

8º FÓRUM GESTÃO DE DADOS DE INVESTIGAÇÃO

25 e 26 DE NOVEMBRO 2021
UNIVERSIDADE DE COIMBRA



Organização



Patrocínio



14h00-16h00 | **WORKSHOP**

CURADORIA DE DADOS EM REPOSITÓRIOS: FLUXOS DE TRABALHO E SERVIÇOS

Moderadores: Pedro Príncipe e Rita Gonçalves

Organização:

Grupo de Trabalho Fórum GDI – Repositórios de dados: tecnologia, organização e certificação

Enquadramento e Objetivos

**Organizado no âmbito do Grupo de Trabalho
“Repositórios de Dados: Tecnologia, organização e certificação”**

Objetivos - partilhar:

- práticas de gestão e curadoria de dados, enquadrando o papel que os repositórios podem desempenhar no ciclo de vida dos dados e nos serviços prestados pelas instituições;
- características, funcionalidades e casos de uso de plataformas e serviços de repositórios para instituições.

AGENDA

14h00: Apresentação do programa e objetivos do workshop

- Mentimeter sobre serviços de repositórios de dados (Rita Gonçalves)

14h10: Os repositórios no ciclo de vida dos dados

- Workflows de curadoria, publicação de dados e cenários de uso (Pedro Príncipe)

14h30: Por onde andam os dados dos investigadores portugueses?

- Identificação de fontes de informação sobre dados de investigadores portugueses (Filomena Borba, Salima Rehemtula e André Vieira)

14h40: Plataformas e serviços de repositórios para instituições: características, demonstrações e casos de uso

- Caracterização de softwares de gestão e publicação de dados para instalação própria (Rita Gonçalves)
- Serviços genéricos: o caso do Zenodo (Pedro Príncipe) e Figshare (Mark Hahnel, CEO e fundador do Figshare)
- Serviços temáticos: o caso do GBIF - the Global Biodiversity Information Facility (Rui Figueira, Coordenador do Nó Português do GBIF, Instituto Superior de Agronomia)
- Software Dataverse para repositórios institucionais (André Vieira e José Carvalho)

15h50: Identificação das principais necessidades na implementação de serviços de repositórios em instituições e na publicação de dados pelos investigadores

- Notas finais e mentimeter

Vá para www.menti.com e use o código 6773 6660





8º FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

25 e 26 DE NOVEMBRO 2021
UNIVERSIDADE DE COIMBRA

Os repositórios no ciclo de vida dos dados

Pedro Príncipe



8º FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

25 e 26 DE NOVEMBRO 2021
UNIVERSIDADE DE COIMBRA

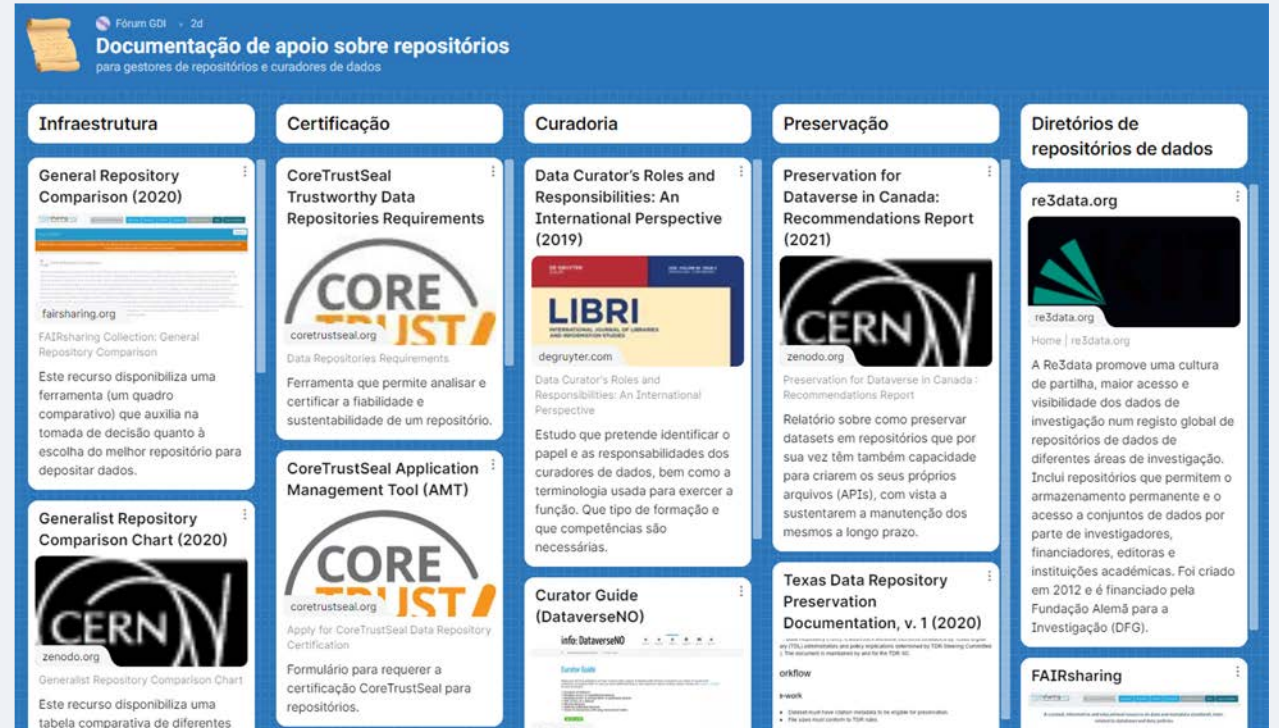
Por onde andam os dados dos investigadores portugueses?

Filomena Borba

Plataformas e serviços de repositórios para instituições

Documentação de apoio sobre repositórios para gestores de repositórios e curadores de dados

https://padlet.com/forumgdi/documentacao_repositorios



The screenshot shows a Padlet board with the following content:

- Infraestrutura**
 - General Repository Comparison (2020)** (fairsharing.org): Ferramenta que permite analisar e certificar a fiabilidade e sustentabilidade de um repositório.
 - Generalist Repository Comparison Chart (2020)** (zenodo.org): Este recurso disponibiliza uma tabela comparativa de diferentes repositórios.
- Certificação**
 - CoreTrustSeal Trustworthy Data Repositories Requirements** (coretrustseal.org): Ferramenta que permite analisar e certificar a fiabilidade e sustentabilidade de um repositório.
 - CoreTrustSeal Application Management Tool (AMT)** (coretrustseal.org): Formulário para requerer a certificação CoreTrustSeal para repositórios.
- Curadoria**
 - Data Curator's Roles and Responsibilities: An International Perspective (2019)** (degruyter.com): Estudo que pretende identificar o papel e as responsabilidades dos curadores de dados, bem como a terminologia usada para exercer a função. Que tipo de formação e que competências são necessárias.
 - Curator Guide (DataverseNO)** (info-dataverseNO): Guia para curadores de dados.
- Preservação**
 - Preservation for Dataverse in Canada: Recommendations Report (2021)** (zenodo.org): Relatório sobre como preservar datasets em repositórios que por sua vez têm também capacidade para criarem os seus próprios arquivos (APIs), com vista a sustentarem a manutenção dos mesmos a longo prazo.
 - Texas Data Repository Preservation Documentation, v. 1 (2020)** (orkflow): Documentação sobre preservação de dados.
- Diretórios de repositórios de dados**
 - re3data.org**: A Re3data promove uma cultura de partilha, maior acesso e visibilidade dos dados de investigação num registo global de repositórios de dados de diferentes áreas de investigação. Inclui repositórios que permitem o armazenamento permanente e o acesso a conjuntos de dados por parte de investigadores, financiadores, editoras e instituições académicas. Foi criado em 2012 e é financiado pela Fundação Alemã para a Investigação (DFG).
 - FAIRsharing**: Plataforma de repositórios de dados.

Plataformas e serviços de repositórios para instituições

Caracterização de softwares de gestão e publicação de dados para instalação própria

<https://bit.ly/3r2tO34>

Softwares de gestão e publicação de dados de investigação

Tabela comparativa de características e funcionalidades de quatro softwares para o depósito, gestão e publicação de dados de investigação disponíveis para instalação própria.

	DATAVERSE	INVENIO RDM	CKAN	DSPACE
open source	✓	✓	✓	✓
identificadores persistentes de registos	DOI; Handle	DOI	DOI	Handle; DOI
outros identificadores persistentes (metadados)	ORCID; ISNI; ...	ORCID; ROR; ...	ORCID	ORCID
esquemas de metadados disciplinares	Geoespaciais; Ciências Sociais e Humanas; Astronomia e Astrofísica; Ciências da Vida	×	×	×
integração de outros esquemas de metadados	✓	—	✓	✓
opções de licença	Dedicção ao Domínio Público (CC0)	Licenças Creative Commons	Creative Commons; Open Data Commons; ...	Licenças Creative Commons
personalização de termos de uso	✓	✓	✓	✓
definição de condições de acesso aos ficheiros	Acesso aberto, restrito, embargado ou fechado	Acesso aberto, restrito, embargado ou fechado	Acesso aberto, restrito ou fechado	Acesso aberto, restrito, embargado ou fechado
formatos de exportação de metadados	Dublin Core; DDI; DDI HTML; Codebook; DataCite 4; JSON; OAI; ORE; OpenAIRE; Schema.org; JSON-LD	JSON; Citation Style Language; JSON; DataCite; JSON/XML; Dublin Core	JSON	CSV; OAI
vocabulários controlados	✓	✓	✓	✓
integração com outros sistemas	✓	✓	✓	✓
gestão de papéis de utilizadores	✓	—	✓	✓
versionamento	✓	✓	✓	✓
número de instalações (registadas em re3data.org)	112	—	79	102

novembro 2021 | versão 1.0
tabela detalhada disponível em: <https://bit.ly/3r2tO34>



Plataformas e serviços de repositórios para instituições

**Serviços genéricos:
o caso do Zenodo (Pedro Príncipe)
e Figshare (Mark Hahnel)**

Plataformas e serviços de repositórios para instituições

Serviços temáticos:

**o caso do GBIF - the Global Biodiversity Information Facility
(Rui Figueira, Coordenador do Nó Português do GBIF,
Instituto Superior de Agronomia)**

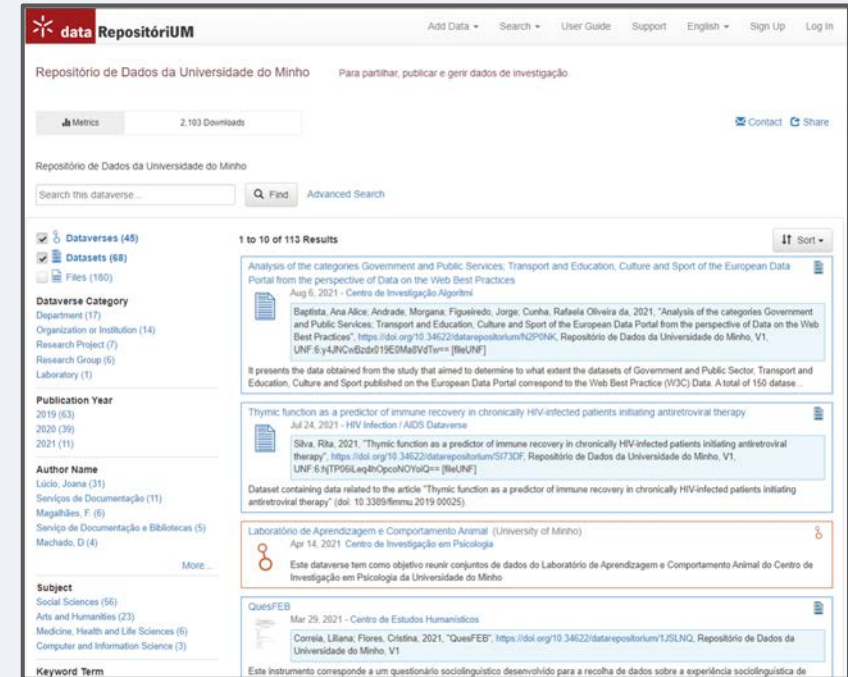
Plataformas e serviços de repositórios para instituições

Software Dataverse para repositórios institucionais

O exemplo do DataRepositóriUM, o repositório de dados da Universidade do Minho

Demonstração:

- **Publicação de dados ao longo do ciclo de vida**
- **Metadados para descrição dos conjuntos de dados**
- **Controlo de acesso aos dados**
(*Tão Aberto Quanto Possível, Tão Fechado Quanto Necessário*)
 - Guestbook, Acesso privado, Acesso restrito a ficheiros
- **Páginas dedicadas a projetos de investigação**
 - Coleção específica e personalizável para a publicação dos dados resultantes do projeto
- **Ligação entre Dados e Publicações**
- **Gestão de utilizadores e permissões**



The screenshot displays the DataRepositóriUM website. At the top, there is a navigation bar with links for 'Add Data', 'Search', 'User Guide', 'Support', 'English', 'Sign Up', and 'Log In'. Below the navigation bar, the page title is 'Repositório de Dados da Universidade do Minho' with a subtitle 'Para partilhar, publicar e gerir dados de investigação'. A 'Metrics' section shows '2,163 Downloads'. A search bar is present with a 'Find' button and an 'Advanced Search' link. On the left, there are filters for 'Dataverses (45)', 'Datasets (68)', and 'Files (180)'. Below these are filters for 'Dataverse Category', 'Publication Year', 'Author Name', and 'Subject'. The main content area shows '1 to 10 of 113 Results'. The first result is 'Analysis of the categories Government and Public Services, Transport and Education, Culture and Sport of the European Data Portal from the perspective of Data on the Web Best Practices' by Baptista, Ana Alice, Andrade, Morgana, Figueiredo, Jorge, Cunha, Rafaela Oliveira da, dated Aug 5, 2021. The second result is 'Thymic function as a predictor of immune recovery in chronically HIV-infected patients initiating antiretroviral therapy' by Silva, Rita, dated Jul 24, 2021. The third result is 'Este dataverse tem como objetivo reunir conjuntos de dados do Laboratório de Aprendizagem e Comportamento Animal do Centro de Investigação em Psicologia da Universidade do Minho' dated Apr 14, 2021. The fourth result is 'QuesFEB' dated Mar 29, 2021.

Plataformas e serviços de repositórios para instituições

Software Dataverse para repositórios institucionais

O exemplo do DataRepositóriUM, o repositório de dados da Universidade do Minho

Gestão de utilizadores

- **Registo individual** de utilizadores
- **Acesso controlado** de novos utilizadores **através de validação** realizada pela equipa de gestão do repositório
- **Moderação das permissões** realizada pelos **curadores dos Dataverses** e equipa de **gestão do repositório**

Processo de curadoria na Criação de Publicação de Datasets

- Autonomia atribuída aos autores/utilizadores de acordo com as permissões atribuídas
 - **Publicação:** Curador
 - **Criação e Edição:** Colaborador e Criador

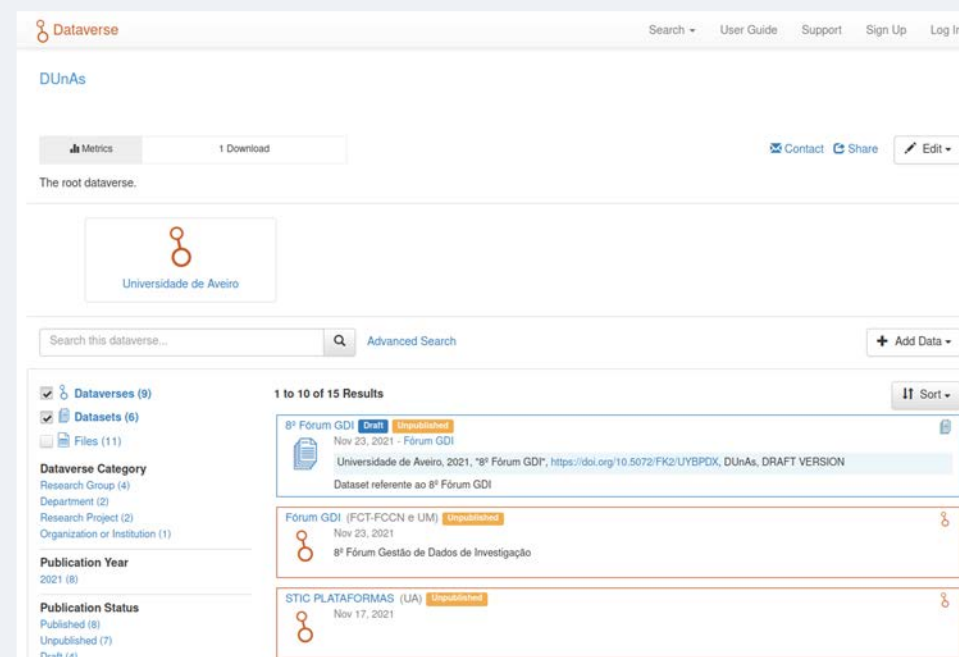
Plataformas e serviços de repositórios para instituições

Software Dataverse para repositórios institucionais

O exemplo do DUnAs, o repositório de dados da Universidade de Aveiro
(em desenvolvimento)

Demonstração:

- Opções de autenticação
- Dataset Harvesting
- Customização



The screenshot displays the Dataverse DUnAs interface. At the top, the Dataverse logo and navigation links (Search, User Guide, Support, Sign Up, Log In) are visible. The main header identifies the repository as 'DUnAs'. Below this, there are metrics (1 Download) and action buttons (Contact, Share, Edit). The content area shows 'The root dataverse.' with the Universidade de Aveiro logo. A search bar is present with the text 'Search this dataverse...' and an 'Advanced Search' link. On the left, a sidebar provides filters for 'Dataverses (9)', 'Datasets (6)', and 'Files (11)', along with 'Dataverse Category' (Research Group, Department, Research Project, Organization or Institution) and 'Publication Year' (2021). The main results area shows '1 to 10 of 15 Results' with a 'Sort' dropdown. Three results are visible, each with a 'Draft' or 'Unpublished' status and a 'Dataverse' icon.

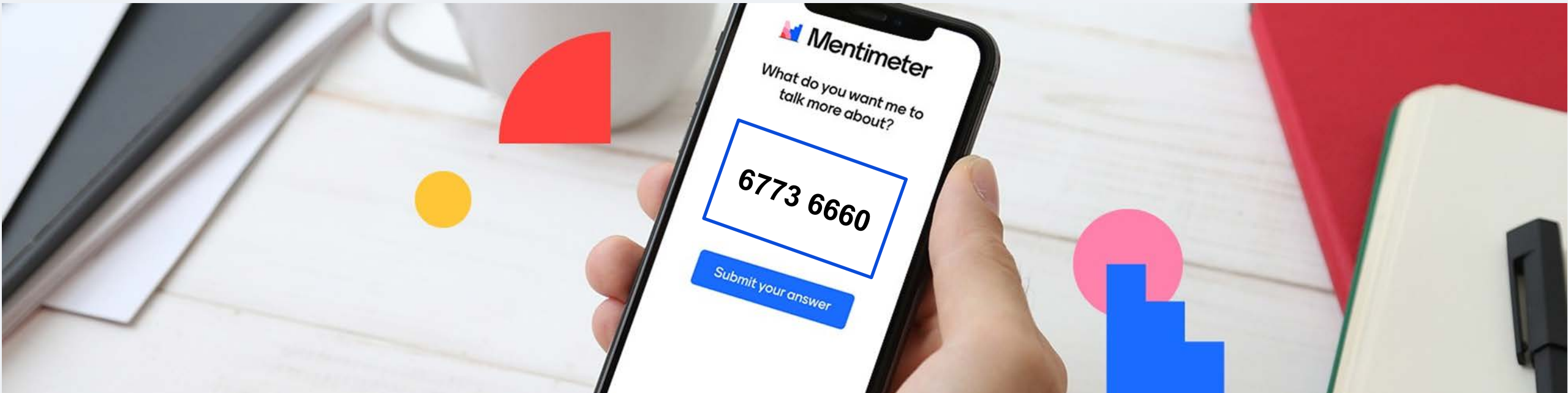


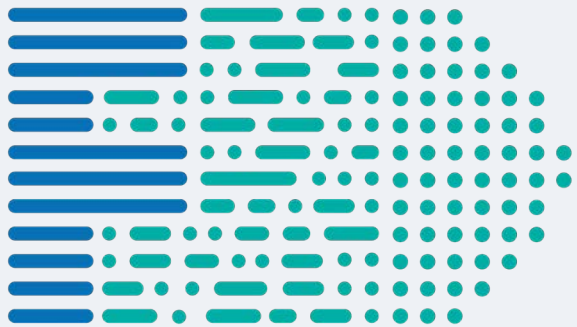
8º FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

25 e 26 DE NOVEMBRO 2021
UNIVERSIDADE DE COIMBRA

**Identificação das principais necessidades
na implementação de serviços de
repositórios em instituições e na
publicação de dados pelos investigadores**

Vá para www.menti.com e use o código **6773 6660**





8º FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

25 e 26 DE NOVEMBRO 2021
UNIVERSIDADE DE COIMBRA

Muito obrigado!