



DatarepositóriUM: uma visão 360° ao estabelecimento de um repositório digital de dados de investigação

Pedro Príncipe^a, Paula Moura^b, André Vieira^c, Antónia Correia^d

^a*Universidade do Minho, Portugal, pedro.principe@usdb.uminho.pt*

^b*Universidade do Minho, Portugal, paula.moura@usdb.uminho.pt*

^c*Universidade do Minho, Portugal, andre.vieira@usdb.uminho.pt*

^d*Universidade do Minho, Portugal, antonia.correia@usdb.uminho.pt*

Resumo

Esta apresentação tem como objetivo dar uma visão 360° ao estabelecimento de um serviço de repositório digital institucional de dados de investigação. A disponibilização e funcionamento do serviço DatarepositóriUM – Repositório de Dados da Universidade do Minho – surge no contexto da relevância atual das instituições de investigação e ensino superior conceberem estratégias de atuação que maximizem o valor dos dados gerados nas atividades de investigação que realizam, procurando assegurar boas práticas de gestão de dados. As bibliotecas e os arquivos assumem particular responsabilidade no estabelecimento das infraestruturas e serviços necessários à realização das estratégias e políticas de atuação institucional da gestão e abertura de dados de investigação. A sua atuação é de particular relevo na curadoria de dados, assegurando a qualidade de metadados necessária para que os conjuntos de dados possam ser localizados, a sua usabilidade ao longo do tempo e a garantia do potencial de reprodutibilidade. O Repositório de Dados institucional da UMinho serve para partilhar, publicar e gerir dados de investigação gerados e recolhidos pela atividade dos investigadores e nas unidades de investigação da Universidade do Minho. Nesta apresentação está documentada a sua missão e objetivos, a normas de utilização e termos de uso, as estratégias de disseminação do serviço e as atividades em curso para integração em sistemas internacionais e a obtenção de certificação.

Palavras-chave: Repositório digital, Repositório de Dados, Dados de Investigação.

Introdução

As instituições de investigação e ensino superior reconhecem atualmente a importância de conceber estratégias de atuação que maximizem o valor dos dados gerados nas atividades de investigação que realizam, procurando responder a requisitos dos organismos financiadores da ciência sobre a disponibilização a abertura de dados e apoiando a comunidade em assegurar a preservação dos dados gerados no âmbito da atividade e projetos de investigação.

As bibliotecas e os arquivos assumem particular responsabilidade no estabelecimento das infraestruturas e serviços necessários à realização das estratégias e políticas de atuação institucional para a gestão e abertura de dados de investigação. Neste contexto, a disponibilizarão de serviços de repositórios de dados que promovam boas práticas de curadoria de dados, mecanismos para publicação ágil de conjuntos de dados e ferramentas para assegurar a preservação dos dados, é de particular importância.

Esta apresentação visa detalhar e documentar a disponibilização e funcionamento do serviço

DataRepositóriUM – Repositório de Dados da Universidade do Minho para partilhar, publicar e gerir dados de investigação.ⁱ

Missão e Objetivos

Repositório de Dados institucional da UMinho serve para partilhar, publicar e gerir dados de investigação gerados e recolhidos pela atividade dos investigadores e nas unidades de investigação da Universidade do Minho (figura 1). Apresenta-se com os seguintes objetivos:

- Depositar ficheiros em *datasets* e coleções organizadas
- Preservar ficheiros e *datasets*
- Apoiar o cumprimento dos requisitos de financiadores
- Registrar identificadores persistentes (DOI)
- Permitir o acesso controlado e sob identificação prévia
- Associar licenças de reuso e explicitar os termos de uso
- Documentar os *datasets* com metadados disciplinares
- Facilitar a disseminação dos resultados de investigação
- Viabilizar a citação de dados e o crédito aos produtores
- Facilitar a disponibilização de dados aos editores



FIGURA 1 – Resumo dos principais objetivos do DataRepositóriUM.

Termos de uso e regras gerais de utilização

A disponibilização de um serviço de repositório de dados requer o estabelecimento de regras de utilização e termos de usoⁱⁱ que definam de forma muito objetiva as regras gerais de conduta, os serviços e as restrições dos carregamentos realizados pelo utilizador, o contrato de licença de utilização dos dados submetidos pelo utilizador, entre outros. No DataRepositóriUM, o utilizador concede todas as permissões e licenças necessárias para tornar o conteúdo submetido ou depositado disponível para arquivo, preservação e acesso. Isto inclui, sem restrições, permissão para:

1. Disseminar cópias do conteúdo em vários formatos de distribuição de acordo com os termos padrão de uso do DataRepositóriUM;
2. Promover e divulgar o conteúdo em qualquer campanha publicitária (em qualquer forma) para o DataRepositóriUM;

3. Descrever, catalogar e documentar os depósitos / submissões do utilizador;
4. Armazenar, traduzir, copiar ou reformatar o Conteúdo de qualquer forma para garantir sua preservação e acessibilidade futuras, e melhorar a usabilidade e/ou proteger a confidencialidade dos respondentes;
5. Incorporar Metadados ou documentação no Conteúdo em catálogos de acesso público.

Estes termos de uso definem o âmbito do serviço que dá a possibilidade de publicar carregamentos realizados pelo utilizador e de os expor, organizar, aceitar e disseminar através da criação de uma coleção. O utilizador reconhece que o DataRepositóriUM não endossa, nem assume responsabilidade por qualquer carregamento dos utilizadores. Existem algumas restrições das quais se destacam as que o utilizador terá de garantir ao DataRepositóriUM que os depósitos não infringem os direitos de autor ou outros direitos de propriedade intelectual, incluindo, entre outros, patentes, marcas registradas, segredos comerciais, direitos de publicidade ou outros direitos de terceiros, que os depósitos não violam nenhuma lei e que devem ser esvaziados de todas as informações identificáveis (dados pessoais conforme a alínea 1) do artigo 4º do RGPD).

Divulgação e suporte ao uso do serviço de repositório de dados

A disponibilização de um novo serviço numa instituição requer um plano de comunicação e de suporte para o tornar útil e passível de utilização pela comunidade. A equipa das Bibliotecas da UMinho, responsável pela coordenação do serviço, tem em curso atividades de divulgação que incluem a realização de *webinars*ⁱⁱⁱ, sessões de divulgação (DataRepositóriUM roadshows), inclusão do serviço nas sessões do programa de formação para a Ciência Aberta e preparação de materiais de apoio, como guias online^{iv} e tutoriais vídeo^v.

O principal objetivo das ações de comunicação e *helpdesk* é destacar qual a utilidade do serviço, as vantagens, as formas de adesão e os benefícios, nomeadamente:

- Gestão de dados segura – possibilita o estabelecimento de limitações ao uso ou acesso aos dados, através da criação de termos de uso e restrições e fornece igualmente uma cópia de segurança para guarda.
- Partilha efetiva – é uma forma adequada de divulgar dados, facilitando a colaboração de uma equipa de investigação dentro de um espaço seguro.
- Acompanhamento das mudanças – fornece controle de versão de ficheiros.
- Acesso e preservação a longo prazo – a utilização de identificadores persistentes e a normalização dos dados asseguram um acesso fiável, garantindo a prevenção da obsolescência dos dados.
- Organização e compatibilidade – pode-se criar o próprio arquivo de dados pessoais na web, em conformidade com os padrões internacionais de dados.
- Poupança de tempo – tem uma interface fácil de usar para partilhar dados.
- Aumento da visibilidade da pesquisa – aumentar o reconhecimento académico pelo trabalho além das publicações científicas.
- Cumprimento dos requisitos do financiador – as agências de financiamento exigem aos investigadores o depósito num repositório dos dados que são recolhidos como parte do seu projeto de investigação.

No *website* do serviço são destacadas de forma visual e de leitura fácil as vantagens da sua utilização, onde se procura deixar claro os ganhos imediatos do seu uso (figura 2).



FIGURA 2 - Identificação das vantagens do uso do DataRepositóriUM. <https://datarepositorium.uminho.pr/about>

Curadoria de dados

A curadoria de dados é um dos aspetos chave da gestão de dados de investigação numa instituição e que determina de forma crítica os processos de documentação, preservação e publicação dos dados. O DatarepositóriUM serve a operacionalização de um fluxo de curadoria de dados bem articulado no ciclo de vida de dados, auxiliando na melhoria da qualidade dos metadados, descrevendo de forma adequada o seu conteúdo, criando uma estrutura coerente com documentação explicativa e relacionando-os com outros resultados de investigação.

A equipa de suporte à publicação de dados procura assegurar materiais de apoio e atividades de validação de metadados que assegurem diferentes níveis de curadoria na gestão dos dados, nomeadamente:

- assegurar tarefas de curadoria para os *datasets* serem publicados no DatarepositóriUM com um nível mínimo de qualidade necessário para que estes possam ser localizados.
- assegurar tarefas de curadoria que potenciem a descoberta e ajuda a garantir a usabilidade e potencial de reprodutibilidade dos conjuntos de dados ao longo do tempo.
- assegurar tarefas de curadoria que garantam medidas de preservação dos *datasets*.

Integrações e Serviços Externos

O DataRepositóriUM é um serviço institucional que está registado e aceite nos principais diretórios mundiais ou bases de dados europeias de referência, nomeadamente os diretórios RE3DATA^{vi} e FAIRSHARING^{vii} e no agregador europeu OpenAIRE^{viii} (figura 3).



FIGURA 3 – Identificação dos serviços onde o DataRepositóriUM está registado e que lhe conferem confiabilidade

Certificação Core Trust Seal

Por último, destacamos a importância de credibilizar os serviços digitais de gestão de dados de investigação, já que a utilização de um repositório de dados está muito dependente do grau de confiabilidade que os potenciais utilizadores atribuem à ferramenta. Deste modo, estão a ser levadas a cabo as atividades com vista à obtenção da certificação CoreTrustSeal^{ix} para o DataRepositóriUM. Para além do cumprimento de requisitos técnicos e do claro estabelecimento das normas gerais de utilização e dos termos de uso, já muito bem definidas atualmente, estão a ser trabalhadas as políticas de curadoria e de preservação para que os requisitos da certificação possam ser satisfeitos de forma cabal e com avaliação positiva.

Conclusão

Com esta apresentação esperamos dar uma visão 360º ao estabelecimento de um serviço de repositório digital institucional para apoiar os investigadores e unidades de investigação e desenvolvimento da Universidade do Minho na publicação e preservação de conjuntos de dados de investigação, procurando valorizar os dados de como resultados com maior relevância e centralidade na visibilidade da investigação e passíveis de serem citáveis e medidos para efeitos de avaliação da ciência.

Referências bibliográficas

Guias (in)formativos: Guia do DataRepositóriUM: Home. (2022, julho). Home - Guia Do DataRepositóriUM - Guias (in)Formativos at Serviços de Documentação Da Universidade Do Minho. <https://guias.sdum.uminho.pt/datarepositorium>

Data Repositories Requirements. (2017, January 2). CoreTrustSeal. <https://www.coretrustseal.org/why-certification/requirements/>

DataRepositóriUM | re3data.org. (n.d.). DataRepositóriUM | re3data.org. <https://www.re3data.org/repository/r3d100013173>

DataRepositóriUM. (n.d.). FAIRsharing. <https://fairsharing.org/3845>

DataRepositóriUM. (n.d.). OpenAIRE - Explore. <https://explore.openaire.eu/search/dataprovider?pid=r3d100013173>

Unidade de Serviço de Bibliotecas e Documentação da Universidade do Minho. (2023, fev.). *DataRepositóriUM.* DataRepositóriUM. Retrieved February 17, 2023, from <https://datarepositorium.uminho.pt/>

Unidade de Serviço de Bibliotecas e Documentação da Universidade do Minho. (2023). *DataRepositóriUM: Para partilhar, publicar e gerir dados de investigação.* DataRepositóriUM. Retrieved February 17, 2023, from <https://datarepositorium.uminho.pt/about/>

Unidade de Serviço de Bibliotecas e Documentação da Universidade do Minho. (2023). *Termos de uso gerais do DataRepositóriUM da Universidade do Minho*. DataRepositoriUM. Retrieved February 17, 2023, from <https://datarepositorium.uminho.pt/about/termosuso.html>

ⁱ <https://datarepositorium.uminho.pt>

ⁱⁱ <https://datarepositorium.uminho.pt/about/termosuso.html>

ⁱⁱⁱ <https://datarepositorium.uminho.pt/about/webinars.html>

^{iv} <https://guias.sdum.uminho.pt/datarepositorium>

^v <https://www.youtube.com/playlist?list=PLIm8KTWNUm7WQariy5VqPcCSVlf2sU2Uu>

^{vi} <https://www.re3data.org/repository/r3d100013173>

^{vii} <https://fairsharing.org/3845>

^{viii} <https://explore.openaire.eu/search/dataprovider?pid=r3d100013173>

^{ix} <https://www.coretrustseal.org/why-certification/requirements/>