

# RepositóriUM: criação e desenvolvimento do Repositório Institucional da Universidade do Minho

*Eloy Rodrigues, Matilde Almeida, Ângelo Miranda, Augusta Guimarães, Daniela Castro*

Serviços de Documentação da Universidade do Minho

Campus de Gualtar

4710-057, Braga

Tel: 253604150

E-mail: {eloy, matilde, amiranda, augusta, dcastro}@sdum.uminho.pt

## RESUMO

Nos últimos anos, os repositórios institucionais têm sido alvo de grande atenção por parte de universidades e bibliotecas universitárias. Os repositórios institucionais inserem-se no movimento vulgarmente conhecido por *Open Access Initiative*, que visa promover o acesso livre e irrestrito à literatura científica e académica, favorecendo o aumento do impacto do trabalho desenvolvido pelos investigadores e instituições, e contribuindo para a reforma do sistema de comunicação científica, reassumindo o controlo académico sobre a publicação, aumentando a competição e reduzindo o monopólio das revistas científicas das editoras comerciais.

Neste contexto, os Serviços de Documentação da Universidade do Minho definiram como objectivo estratégico para 2003 a constituição de um Repositório Institucional – o RepositóriUM, com o propósito de armazenar, preservar e divulgar a produção intelectual das diversas comunidades científicas desta Universidade. O presente trabalho descreve a forma como o RepositóriUM foi implementado na plataforma *Dspace*, as funcionalidades do sistema, a estratégia de divulgação junto das comunidades científicas da Universidade do Minho, o processo de depósito ou auto-“arquivo” dos documentos e a aplicação do esquema de metadados Dublin Core na descrição dos documentos digitais depositados.

Conclui-se apresentando os resultados e os ensinamentos obtidos neste primeiro ano do projecto, bem como as principais linhas que se irão prosseguir no desenvolvimento do RepositóriUM.

**PALAVRAS-CHAVE:** repositórios institucionais - comunicação científica - Open Access Initiative -Dspace

## INTRODUÇÃO

Nos primeiros meses de 2003, no quadro da elaboração do plano de actividades dos Serviços de Documentação da Universidade do Minho (SDUM) para esse ano, e da preparação da candidatura da Universidade ao Programa E-U Campus Virtual, foi decidido proceder à criação de

um repositório institucional. Deste modo, a Universidade do Minho juntou-se às centenas de Universidades em todo o mundo que, em especial nos últimos dois anos, têm promovido o aparecimento destes sistemas de informação.

O que são os repositórios institucionais [1]? Basicamente, são colecções digitais que armazenam, preservam, divulgam e dão acesso à produção intelectual de comunidades universitárias. Ao fazê-lo, os repositórios pretendem intervir e dar resposta a duas questões estratégicas que as universidades enfrentam:

- contribuir para o aumento da visibilidade, estatuto, imagem e “valor” público da instituição, servindo como indicador tangível da qualidade dessa universidade e demonstrando a relevância científica, económica e social das suas actividades de investigação e ensino;
- contribuir para a reforma do sistema de comunicação científica, expandindo o acesso aos resultados da investigação, reassumindo o controlo académico sobre a publicação científica, aumentando a competição e reduzindo o monopólio das revistas científicas, o que se pode traduzir também em economias para as universidades e as bibliotecas que as servem.

No que diz respeito ao primeiro aspecto, os repositórios institucionais servem não apenas para preservar a produção intelectual de uma dada comunidade académica, mas sobretudo para difundir e dar projecção a essa produção junto de outras comunidades universitárias e científicas, e da sociedade em geral. No actual momento, a produção intelectual de uma determinada comunidade universitária encontra-se dispersa por milhares de revistas científicas e actas de conferências. Se é verdade que a publicação nessas revistas científicas (em particular as com maior factor de impacto) se reflecte positivamente no prestígio da Universidade e na sua capacidade para atrair recursos financeiros, a constituição de um repositório institucional, ao concentrar a produção dos investigadores dessa Universidade, torna-a mais visível e facilita a demonstração do seu valor científico, cultural, social e económico.

Quanto ao segundo aspecto, os repositórios

institucionais inserem-se num movimento mais amplo de acesso livre ao conhecimento científico e académico. De acordo com a Declaração de Budapeste [2], que impulsionou este movimento vulgarmente conhecido por Open Access Initiative (BOAI) e que aqui designaremos por Acesso Livre [3], o encontro entre uma velha prática (a disponibilidade dos investigadores para publicarem os resultados do seu trabalho – em revistas científicas ou actas de congressos – sem esperarem qualquer pagamento) e novas tecnologias (a Internet e as ferramentas e protocolos a ela associados) pode promover um inestimável bem: o acesso livre e irrestrito à literatura científica por parte dos cientistas e académicos, professores, estudantes e público em geral.

São geralmente consideradas duas vias (paralelas e não antagónicas) para o acesso livre ao conhecimento:

- revistas com acesso livre, onde os artigos ficam disponíveis sem restrições desde a sua publicação;
- o auto-arquivo pelos autores dos seus trabalhos em repositórios institucionais livremente acessíveis;

Assim, ao encorajar os seus investigadores e bolseiros a depositar os seus trabalhos no repositório institucional, a Universidade do Minho pretende promover o Acesso Livre, a distribuição irrestrita, a interoperabilidade e a preservação a longo prazo. Pretende também reunir num único sítio o conjunto das publicações científicas da U.M., contribuindo desse modo para o aumento da sua visibilidade e impacto e garantindo a preservação da memória da Universidade.

Aumentar a visibilidade e o impacto da investigação desenvolvida é do interesse óbvio da Universidade, das suas unidades orgânicas (centros de investigação), bem como dos docentes e investigadores individualmente. E hoje há cada vez mais evidência que os artigos livremente disponíveis na Internet têm maior impacto do que os outros [4].

Para além do aumento do impacto da produção intelectual, promovendo a sua acessibilidade, a constituição do RepositóriUM poderá também revelar outras vantagens para os investigadores e a instituição. Por exemplo, o repositório institucional poderá ser útil nos processos de avaliação das unidades de investigação, assegurando que a produção científica, bem como dados bibliográficos e bibliométricos, está facilmente disponível. E a partir do repositório os investigadores poderão obter ou gerar relatórios de actividade, estatísticas de acesso aos seus documentos, criação de listas de publicações, etc.

Refira-se, finalmente, que uma das motivações para a criação do repositório institucional da U.M. foi também

a procura de uma solução para o armazenamento, disponibilização e preservação das teses e dissertações aprovadas na Universidade do Minho, em formato digital.

## **A PLATAFORMA PARA A CRIAÇÃO DO REPOSITÓRIUM: O DSPACE**

Após a decisão da criação do RepositóriUM, foi realizado um estudo prévio das plataformas então existentes para a constituição e gestão de repositórios institucionais. Em resultado da análise das suas funcionalidades, bem como dos objectivos definidos na Universidade, foi seleccionada a plataforma DSpace, para a implementação do repositório institucional da U.M.

O DSpace é um projecto das bibliotecas do Massachusetts Institute of Technology (MIT) para recolher, preservar, gerir e disseminar a produção intelectual dos seus investigadores. Ele é o resultado de um esforço conjunto de investigação e desenvolvimento do MIT e da Hewlett-Packard (HP). O sistema foi disponibilizado publicamente em Novembro de 2002 de acordo com os termos da *BSD open source* license [5]. Está actualmente em funcionamento no MIT e em diversas outras universidades dos Estados Unidos e da Europa.

Esta plataforma de *software open source* possibilita às instituições:

- recolher e descrever trabalhos em formato digital utilizando um processo de depósito de documentos que pode ser constituído por vários passos;
- disponibilizar estes trabalhos através da *web* através de um sistema de pesquisa e recuperação da informação;
- preservá-los a longo prazo.

O DSpace é constituído por um conjunto de ferramentas para administrar e disseminar conteúdos digitais de acordo com o *OAIS reference model* [6].

### **Organização**

A forma como os dados estão organizados pretende reflectir a estrutura da instituição que utiliza o sistema. Como podemos ver na figura 1, o DSpace está orientado às comunidades. Estas correspondem tipicamente a laboratórios, centros de investigação, departamentos ou outro tipo de unidades orgânicas que constituem a instituição.

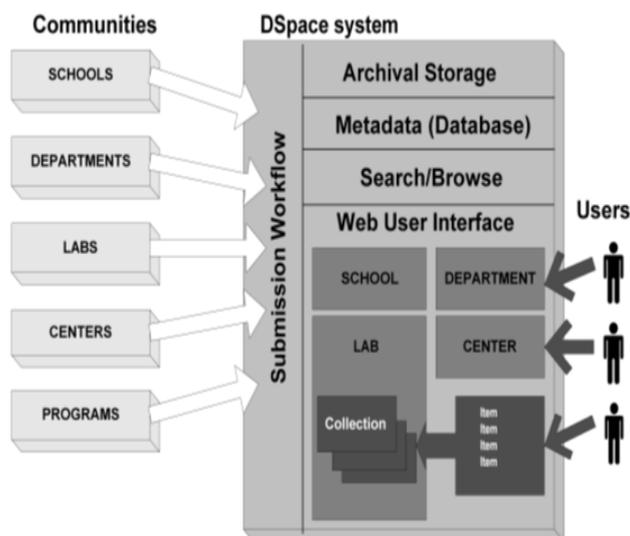


Figura 1: Modelo de Informação do DSpace

### Metadados

Para a descrição dos documentos, o DSpace utiliza a recomendação de metadados *Qualified Dublin Core* baseado no formato de registo sugerido pela *Libraries Working Group Application Profile*. Apenas três campos são obrigatórios: título, idioma, e data de depósito, todos os outros campos são opcionais. Entre outros, existem campos adicionais para resumos dos documentos, palavras-chave, metadados técnicos e metadados para direitos de autor.

Além deste esquema de metadados descritivo, o sistema armazena também outros tipos de metadados, como por exemplo, tamanho em *bytes* dos ficheiros, *MIME types*, metadados específicos para preservação digital, etc.

### Processo de Depósito e Workflow

O DSpace é o primeiro repositório digital *open source* a apresentar uma solução para resolver o problema complexo de como integrar e implementar os diferentes processos de depósito necessários a um sistema multidisciplinar. As diversas comunidades, representando escolas, departamentos, centros e laboratórios de investigação, têm diferentes ideias de como, por quem, e com que restrições os materiais devem ser depositados no repositório. Quem deve ter permissão para depositar? Que tipo de documentos podem ser depositados? Quem pode rever ou aprovar o depósito? Para que colecções uma pessoa pode depositar materiais? Quem pode ver os documentos no repositório? Todas estas questões são respondidas pelos responsáveis da cada comunidade, e posteriormente definidas e implementadas no sistema. Isto é feito à custa da definição de papéis (ex.: *submitters*, *reviewers*, *metadata editors*) com determinadas permissões que são atribuídos aos utilizadores do sistema. Estas permissões podem ser definidas individualmente para comunidades, colecções e documentos. O sistema inclui também um mecanismo para importar os documentos em lote.

### Interoperabilidade

Para concretizar o objectivo da interoperabilidade com outros sistemas, o DSpace implementa o protocolo *Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)*.

O OAI-PMH é um protocolo que permite aos fornecedores de informação disponibilizar e expor pela Internet os metadados de cada um dos seus recursos. Desta forma os metadados ficam disponíveis para serem recolhidos por serviços especializados em indexação de recursos científicos e passam a constituir as bases de dados desse tipo de serviços.

### Identificadores Persistentes

Um dos objectivos e requisitos dos repositórios digitais é referenciar de uma forma persistente os seus recursos de forma a permitir a pesquisa e recuperação dos mesmos num futuro distante. Em particular, é crucial que citações para materiais arquivados permanecem válidas por longos períodos de tempo. Os investigadores necessitam de referências estáveis para os seus trabalhos. Os utilizadores da Internet aprenderam rapidamente que os *sites* (referenciados pelas suas URLs) podem desaparecer ou serem reconfigurados sem qualquer tipo de aviso, o que faz com que as suas referências para recursos digitais de investigação não sejam fiáveis a longo prazo.

Para ajudar a resolver este problema, uma das funcionalidades do DSpace é a criação de identificadores persistentes para cada item, colecção e comunidade armazenada no sistema. Para isso é necessário implementar um mecanismo independente do local físico de armazenamento da informação. O sistema usa o CNRI *Handle System* [7] para criar estes identificadores.

### IMPLEMENTAÇÃO DO REPOSITÓRIUM

A implementação do RepositóriUM, através do sistema DSpace, iniciou-se em Maio de 2003. O plano de implementação consistiu em 4 fases principais:

#### Instalação, Configuração e Tradução

Inicialmente procedeu-se à instalação da infra-estrutura física e lógica (sistema operativo, servidor web, base de dados, etc.) necessárias ao pacote original do sistema, tal como ele é distribuído pelo MIT. Após instalação do sistema DSpace, e efectuada uma análise mais aprofundada de todas as suas funcionalidades, foram iniciadas as tarefas de configuração, personalização e tradução do interface gráfico para a língua portuguesa. Esta fase decorreu em Maio e Junho de 2003.

#### Carregamento de Teses e Dissertações

Como foi referido, os problemas relacionados com as teses e dissertações aprovadas pela Universidade do Minho foram uma das motivações para a constituição do RepositóriUM. Assim, esta foi definida como a área prioritária para o início do carregamento de conteúdos no sistema.

De Julho a Setembro de 2003 foram efectuados vários

apelos, através da lista de distribuição UM-Net, aos doutorados e mestres pela Universidade do Minho para entregarem as suas teses e dissertações. Pelo facto de só se ter divulgado o apelo através da lista de distribuição da U.M., apenas os mestres e doutorados que continuam ligados à Universidade (como docentes e investigadores) foram contactados. Foram disponibilizadas várias vias para que os autores entregassem os seus documentos: Correio electrónico, FTP, entrega directa. Foram também disponibilizados serviços de digitalização para aqueles que não tivessem os seus documentos em formato digital.

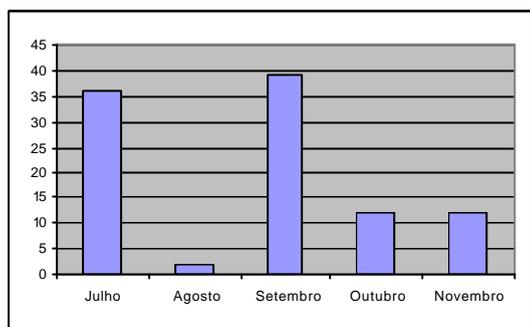


Figura 2: Distribuição Mensal das T&D

As figuras 2 e 3 ilustram o ritmo e a via de entrega das teses e dissertações para o RepositóriUM.

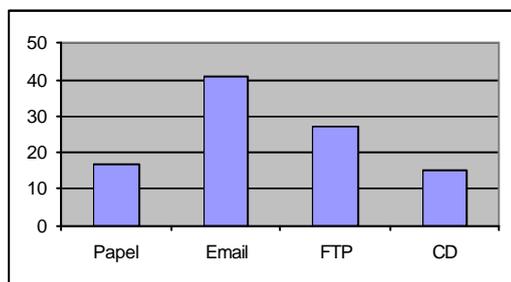


Figura 3: Por via utilizada

#### Constituição de Comunidades Piloto

A terceira fase do processo, foi a constituição de comunidades piloto, no sentido de testar a utilização do sistema com outros tipos de documentos e com utilizadores externos aos Serviços de Documentação. Com base em diversos critérios (desde a diversidade de áreas científicas, localização geográfica e tipo de unidade orgânica, até à relevância da sua produção científica) foram identificadas 6 unidades orgânicas, a quem foram endereçados convites para serem comunidades piloto no RepositóriUM. Das 6 unidades contactadas 4 (Centro de Engenharia Biológica, Departamento de Engenharia de Polímeros, Departamento de Sistemas de Informação e Núcleo de Estudos de Economia e Gestão) aceitaram o convite. Assim, foram constituídas e configuradas no RepositóriUM essas 4 comunidades piloto.

Cada comunidade definiu as suas colecções, os

processos de depósito e aceitação e políticas de acesso. Apesar de se terem realizado alguns depósitos directos de documentos pelos seus autores (ou seja, *auto-arquivo*), nesta fase a esmagadora maioria dos documentos foram carregados em "lote".

Como se pode constatar na figura 4, mais de 1/3 dos documentos inseridos foram produzidos por uma das comunidades. Refira-se que a comunidade "SDUM" foi constituída para reunir o conjunto das teses de mestrado e doutoramento depositadas no RepositóriUM.

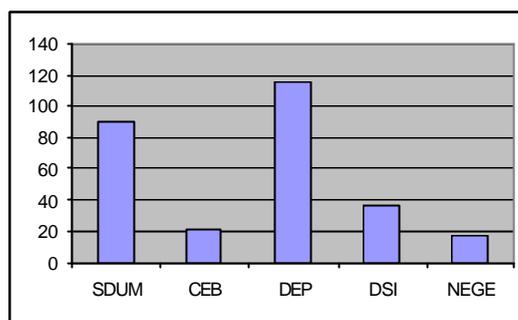


Figura 4: Documentos por Comunidade

Quanto ao tipo de documentos depositados nesta fase, a esmagadora maioria (67,5%) foram artigos e comunicações, tendo as teses de doutoramento (17,9%) e as dissertações de mestrado (14,6%) um peso semelhante (ver figura 5).

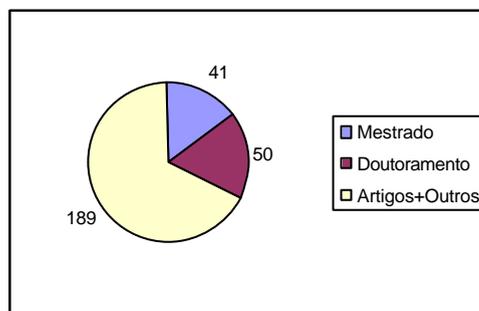


Figura 5: Documentos por Tipo

#### Abertura ao Público

A última fase da implementação foi a sua abertura oficial. No dia 20 de Novembro de 2003, através de uma pequena cerimónia presidida pelo Reitor da Universidade do Minho, o RepositóriUM foi disponibilizado publicamente, ficando acessível para toda a Universidade do Minho e para o público em geral. O RepositóriUM abriu com um total de 280 documentos.

## DESAFIOS, PROBLEMAS E SOLUÇÕES NA CONSTITUIÇÃO DO REPOSITÓRIUM

As iniciativas relacionadas com o movimento OAI têm vindo a ter um impacto significativo nos sistemas de informação em Ciência e Tecnologia (C&T), reflectindo-se quer no sector editorial e nas comunidades científicas, como no domínio das bibliotecas universitárias. Isto mesmo pôde ser constatado na nossa experiência.

Do ponto de vista do tratamento técnico documental, o processo de “auto-arquivo” de documentos em formato electrónico confrontou-nos com uma nova abordagem ao tradicional circuito do documento.

A equipa dos SDUM procurou ajustar-se a esta nova realidade emergente, promovendo por um lado a rápida disseminação dos resultados científicos gerados na Universidade e a visibilidade dos seus investigadores, e garantindo, por outro, a qualidade e consistência dos metadados, de modo a conferir, dentro do possível, a máxima precisão à pesquisa.

### O esquema de metadados Dublin Core

Como já foi referido, o Dublin Core (DC) foi o esquema de metadados adoptado na implementação do RepositóriUM.

O Dublin Core (DC) constitui um projecto de criação de uma estrutura para categorizar recursos electrónicos, tal como acontece, por exemplo, com os catálogos para os recursos “físicos”. Deste modo, os seus quinze elementos de descrição de documentos electrónicos - title, creator, subject, description, publisher, contributor, date, type, format, identifier, language, source, relation, coverage, rights - facultam a informação básica sobre os mesmos.

Nascido de um encontro realizado em Dublin (Ohio), em 1995, o DC foi concebido para descrever recursos da Internet e responder à necessidade de se criar uma nomenclatura comum de metadados.

As virtualidades deste esquema de metadados desde cedo despertaram a atenção de instituições tradicionalmente dedicadas ao tratamento de informação, tais como as bibliotecas. As vantagens do DC mais comumente apontadas são:

- a simplicidade,
- a interoperabilidade semântica,
- consenso internacional,
- a extensibilidade,
- a modularidade de metadados na web.

De facto, a flexibilidade e amplitude do esquema de metadados DC possibilitaram a sua utilização em larga escala, contribuindo simultaneamente para a sua rápida disseminação. No entanto, para que pudesse ser usado nas mais diversas situações, o DC acabou por definir de forma imprecisa, quer em termos semânticos quer de sintaxe, os seus elementos de descrição. Deste modo, o que constituía para o DC uma vantagem, permitindo a

sua fácil utilização, levou, por vezes, a interpretações distintas e a uma utilização descoordenada dos mesmos.

Por forma a colmatar estas situações e a dar maior precisão semântica aos seus elementos, o DC associou a quase todos eles qualificadores de refinamento de elemento e/ou de esquema de codificação. Os primeiros constituem termos associados ao elemento que especificam com maior precisão o seu significado. Por exemplo, o elemento *Date* tem como qualificadores de refinamento de elemento os termos *Issued*, *Created*, *Available*, *Modified* e *Valid*). Por seu lado, os qualificadores de esquema de codificação identificam esquemas que ajudam na interpretação do valor de um elemento. Por exemplo, o mesmo elemento *Date* tem como qualificadores de esquema de codificação os esquemas *W3C-DTF* e *DCMI Period*) (Fig. 6).

Registo Completo

Campo DC	Valor
contributor.author	Afonso, I. M.
contributor.author	Hes, Lubos
contributor.author	Maia, J. M.
contributor.author	Melo, L. F.
date.accessioned	2003-10-24T16:49:20Z
date.available	2003-10-24T16:49:20Z
date.issued	2003
identifier.citation	"Journal of Food Engineering". 57:2 (April 2003)
identifier.uri	http://hdl.handle.net/1822/92
description.abstract	In the present work an experimental investigation was carried out in order to obtain a correlation for the determination of convective coefficients of stirred yoghurt in a plate heat exchanger. The study was carried out in order to characterise the convective behaviour, evaluating its dependency both on temperature and on the material shear rate. A shift in the temperature dependence of the convective coefficient was observed at 25 °C. It is also shown that the material shear rate dependence of the convective coefficient changes from a Bingham fluid to a power law fluid in excess of approximately 6.7 Pa. A behaviour of the non-Newtonian stirred yoghurt in a plate heat exchanger is also described. The convective heat transfer coefficient was obtained from the thermal entry length due to the short length of the plate heat exchanger.
format.extent	254951 bytes
format.mimetype	application/pdf
language.iso	en
publisher	Elsevier
title	Heat transfer and rheology of stirred yoghurt in a plate heat exchanger
type	TEXT
type	Artigo

Aparece nas Coleções: [IPC - Artigos \(papers\)](#)

Ficheiro	Tamanho	Formato
<a href="#">04.pdf</a>	248KB	Adobe PDF

[Ver/Abrir](#)

Fig. 6 – Exemplo de registo Dublin Core no RepositóriUM: elementos e qualificadores

### O Circuito de depósito (Workflow)

O circuito de depósito suporta o auto-arquivo (*self archive*) do “documento” digital, por iniciativa do próprio autor (o depositante). O RepositóriUM tem um circuito de depósito (*workflow*) muito simples, dividido em 7 etapas (Fig.7), que conduzem o depositante ao longo de todo o processo de um modo bastante intuitivo e flexível. Em qualquer ponto do processo de depósito, o utilizador tem a possibilidade de o suspender e conservar o trabalho, com a garantia de que os dados introduzidos serão salvaguardados até que o processo

seja de novo retomado. O sistema permite exibir ao utilizador autenticado uma área pessoal, denominada "Meu RepositóriUM", onde encontrará uma lista dos depósitos em curso, a ligação para a lista dos documentos que já depositou e foram aceites no RepositóriUM, e uma lista de depósitos que aguardam alguma intervenção sua, como por exemplo revisão.



**Fig. 7 - Barra de progresso do processo de depósito.**

Para dar início ao processo de depósito é necessário, em primeiro lugar, seleccionar a comunidade e a colecção para a qual se vai depositar o item. Às comunidades é dada a liberdade de criarem um número ilimitado de colecções, que podem ser organizadas em torno de um tópico, por tipo de "documento" (tal como teses e dissertações, artigos, imagens, datasets, etc.) ou através de um sistema de classificação.

Nos parágrafos seguintes são descritas as etapas que constituem o processo de depósito.

1ª, 2ª e 3ª Etapas – Estas etapas dizem respeito à descrição do item. Na 1ª etapa o sistema coloca perguntas básicas sobre o documento, como por exemplo se é composto por múltiplos ficheiros. Em função das respostas são criados e apresentados dois formulários para a introdução de metadados. O primeiro formulário (2ª etapa) agrupa dados essenciais, como autor, título, data da publicação ou distribuição pública, editor, tipo de documento/informação, etc., e o segundo formulário (3ª etapa) agrupa dados como palavras-chave, o resumo, informação sobre patrocínios e notas adicionais.

Cabe aqui referir a importância da sensibilização dos utilizadores para a qualidade e profundidade dos metadados, para efeitos de difusão e recuperação das suas publicações. Igualmente importante foi a inclusão de ajudas online em contexto e de exemplos elucidativos de como preencher cada campo, orientando passo a passo um trabalho a que os autores não estão habituados.

4ª Etapa - Nesta etapa é apresentado ao utilizador o formulário de carregamento dos ficheiros. No caso do utilizador ter indicado no início do processo de depósito que o documento é composto por múltiplos ficheiros, por exemplo um documento que contenha ficheiros de texto, vídeo, e dados, é-lhe recomendado que preencha um

campo de "Descrição do Ficheiro". A informação introduzida nesse campo, embora parecendo trivial, serve para ajudar os utilizadores a perceberem que tipo de informação está armazenada em cada ficheiro. Se o item for composto por vários ficheiros de diferentes formatos a descrição poderá ser do tipo: Documento principal, Imagem, Programa de computador, Dados; se o documento for composto por vários ficheiros de texto a informação poderá ser do tipo: Capa, Páginas preliminares, Introdução, Capítulo, Conclusão, Bibliografia, etc.

5ª Etapa – Na etapa de verificação o utilizador tem a oportunidade de rever todos os dados introduzidos e efectuar eventuais correcções.

6ª Etapa – Etapa de concessão de licença de distribuição não-exclusiva à Universidade do Minho, para arquivar e tornar acessível o documento em suporte digital. Os itens

podem ser depositados no RepositóriUM de acordo com diferentes estatutos, estabelecidos em função de políticas mais ou menos restritivas e tendo em consideração condicionantes de *copyright*. Apesar da Universidade e os Serviços de Documentação apoiarem os princípios do Open Access, e estimularem a disponibilização da produção científica de forma irrestrita e imediata, o RepositóriUM implementa controlos de acesso, para limitar o acesso a documentos que exijam restrições, nomeadamente artigos publicados em revistas comerciais que o exijam.

7ª Etapa – A última etapa, que conclui o acto de depósito.

Após a conclusão do processo de depósito (auto-arquivo), existe um processo de aceitação que pode ter até três passos (Fig. 8). É dada à entidade responsável pela administração da colecção a possibilidade de rever o depósito, de acordo com a política ou exigências que definiu para cada colecção.

Os documentos depositados no RepositóriUM passam sempre por, pelo menos, um passo, que é o da validação dos metadados, que em todos os casos será sempre da responsabilidade dos SDUM.

Para além deste, os documentos podem passar ainda por um ou dois passos intermédios de revisão, de acordo com a decisão dos responsáveis das comunidades. Cada comunidade tem a possibilidade de definir políticas e fluxos de trabalhos diferentes para cada uma das suas colecções.

Aspectos como, quem pode depositar documentos, com que restrições, que tipo de documentos pode ser depositado, em que casos é necessário nomear um revisor, um grupo de revisores ou um coordenador, são algumas das questões que têm de ser definidas conjuntamente entre os representantes de cada comunidade científica e a equipe de trabalho dos SDUM, para ser levada a cabo a modelação do *workflow* de cada colecção. Como exemplo, podemos ter uma comunidade que decide que para a colecção *preprints* todos os membros têm permissões para submeter directamente, sem necessidade de passar por passos de revisão, e decide que para a colecção de documentos de trabalho (*working papers*) é necessário estabelecer um controlo de qualidade, definindo um ou dois passos intermédios de revisão e aprovação.

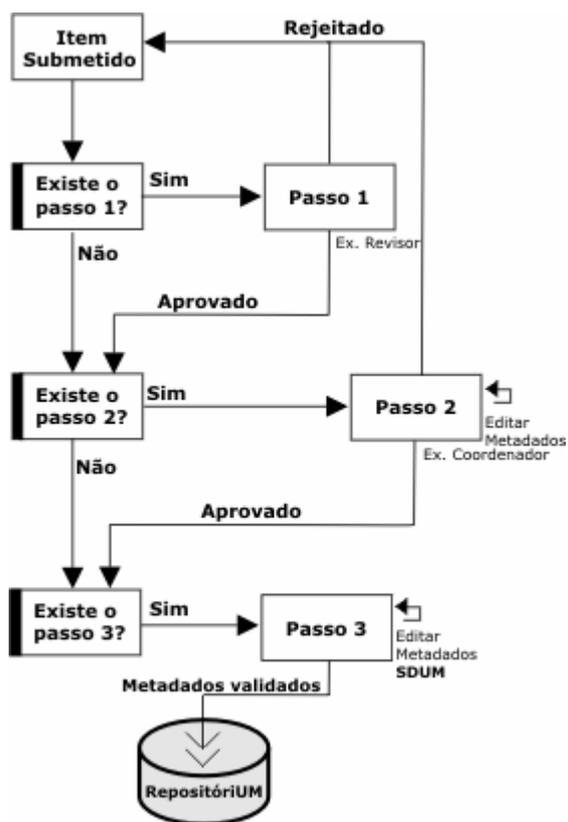


Fig. 8 – Diagrama processo de aceitação.

Os utilizadores a quem foram atribuídos papéis de revisão e aprovação em cada um dos passos recebem uma notificação via correio electrónico quando existe um novo documento para revisão/aprovação. O depositante é igualmente notificado, seja informando-o da existência de algum problema com o seu depósito detectado na revisão, seja informando-o de que o documento passou a integrar a colecção e já se encontra acessível no RepositóriUM.

#### Validação dos metadados: problemas e soluções

A validação dos metadados associados aos itens, passo obrigatório para a conclusão do processo de depósito e disponibilização dos documentos é sempre efectuada por técnicos superiores de biblioteca e documentação dos Serviços de Documentação da U.M.

Com a adopção do esquema de metadados Dublin Core cedo os responsáveis pelo processo de validação se viram confrontados com questões atidas ao controlo de autoridade e demais aspectos essenciais relacionados com o garante da integridade e consistência dos dados no novo sub-sistema de informação implementado.

O procedimento é iniciado com a edição do item depositado, procedendo-se à verificação/revisão dos metadados introduzidos. Concluída a verificação / revisão, o registo é considerado validado, sendo-lhe atribuído pelo sistema o respectivo URL.

No caso das teses e dissertações o processo fica

concluído com a intervenção do Sector de Catalogação, que efectua a verificação e actualização do catálogo bibliográfico da Universidade do Minho, preenchendo, no caso dos documentos já catalogados, o campo 856 relativo ao acesso electrónico, ou procedendo ao registo bibliográfico do novo documento.

Convém a propósito mencionar que o processo inicial de introdução dos metadados da colecção das teses e dissertações digitais foi efectuada pelos SDUM, o que implicou a criação de um Workflow específico.

Da mesma forma, e no sentido de facilitar o trabalho de depósito aos autores nesta fase de arranque do RI, foi acordado entre os SDUM e as comunidades a possibilidade de enviarem, em lote, parte das suas publicações, cabendo toda a tarefa de inserção dos metadados aos técnicos dos SDUM.

O tratamento dos documentos introduzidos no RepositóriUM suscitou algumas questões e problemas técnicos. Desde logo, a construção do registo de autoridade dos autores da U.M., que surgiam referenciados tanto pelo nome completo (no caso das teses de doutoramento, dissertações de mestrado ou provas de aptidão pedagógica) como pelo apelido e iniciais ( nos artigos, relatórios, working papers, etc.), apresentando neste último caso variações consideráveis. Tomemos como exemplo o registo de autoridade do autor referenciado na Fig. 9.

Etiquete	100	---	---	-----nem-----a-----
Dados Gerais	100	---	a	20001121--apor-0103---ba
Regros	152	---	a	RPC
Nome Pessoa	200	1	a	Amaral,
			b	Luis
Ver - Pessoa	400	1	a	Amaral,
			b	L.
Ver - Pessoa	400	1	a	Amaral,
			b	L. A.
Ver - Pessoa	400	1	a	Amaral,
			b	Luis A.
Ver - Pessoa	400	1	a	Amaral,
			b	Luis M.
Ver - Pessoa	400	1	a	Amaral,
			b	Luis A. M.
Ver - Pessoa	400	1	a	Amaral,
			b	Luis Alfredo Martins do
Fonte origem	801	0	a	PT
			b	SDUM
			c	20030101

Fig. 9 – Registo de Autoridade

A opção pareceu-nos desde logo clara, ou seja seguir os princípios estabelecidos nas RPC, optando-se pela forma de nome pelo qual o autor é mais conhecido e referenciado no meio académico (dado nem sempre fácil de determinar).

Outras questões foram surgindo, como as relativas à data de edição (no caso de teses o dissertações). Nem sempre a data apresentada na folha de rosto era a da defesa pública e, nesses casos, a opção foi a data da defesa. Os patrocinadores e respectiva descrição suscitaram igualmente algumas dúvidas, prontamente resolvidas pela equipa.

## CONCLUSÕES E TRABALHO FUTURO

A primeira etapa na constituição do RepositóriUM, que se concluiu no final de 2003, ao reunir um conjunto de resultados de investigação de quatro comunidades piloto, e ao recolher teses e dissertações defendidas no passado, disponibilizadas pelos próprios autores, teve como principal objectivo testar e experimentar o conceito dos repositórios institucionais (RI) no contexto particular da Universidade do Minho.

A estratégia adoptada foi recolher um conjunto significativo de recursos de diversas áreas que possibilitassem a criação de alguma “massa crítica” que permitisse a avaliação do objectivo proposto e a definição e planeamento das próximas fases do projecto. De resto, esta tem sido a estratégia adoptada por grande parte das instituições no lançamento dos seus RI's.

Neste sentido, tendo em conta o razoável envolvimento das comunidades piloto, o reconhecimento e aceitação por parte dos investigadores e docentes envolvidos dos seus benefícios e vantagens, os excelentes índices de utilização do sistema (ver Figura 10) e o seu impacto imediato na comunidade académica e público em geral, bem como o facto de o RepositóriUM constituir o primeiro sistema de informação do projecto E-UM a entrar em funcionamento, pode-se afirmar que, genericamente, o projecto é um sucesso e que o conceito de repositório institucional está na Universidade do Minho para durar.

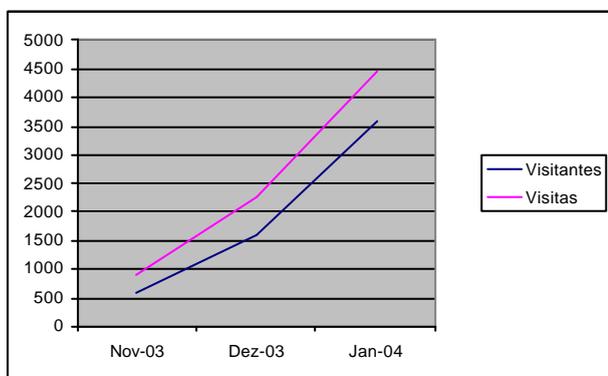


Fig. 10 – Acessos ao RepositóriUM

No entanto, estes primeiros meses revelaram também o aparecimento de diversos problemas, dificuldades e entraves. Problemas decorrentes tanto de questões sociais, culturais e legais bem como de questões de ordem mais operacional e prática. Pelo acompanhamento e participação activa que temos mantido das iniciativas do movimento de Acesso Livre, especialmente na vertente de implementação de RI's, estes problemas são comuns a todas as instituições e tem merecido a atenção das

comunidades envolvidas na promoção e desenvolvimento desta iniciativa.

Do conjunto de problemas e questões podem destacar-se:

- as relacionadas com a propriedade intelectual e direitos de autor (copyright) dos resultados da investigação e das publicações;
- a adopção, promoção e envolvimento dos vários interessados, como sejam, as próprias instituições, os produtores de investigação científica e as bibliotecas;
- o controlo de qualidade dos conteúdos dos repositórios e de que forma é que eles poderão acrescentar valor à investigação como um todo e à sociedade em geral;
- a forma como estes sistemas se devem enquadrar no sistema de comunicação científica e académica tradicional.

Uma das dificuldades, cuja superação é essencial para garantir o sucesso da constituição de um repositório institucional e que também se evidenciou na nossa experiência, é a obtenção do envolvimento adequado dos autores e produtores dos resultados de investigação. Os autores, quer como produtores quer como consumidores de informação, são os primeiros interessados num RI e quem poderá retirar benefícios mais imediatos.

Mas o que constatamos na prática foi que muitos não estão conscientes do grande potencial que estes sistemas podem representar. Como indivíduos, no contexto académico, as pessoas estão já sobrecarregadas com tarefas e podem olhar para esta iniciativa apenas como mais uma fonte de trabalho que lhes irá ocupar o seu precioso tempo. Por outro lado, muitos autores não tem consciência nem se apercebem no seu dia-a-dia, dos problemas e contradições do sistema de comunicação científica e académica tradicional, que estão na origem do aparecimento das iniciativas “Open Access”.

Por isso, uma boa estratégia de divulgação e promoção é um factor crítico para o sucesso na implementação de um RI, uma vez que são imprescindíveis para alterar alguns aspectos sociais e culturais que podem originar um fraco envolvimento por parte dos académicos.

Assim, nos primeiros meses de 2004 foram e continuarão a ser realizadas diversas acções (apresentações, reuniões, artigos, etc.) de promoção do RepositóriUM junto das unidades orgânicas da Universidade que ainda não o utilizam, e do auto-arquivo das publicações, junto do conjunto dos docentes e investigadores da Universidade do Minho. O objectivo é que no final de 2004 o RepositóriUM reúna, para além de documentos de anos anteriores, uma parte muito significativa da produção científica dos docentes e investigadores da Universidade do Minho no presente ano.

Para além das acções de promoção do auto-arquivo junto dos autores, tal como a experiência tem demonstrado em outras Universidades, será importante que progressivamente os departamentos/centros de

investigação e a própria Universidade adoptem formalmente políticas que tornem obrigatório o depósito da produção científica no RepositóriUM.

Por outro lado, é necessário ter em consideração as necessidades dos autores desenvolvendo serviços que facilitem o depósito dos seus trabalhos, e que recompensem os minutos adicionais que devem gastar com o auto-arquivo da sua produção científica. Nesse sentido, está já a ser concluído um serviço de apoio e esclarecimento das questões relacionadas com os direitos de autor (copyright) dos seus artigos.

Ao mesmo tempo, através da auscultação dos membros das comunidades que usam o RepositóriUM, serão definidos um conjunto de serviços de valor acrescentado para os autores (relatórios, estatísticas, geração de documentos, etc.) que irão ser desenvolvidos e disponibilizados até ao final de 2004.

Para além destas áreas consideradas prioritárias, prevê-se ainda para o presente ano:

- disponibilização de funcionalidades de pesquisa em outros repositórios e arquivos de livre acesso integradas na interface do RepositóriUM;

- instalação de uma nova versão base do sistema DSpace, que elimina algumas lacunas na versão actual;

Finalmente, consideramos que é também de grande importância a promoção do conceito dos repositórios institucionais e do movimento de OAI no conjunto da comunidade científica e académica portuguesa. Nesse sentido, para além de diversas outras acções, desenvolvemos no nosso Website uma área sobre as questões do Livre Acesso e repositórios institucionais, ([www.sdum.uminho.pt/site/acessolivre/](http://www.sdum.uminho.pt/site/acessolivre/)) onde iremos progressivamente reunindo informações sobre as questões estratégicas e políticas, normas, protocolos e ferramentas, exemplos de repositórios, etc.

[3] - De acordo com da Budapest Open Access Initiative a definição de Open Access é a seguinte: “(...)the free availability of peer-reviewed literature on the public internet, permitting any user to read, download, copy, distribute, print, search, or link to the full texts of the articles(...)” Budapest Open Access Initiative disponível em: <http://www.soros.org/openaccess/read.shtml>

[4] – Uma análise de 119. 924 artigos de conferência em informática e áreas relacionadas concluiu que o número médio de citações para artigos *offline* é de 2.74, enquanto a média de citações de artigos *online* é de 7.03, um aumento de 336% - Ver LAWRENCE, Steve - **Free online availability substantially increases a paper's impact**. *Nature*. vol. 411, nr. 6837 (2001), p. 521. Outros estudos, na área da física, confirmam esta tendência. Está a ser desenvolvido actualmente um estudo em larga escala sobre o impacto dos artigos disponíveis de acordo com o Open Access.

[5] - BSD open source license - Berkeley Standard Distribution License. Licença que define os termos da utilização e distribuição para software de domínio público. Aprovada pelo Open Source Initiative (OSI). Ver <http://www.opensource.org/licenses/bsd-license.php>

[6] - OAIS Reference Model - Reference Model for an Open Archival Information System, ver <http://www.ccsds.org>.

É uma *framework* conceptual para sistemas de arquivo dedicados à preservação e acesso no longo-prazo a recursos digitais.

[7] - CNRI Handle System - Ver <http://www.handle.net>

## NOTAS

[1] - Para a explicitação do conceito usamos o documento The Case for Institutional Repositories: A SPARC Position Paper, disponível em:

<http://www.arl.org/sparc/IR/ir.html>

[2] - Budapest Open Access Initiative disponível em:

<http://www.soros.org/openaccess/read.shtml>