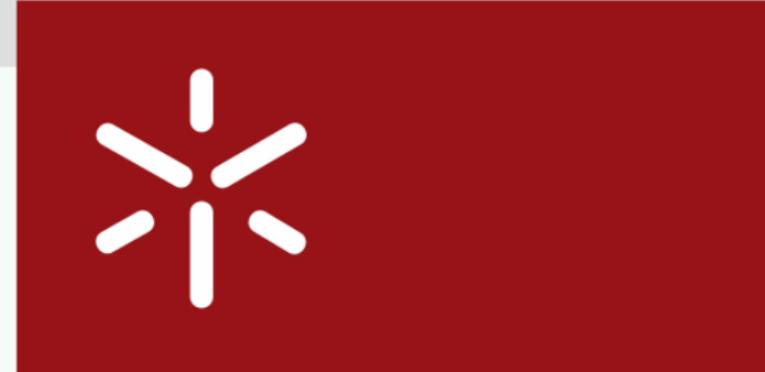
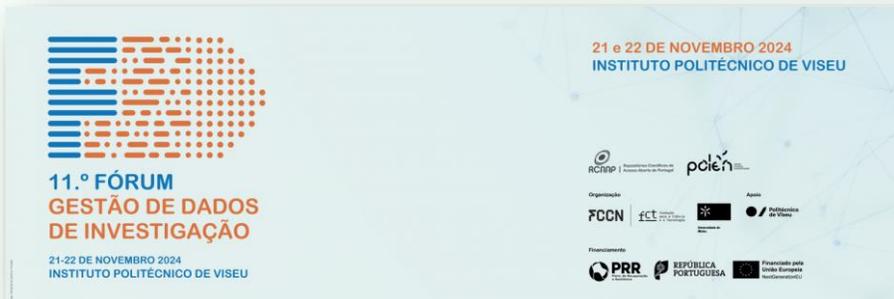


GESTÃO DE INFORMAÇÃO CIENTÍFICA,
REPOSITÓRIOS E CIÊNCIA ABERTA

Pedro Principe, André Vieira, Paula Moura



Universidade do Minho

Serviço de Documentação e Bibliotecas

DataRepositóriUM

Destaque às funcionalidades do Dataverse em uso
no Repositório Institucional de Dados da UMinho

22 de novembro de 2024



DataRepositóriUM

para boa gestão de dados e
prática de Ciência Aberta

Números a 21/11/2024

94 Dataverses + **8** por publicar

235 Datasets + **154** por publicar

1707 Files + **555** por publicar

29750 Downloads

<https://datarepositorium.uminho.pt>



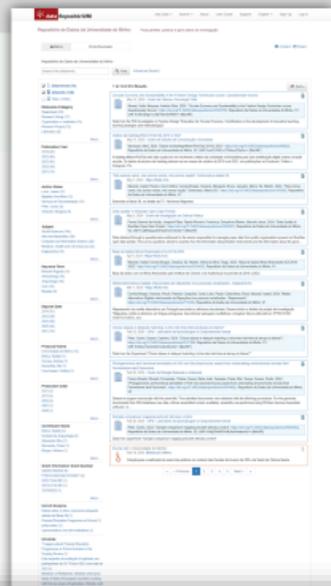
PLATAFORMA

software



referência

<https://dataverse.harvard.edu/>



DATAVERSE É UM REPOSITÓRIO

localizar, partilhar, citar e publicar dados

DATAVERSE É UM SOFTWARE

Plataforma de software repositório

DATAVERSE É UMA COMUNIDADE

Para promover a gestão e partilha de dados a nível mundial

<https://dataverse.org/>

Open source research data repository software

- Researchers** Enjoy full control over your data. Receive web visibility, academic credit, and increased citation counts. A personal Dataverse collection is easy to set up, allows you to display your data on your personal website, can be featured uniquely on your research program, makes your data more discoverable to the research community, and optimizes data management plans. Want to set up your personal Dataverse collection?
- Journals** Seamlessly manage the submission, review, and publication of data associated with published articles. Establish an end-to-end link between journals and your personal Dataverse collections. Participate in the open data movement by using a Dataverse collection as part of your journal data policy or list of repository recommendations. Want to find out more about journal Dataverse collections?
- Institutions** Establish a research data management solution for your community. Federate with a growing list of Dataverse repositories worldwide for increased discoverability of your community's data. Participate in the drive to set norms for storing, preserving, citing, reusing, and analyzing research data. Want to install a Dataverse repository?
- Developers** Participate in a vibrant and growing community that is helping to drive the norms for sharing, generating, citing, reusing, and analyzing research data. Contribute code extensions, documentation, testing, and/or standards. Integrate research analysis, visualization and exploration tools, or other research and data workflow systems into the Dataverse Project. Want to contribute?

DATAVERSE REPOSITORIES - A WORLD VIEW

Roadmap: The Dataverse Project

Strategic Goals

The Strategic Goals of the Dataverse Project are to:

- Grow the Dataverse community
- Empower the open source community to explore and implement new Dataverse applications, tools, and services
- Develop the capability to handle sensitive data and big data
- Expand data and metadata features for existing and new disciplines
- Expand archival and preservation features
- Increase interoperability through the implementation of standards
- Increase contributions from the open source development community
- Improve the Dataverse user experience

Roadmap

The Dataverse UI is being redesigned as a single-page application (SPA) in which internal and external Dataverse services will be delivered solely through improved APIs. This change will improve UI responsiveness and empower the Dataverse community to develop their own UIs and applications using extended Dataverse API endpoints.

Read more about our roadmap for the Dataverse infrastructure redesign in: [Restructuring the Dataverse UI as a Single-Page Application](#)

Current and Past Projects

- Generalist Repository Ecosystem Initiative (GREI)** Multi-year grant starting in 2022 to develop collaborative approaches for data management and sharing through inclusion and enrichment of generalist repositories in the NIH data ecosystem, including the Harvard Dataverse Repository. The [announcement](#) has details of areas being explored.
- FAIR Principles for Geospatial Data** Supplemental grants to expand Harvard Dataverse Repository's capacity to store and share geospatial public health data. For details, see: [National Cohort Studies of Alzheimer's Disease, Delayed Dementias and Air Pollution](#) (Aims 1-3) and [Approaches for AI/ML Readiness for Wildlife Exposures](#) (Aim 3).

Project Board

Our [project board on GitHub](#) is open for public review and comment. This is where we track development and QA activities. Often these items are discussed during [community calls](#), but please feel free to give feedback and [join](#) our community through any [channel](#).

Recent Releases

- 6.4** This release includes several bug fixes and many new features. Features include enhanced DataCite metadata, including "Relation Type", support for all ISO 639-3 languages, support for [Crosswalks](#), and support for dataset types like "software" or "workflow".
- 6.3** Updates include a new [Contributor Guide](#); improvements to search performance, controlled vocabulary support, and RO-Crate detection; support for file-level retention periods; and sitemaps support for more than 50,000 items.
- 6.2** This release contains a number of updates and new features including search and facet by license, rate limiting, support for using multiple PID providers, and more.
- 6.1** Updates include new guestbook at request and collection-level storage quotas features, improved Globus support, support for JSON validation of datasets prior to creation, new or improved API endpoints, and improvements to OpenID Connect (OIDC) and SSO.
- 6.0** This is a platform upgrade release. Payara, Solr, and Java have been upgraded. No features have been added to the Dataverse software itself. Only a handful of bugs were fixed.
- 5.14** Improved indexing performance, MDC (Make Data Count) and legacy access count support, Signposting for Dataverse, Permalinks support, Datasets API extended 2023/08/04.
- 5.13** Schema.org improvements, DVWebloader, NetCDF/HDF5 support, Geospatial Search, CodeMeta 2.74/23.

<https://www.iq.harvard.edu/roadmap-dataverse-project>

Data RepositóriUM – para partilhar, publicar e gerir dados

Depósito de ficheiros em coleções organizadas no ciclo de vida da investigação



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- No processamento dos dados posso depositar e documentar os dados (*atenção, não é serviço de box*)
- Ter uma coleção organizado com datasets de um projeto específico
- Processar dados e documentar e depois partilhar em equipa para acesso limitado
- Apoiar a execução de um plano de dados
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Registo de identificadores persistentes (DOI)



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Promove a citação de dados
- Apoiar a publicação de dados ligados a artigos e resultados em conferências (ex. em poster ou apresentações)
- Requisito básico dos princípios dos dados FAIR (findable & accessible)
- Requisito para report em projetos financiados (FCT, CE)
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Apoiar o cumprimento dos requisitos de financiadores



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Com facilidade podemos incluir registos de dados nos reports dos projetos (H2020, HE, FCT...).
- Incluir a referência ao repositório da elaboração de um plano de gestão de dados.
- Indicar que a documentação dos datasets é devidamente assegurada pelos metadados no repositório
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Depositar e publicar dados como parte integrante de teses e dissertações



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Facilitar a documentação dos dados ao longo do ciclo de vida de investigação numa tese de doutoramento.
- Apoiar a gestão dos dados por via da disponibilização de uma coleção para organizar os datasets.
- Registrar os dados e depois ligar ao produto final (tese)
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Facilitar a disponibilização de dados aos editores



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- O repositório apoia o cumprimento dos requisitos de editores para a disponibilização de dados.
- Permite enviar ou disponibilizar dados para processo de revisão por pares.
- Facilita ligação dos dados aos resultados presentes nos artigos publicados.
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Documentar os datasets com metadados gerais e disciplinares



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Documentar de forma adequada os datasets gerados na investigação.
- Usar standards de metadados para descrever os conjuntos de dados.
- Poder referenciar o DataRepositóriUM no plano de gestão de dados ou numa proposta de projeto (FCT ou CE).
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Facilitar a disseminação dos resultados de investigação



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Ter um link (URL) para a coleção dos dados de um projeto.
- Poder incluir imagens (logos da instituição, centro ou do financiador) nas páginas das coleções de dados.
- Divulgar os DOI dos conjuntos de dados em conferências ou em reuniões de grupo.
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Registrar e licenciar os dados para referência, reuso e citação



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Associar licenças para reuso.
- Promover a referência e a citação dos dados
- Explicitar os termos de uso.
- Apoiar os investigadores a cumprirem requisitos de reutilização incluídos nas diretrizes de financiadores.
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Permitir acesso controlado aos dados ou sob identificação prévia



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Facilitar de forma clara o princípio de "As Open as possible, as closed as necessary"
- Apoiar os investigadores a cumprirem requisitos mesmo que necessitem ter dados com acesso limitado.
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

Data RepositóriUM – para partilhar, publicar e gerir dados

Preservação de ficheiros e datasets



Digitalbevering.dk

O QUE VEM RESOLVER OU APOIAR:

- Repositório confiável de dimensão institucional
- Assegura padrões de sustentabilidade alinhados com requisitos de financiadores
- ...

Universidade do Minho | Unidade de Serviços de Documentação e Bibliotecas

DESTAQUES A FUNCIONALIDADES EM UTILIZAÇÃO

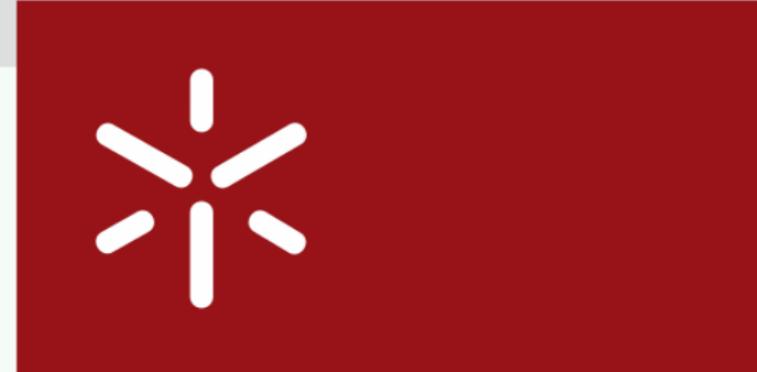
6 MAIS UTILIZADAS OU QUE OS UTILIZADORES MAIS APRECIAM

- Customização das coleções ou dataverses para **projetos**.
- Partilha ou identificação de dados para os **editores/revistas**.
- Especificação de direitos de acesso aos ficheiros diferenciados (acesso restrito com **termos de acesso**).
- **Co-edição** dos datasets (registos).
- Versionamento (histórico documentado das versões dos dados).
- URL para **incorporar dataset** em portais web.

3 FUNCIONALIDADES IMPORTANTES MAS AINDA SUBEXPLORADAS NA GESTÃO

- ✓ Integração de mais vocabulários controlados e metadados disciplinares.
- ✓ Disponibilizar mais possibilidades de personalização de Termos de uso.
- ✓ Integração de APIs para registo de mais identificadores persistentes.





Universidade do Minho

Serviço de Documentação e Bibliotecas

OBRIGADO

datarepositorium@usdb.uminho.pt

