando em Maio de 2003 com a tradução do sistema. O último passo consistiu na abertura a toda comunidade da UMinho, realizada em Novembro de 2003.
Durante estes meses de actividade (de Maio/2003 até Abril/2004) encontramos muitos problemas e adquirimos experiências que podem ser partilhadas com outros cientistas.
A experiência do projecto dá suporte a toda evidência empírica obtida mundialmente sobre sistemas de comunicação e colaborativos. Colaboração, comunicação e partilha de conhecimento não são implementadas apenas através da implementação de sistemas de informação que suportam esse tipo de actividades. Existe a necessidade dos gestores criarem e alimentarem uma cultura de colaboração, comunicação e partilha utilizando instrumentos de gestão. Alterações culturais não são fáceis de realizar e podem ter resultados inesperados, uma vez que se baseiam na alteração de percepções, sentimentos, e entendimentos profundamente enraizados numa comunidade.
Será necessário realizar estudos mais profundos e bem direccionados para tentar encontrar soluções e propor caminhos aos responsáveis pela implementação de repositórios institucionais, especialmente no que diz respeito a aspectos sociais, políticos e comportamentais, nem sempre tomados em consideração pelos especialistas informáticos.
Alguns desses estudos já estão a ser desenvolvidos no departamento de Sistemas de Informação da UMinho.

TRABALHO FUTURO

Existem dois tipos de trabalhos futuros planeados: de desenvolvimento e de investigação. Os esforços de desenvolvimento serão realizados principalmente pelos SDUM e a investigação será realizada principalmente pelos departamentos científicos da UMinho, em particular no Departamento de Sistemas de Informação.
No segundo semestre de 2004, o desenvolvimento do RepositóriUM será encaminhado em várias direcções. A partir de uma experiência anterior, o vector estratégico mais importante é promover a iniciativa de depósito pelo autor, através do auto-arquivo ou do envio mediado de conteúdos. Nesse sentido, os SDUM irão:

- Melhorar a usabilidade da interface de depósito e a informação de suporte (guias para o depósito de diferentes tipos de documentos, exemplos e ajudas contextuais) para o processo de depósito;

- Desenvolver um serviço de ajuda e esclarecimento sobre direitos de autor (*copyright*) integrado no processo de depósito;

- Integrar um passo de licença para o utilizador final (utilizando licenças padrão da *Creative Commons- http://creativecommons.org/*) no processo de depósito;

- Integrar o depósito de teses e dissertações submetidas e aprovadas no RepositóriUM com processos administrativos/académicos e sistemas de informação da Universidade do Minho;

Um outro aspecto estratégico importante é a demonstração do valor e vantagens do RepositóriUM para os investigadores e para as unidades organizacionais onde eles trabalham. As comunidades que já utilizam o sistema estão a ser consultadas em relação aos serviços de valor acrescentado e funcionalidades que elas consideram que será importante incluir no RepositóriUM. A lista final de desenvolvimentos ainda está a ser definida, mas certamente incluirá:

- Estatísticas e relatórios de acesso e *download* de artigos, a nível individual e de comunidade;

- Listas de publicações por autor e unidades orgânicas (criação de páginas *Web* de referências, criação de uma secção para publicação de relatórios periódicos de avaliação produzidos pelos investigadores, etc.);

- Serviços de análise de citação de artigos incluídos no RepositóriUM;

Considerando as actividades de investigação, uma parte do trabalho planeado já está a ser realizado, enquanto outra parte está planeada para ser iniciada brevemente.
A utilização do RepositóriUM está a ser estudada por uma aluna de doutoramento que está a procurar compreender e comparar as atitudes e comportamentos de investigadores de diferentes áreas do conhecimento da Universidade do Minho e de diferentes culturas em relação ao uso de repositórios. Atitudes e comportamentos serão avaliados utilizando-se uma ferramenta de medida em desenvolvimento. Em paralelo, um estudo qualitativo será realizado para identificar as razões que fundamentam os padrões de comportamento identificados pela avaliação.
O Departamento de Sistemas de Informação também participa no desenho de um projecto em I&D proposto para financiamento no quadro do programa IST da União Europeia. Neste projecto, a análise descrita no parágrafo anterior será estendida a outros países europeus. Dessa forma, poder-se-á comparar os padrões de comportamento das mesmas comunidades científicas em diferentes culturas institucionais de diferentes países. Esperamos que a extensão desse estudo ajude a identificar uma lista de factores de sucesso e barreiras à implementação de repositórios institucionais, bem como uma lista de estratégias e alterações que poderão ser adoptadas pelas comunidades académicas para o sucesso da implementação dos repositórios.
Um aluno de mestrado está a desenvolver trabalho com intuito de verificar que retorno deve ser dado a quem utiliza o sistema e que estatísticas devem ser revistas para motivar os utilizadores a incluírem os seus documentos no sistema. Este aluno irá desenvolver e integrar o código necessário para criar estas estatísticas e apresentá-las ao utilizador final de uma forma amigável.

Um outro aluno de doutoramento tem estudado os grupos de elementos de metadados para ambientes de ensino à distância. Uma instanciação do DSpace (Papadocs – http://papadocs.dsi.uminho.pt) foi instalada para armazenamento e disponibilização de trabalhos dos alunos. Foi necessário realizar algumas adaptações ao código do RepositóriUM. Outras adaptações e integração com ambientes de ensino à distância serão estudadas num futuro próximo.
Um outro investigador encontra-se a trabalhar com mecanismos de comunicação informal que serão incluídos no Papadocs e no RepositóriUM. Alguns desenvolvimentos têm sido realizados e estão agora a ser testados. A informação para a comunidade DSpace será divulgada assim que se obtiver uma versão estável.

REFERÊNCIAS

(1) DSpace Federation - DSpace at MIT. (2003). Retrieved 28/01/2004, 2004, from http://DSpace.org/introduction/index.html.
(2) Rodrigues, E.*;* Almeida, M.*;* Miranda, A.*;* Guimarães, A.*;* Castro, D. RepositóriUM : criação e desenvolvimento do Repositório Institucional da Universidade do Minho. In CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 8, Estoril, 2004 - "Nas encruzilhadas da informação e da cultura : (re)inventar a profissão": actas. [CD-ROM]. Lisboa: Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas, 2004. ISBN 972-9067-36-8.
(3) Crow, R. A guide to Institutional Repository Software. 1a. edição. October 2003. Disponível em http://www.soros.org/openaccess/software. Acesso em 23 out. 2003.

---

[i] "Open Source" no original, em Inglês.