



Data RepositóriUM: projeto de implementação do repositório de dados para a Universidade do Minho

Pedro Principe^a, Paulo Agostinho Gomes^b, Eloy Rodrigues^c

^a*Serviços de Documentação da Universidade do Minho, pedroprincipe@sdum.uminho.pt*

^b*Serviços de Documentação da Universidade do Minho, agomes@sdum.uminho.pt*

^c*Serviços de Documentação da Universidade do Minho, eloy@sdum.uminho.pt*

Resumo

Os Serviços de Documentação da Universidade do Minho (SDUM) têm previsto no seu plano de atividades para 2018 a criação de um repositório de dados. Esta iniciativa decorre do seu plano de intervenção para apoiar a gestão de dados na Universidade, iniciado em 2014 com o estudo sobre a análise dos dados produzidos na UMinhoⁱ, e enquadra-se na estratégia institucional proposta pelo grupo de trabalho nomeado pela Reitoria para elaboração de um estudo e recomendações para a gestão de dados de investigação na UMinho. O estudo prospetivo apresenta recomendações no domínio da i) formulação da política para a gestão de dados; ii) disponibilização de infraestrutura e sistemas para dados, iii) criação de serviços e ferramentas de apoio à gestão dos dados, iv) dinamização de atividades de capacitação, e v) apoio a questões legais, de proteção de dados e de propriedade intelectual.

Para a área das infraestruturas, foi estabelecido o objetivo de disponibilizar uma infraestrutura de sistemas e serviços integrados a colocar à disposição da comunidade para a gestão de dados ao longo do ciclo de vida da investigação. Esta infraestrutura incluiu a disponibilização do “*Data RepositóriUM*”, um sistema de repositório de dados para registo e abertura dos dados, permitindo o depósito de dados abertos, embargados ou fechados, com atribuição de identificadores persistentes e a preservação dos conjuntos de dados.

Esta proposta de *flash talk* para o 4º Fórum GDI visa apresentar os trabalhos em curso para a implementação do Repositório de Dados da UMinho – *Data RepositóriUM*, incluindo a caracterização do sistema, a solução do software escolhido e as principais etapas do projeto.

O Repositório de Dados da UMinho está a ser desenvolvido com recurso ao software *Dataverse*, uma plataforma desenvolvida e em utilização pela Universidade de Harvard, que se apresenta como uma solução *open source* de aplicação web para partilha, preservação, citação, exploração e análise de dados de investigação. Os SDUM tiveram oportunidade de explorar outros sistemas de software, nomeadamente os programas *DSpace*, *INVENIO* e *CKAN*, todos com reconhecido potencial como plataforma de repositório de dados, desde os seus requisitos técnicos e de armazenamento, ao potencial de desenvolvimento e adaptação dos sistemas, no entanto, para aplicação à realidade da UMinho, a escolha recaiu sobre o *software Dataverse* em função de alguns fatores diferenciadores, dos quais se destacam: i) a versatilidade na organização de

coleções (designados de *dataverses*), ii) a polivalência na adaptação dos esquemas de metadados disciplinares específicos para cada *dataverse*, e iii) a flexibilidade na definição de papéis e responsabilidades nos workflows de curadoria de dados e na administração dos *dataverses*.

No processo de criação do “Data RepositóriUM”, foram já realizadas as tarefas de instalação dos pré-requisitos e do software, definição de especificação de armazenamento e tarefas de segurança, bem como instalação de sistemas complementares. Está em curso no mês de setembro a definição do sistema de autenticação e a conceção de imagem e configuração de configuração de layout. Foi definida a estrutura de *dataverses* (coleções) e encontra-se em curso a análise dos esquemas de metadados e definição de modelos de termos de uso (setembro-outubro). A fase de teste do sistema com integração de quatro comunidades piloto irá decorrer no último trimestre de 2018, com definição das tarefas de curadoria, o depósito e registo de dados. Este processo de estabelecimento do repositório de dados da UMinho incluiu ainda a formação de curadores nos SDUM e a elaboração do guia de utilização, prevendo-se a apresentação pública do novo serviço em janeiro de 2019.

Palavras-chave: Repositório de Dados, Data RepositóriUM, Dados Abertos.

Designação do projeto/iniciativa

Data RepositóriUM.

Público-alvo

Gestores de repositórios e data centers, bibliotecários, especialistas de informática.

Ligações web úteis

<https://dataverse.org>

ⁱ Relatório do estudo: <http://hdl.handle.net/1822/29948>; Dados: <http://hdl.handle.net/1822/29949>.