



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Repositórios de Dados
CASOS DE ESTUDO

CASOS DE ESTUDO

Casos de repositórios de dados institucionais e temáticos recolhidos pelo Grupo de Trabalho dos Repositórios de Dados do Fórum GDI

Todos os Casos de Estudo podem ser consultados na comunidade do Fórum GDI no Zenodo.

<https://zenodo.org/communities/forumgdi>

Neste momento de partilha, vamos solicitar a colaboração de todos para a apresentação dos mesmos num minuto.

Obrigada



Instituto Hidrográfico



O [Instituto Hidrográfico \(IH\)](#), órgão da Marinha Portuguesa, é reconhecido como Laboratório do Estado, dependendo assim do Ministério da Defesa Nacional em articulação com os Ministérios da Ciência, Tecnologia e Ensino Superior e da Economia e do Mar.

O IH tem por missão fundamental assegurar atividades relacionadas com as ciências e técnicas do mar, tendo em vista a sua aplicação na área militar, e contribuir para o desenvolvimento do País nas áreas científica e de defesa do ambiente marinho.

Hidrográfico+

<https://geomar.hidrografico.pt/>

O Hidrográfico+ é uma infraestrutura geoespacial de dados do meio marinho criada com o objetivo de facilitar o acesso do cidadão a dados produzidos pelo Instituto Hidrográfico e dessa forma fomentar o conhecimento e a ciência oceânica, promover a divulgação da importância dos oceanos, dinamizar a economia azul e contribuir para a sustentabilidade dos oceanos e o desenvolvimento da sociedade. Pretende dar resposta às recomendações e obrigações da legislação comunitária e nacional, nomeadamente, as Diretivas Europeias INSPIRE (2007) e Open Data Directive (2019) que reconhecem o valor dos dados geoespaciais.

A infraestrutura Hidrográfico+ permite a ligação ao repositório DSPACE, onde são catalogadas todas as publicações científicas produzidas pelos colaboradores da instituição.

O repositório é alimentado por diferentes áreas temáticas que correspondem às várias componentes técnicas do Instituto Hidrográfico: Geologia Marinha, Hidrografia, Oceanografia, Química e Poluição Marinha, Segurança da Navegação.

A infraestrutura Hidrográfico+ é constituída por um conjunto de servidores que prestam serviços e dão suporte às funcionalidades. A componente de acesso a dados é feita através de serviços de pesquisa, visualização e descarregamento de dados seguindo as normas OGC. A implementação está baseada em servidores *open source* Geoserver, Geonetwork e ncWMS. Paralelamente e para determinada tipologia de dados é utilizado um servidor Thredds.

A infraestrutura Hidrográfico+ é um dos nodos da rede distribuída do Centro Nacional de Dados Oceanográficos (<https://nodc-portugal.pt/>) implementado de acordo com as boas práticas do IODE/IOC.

Repositórios de dados: caso de estudo Instituto Hidrográfico

Aspetos de políticas e processos

No Instituto Hidrográfico existe um documento com a política de dados que apresenta as linhas de orientação para as práticas de gestão de dados, criando as bases para a utilização e reutilização de um recurso público, seguindo os princípios de reprodutibilidade, transparência e concorrência no meio científico nacional.

Aspetos técnicos da plataforma

A infraestrutura combina *software* comercial, como o Oracle ou o ArcGIS Server, com aplicações de livre acesso, como o PostgreSQL + Postgis, GeoServer, ncWMS Server e Geonetwork.

Os metadados estão de acordo com a ISO 19115 e os serviços com a ISO 19119.

São utilizados os vocabulários controlados:

OHI: <http://iho-ohi.net/S32/engView.php>

Diretiva Inspire: <http://inspire.ec.europa.eu/registry>

NERC Vocabulary Server: <https://vocab.nerc.ac.uk/>

Agregação em diretórios e certificação

O processo de certificação CoreTrustSeal está planeado, mas ainda não foi iniciado.

Links úteis

[Declaração de Acessibilidade e Usabilidade do Hidrográfico+](#)

[Filme institucional do Instituto Hidrográfico](#)

Contactos

centrodados@hidrografico.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

maio 2023

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>

Arquivo Português de Informação Social



O [Arquivo Português de Informação Social \(APIS\)](http://www.apis.ics.ulisboa.pt), sediado no Instituto de Ciências Sociais (ICS-ULisboa), é uma das componentes do [PASSDA](#) (Production and Archive of Social Science Data), que se constitui como uma infraestrutura de recolha, arquivo e disseminação de dados sobre atitudes, valores e comportamentos sociopolíticos.

Para além disso, o APIS é membro da infraestrutura europeia [CESSDA ERIC](#) (Consortium of European Social Science Data Archives).

Mais informação:
[Sobre o APIS](#)
[Estrutura do APIS](#)

APIS

<http://www.apis.ics.ulisboa.pt/>

A principal atividade do APIS consiste em assegurar a continuidade do acesso aos dados produzidos pelos projetos de investigação, através de um conjunto de atividades de curadoria, preservação e disseminação, promovendo ao mesmo tempo a visibilidade dos dados e a sua utilização por terceiros para efeitos de análise secundária, consulta pública e uso pedagógico.

O APIS incentiva e promove o depósito de dados de investigação com origem em inquéritos produzidos nas áreas da Sociologia, Ciência Política e Psicologia Social. Contudo, qualquer investigador na área das ciências sociais poderá enviar dados para o APIS.

Aspetos de políticas e processos

Relativamente a como depositar os dados no APIS, todas as indicações estão disponíveis para consulta no nosso site (procedimentos para o depósito, como preparar os dados e a documentação, os metadados, respeitando os vocabulários controlados, etc.):

<http://www.apis.ics.ulisboa.pt/depositar/>
<http://www.apis.ics.ulisboa.pt/preparar-os-dados/>
<http://www.apis.ics.ulisboa.pt/metadados/>

Repositórios de dados: caso de estudo Arquivo Português de Informação Social

Aspetos técnicos da plataforma

A nossa coleção de estudos pode ser consultada, não só através do portal [RCAAP](#), mas também na plataforma [nesstar](#).

Infraestrutura técnica e de segurança: <http://www.apis.ics.ulisboa.pt/tecnologia-infraestrutura-tecnica-e-seguranca/>

O APIS usufrui do sistema de atribuição de PIDs (handle) do RCAAP.

Quanto aos vocabulários controlados, o APIS, enquanto membro do [CESSDA](#) (Consortium of European Social Science Data Archives), utiliza, para a publicação dos metadados no Nesstar e também no [CESSDA Data Catalogue](#) (CDC), os vocabulários controlados DDI, tendo inclusivamente participado na sua tradução para português.

Links importantes sobre o CESSDA Vocabulary Service:

<https://ddialliance.org/announcement/cessda-vocabulary-service-tool-now-available>
<https://vocabularies.cessda.eu/#!discover>

Agregação em diretórios e certificação

O APIS está registado no [re3data.org](#) e obteve a certificação CoreTrustSeal em setembro de 2021.

Registo do APIS no [re3data.org](#): <http://doi.org/10.17616/R3INJMUK>

Candidatura ao [CoreTrustSeal](#): https://www.coretrustseal.org/wp-content/uploads/2021/09/20210901-apis-arquivo-portugues-de-informacao-social_final.pdf

Links úteis

Estrutura/organograma do APIS:
<http://www.apis.ics.ulisboa.pt/estrutura-do-apis/>

Condições de acesso aos dados:
<http://www.apis.ics.ulisboa.pt/condicoes-acesso/>

Condições de uso dos dados:
<http://www.apis.ics.ulisboa.pt/condicoes-de-uso/>

Contactos

Pedro.ferreira@ics.ulisboa.pt

Patricia.miranda@ics.ulisboa.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

junho 2023

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>



DUnAs Repositório de Dados de Investigação da Universidade de Aveiro

A Universidade de Aveiro (UA) é uma fundação pública com regime de direito privado que tem como missão a intervenção e desenvolvimento da formação graduada e pós-graduada, a investigação e a cooperação com a sociedade. É composta por 16 departamentos, 4 escolas politécnicas, 20 unidades de investigação e 8 serviços.

O repositório DUnAs é um projeto concretizado pelos Serviços de Biblioteca, Informação Documental e Museologia (SBIDM) e pelos Serviços de Tecnologias de Informação e Comunicação (STIC), em colaboração com o Gabinete de Apoio à Investigação da Universidade de Aveiro.

DUnAs

<https://dunas.ua.pt/>

Considerando a previsão da imposição legal por parte das agências de financiamento no que toca à salvaguarda e publicação dos dados decorrentes da investigação científica, identificou-se a necessidade de disponibilizar na UA uma plataforma que permitisse dotar a instituição de uma ferramenta de apoio à investigação nesse âmbito.

Como requisitos, definiu-se que esta plataforma deveria, tanto quanto possível: espelhar a orgânica da instituição; ser flexível relativamente aos cenários identificados como tipologias de projetos de investigação; permitir a liberdade necessária para a gestão de dados e *datasets* em todas as etapas do seu ciclo de vida (com o foco na minimização do esforço do serviço responsável pelo suporte); assegurar uma qualidade de serviço elevada; fornecer uma experiência de utilização inclusiva e intuitiva; minimizar os custos para a instituição. Esta plataforma deveria ainda basear-se num projeto sólido, com uma comunidade alargada e dinâmica, permitir elevada interoperabilidade entre sistemas (quer externos, quer internos) e ser passível de ser submetida a um processo de certificação para atribuição de um selo de qualidade.

Assim, o DUnAs constitui-se como um repositório destinado ao arquivo, publicação e acesso aos dados de investigação gerados no âmbito das atividades científicas e académicas da UA. Por servir uma instituição que leciona e investiga em múltiplas áreas do conhecimento, o DUnAs é um repositório de carácter institucional e multidisciplinar.

Repositórios de dados: caso de estudo DUnAs

Aspetos de políticas e processos

O depósito de dados no DUnAs é realizado em regime de autoarquivo, pelos investigadores ou membros do projeto designados, mediante a atribuição de permissões pelos SBIDM. Um dos critérios de depósito de dados no DUnAs é que os mesmos tenham sido gerados no âmbito de uma unidade ou projeto de investigação da UA e que pelo menos um dos autores do *dataset* tenha afiliação à instituição.

A publicação de *datasets* implica uma submissão prévia para revisão por parte dos SBIDM quanto à qualidade e quantidade de metadados, sendo que, caso não sejam detetadas inconformidades, o *dataset* é publicado pelo serviço. Caso haja alterações ou melhorias a efetuar, o *dataset* é devolvido ao depositante.

A permissão de publicação de *datasets* é exclusiva dos SBIDM.

A veracidade e qualidade dos dados, assim como as questões de proteção de dados pessoais e informação confidencial e de direitos de autor é da total responsabilidade dos depositantes.

Aspetos técnicos da plataforma

Software Dataverse v5.11.1 customizado ao nível do fluxo de publicação.

Customização da interface gráfica.

Configuração das métricas de utilização do repositório e de citação de *datasets*.

Atribuição de identificador único (DOI) a cada *dataset* e ficheiro.

Disponibilização de esquemas de metadados genéricos e disciplinares.

Possibilidade de associação das licenças CC0 e CC BY ou definição de termos de uso personalizados.

Agregação em diretórios e certificação

O repositório encontra-se registado no re3data.org: <http://doi.org/10.17616/R3INJNAA>

Links úteis

[Página web sobre o serviço DUnAs](#)

[Guia temático do DUnAs](#)

[Termos gerais de uso do DUnAs](#)

[Política de depósito no DUnAs](#)

Contactos

Universidade de Aveiro
Campus Universitário de Santiago
3810-193 Aveiro, Portugal
sbidm-dunas@ua.pt
(+351) 234 247 149

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

setembro 2023

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>



DataRepositóriUM

A Universidade do Minho está, atualmente, entre as mais prestigiadas instituições de ensino superior do país, tendo também vindo a afirmar-se progressivamente no panorama internacional.

Tem como um dos seus objetivos estratégicos, consolidar a sua posição como instituição de referência no domínio da Ciência Aberta e dos repositórios científicos, adotando práticas de Ciência Aberta e reforçando a participação em iniciativas nacionais e internacionais.

A Unidade de Serviços de Documentação e Bibliotecas, através do Gabinete de Gestão de Informação Científica, Repositórios e Ciência Aberta, tem como um dos seus objetivos operacionais a consolidação do serviço DataRepositóriUM aumentando a sua utilização.

DataRepositóriUM

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

O DataRepositóriUM é o Repositório de Dados Institucional para partilhar, publicar e gerir dados de investigação gerados e recolhidos pela atividade dos investigadores e nas unidades de investigação da Universidade do Minho, apoiado pela Unidade de Serviços de Documentação e Bibliotecas.

O repositório foi lançado em fevereiro de 2020, possuindo Termos de Utilização próprios e tendo como referência as Normas Gerais de utilização da Comunidade Dataverse. Aceita para publicação dados finais de investigação multidisciplinares de todos os tipos e formatos, com preferência para formatos de dados abertos.

Tem como objetivos: a) Depositar ficheiros em *datasets* e coleções organizadas, b) Preservar ficheiros e *datasets*, c) Apoiar o cumprimento dos requisitos de financiadores, d) Registrar identificadores persistentes (DOI), e) Permitir o acesso controlado e sob identificação prévia, f) Associar licenças de reuso e explicitar os termos de uso, g) Documentar os *datasets* com metadados disciplinares, h) Facilitar a disseminação dos resultados de investigação, i) Viabilizar a citação de dados e o crédito aos produtores, e j) Facilitar a disponibilização de dados aos editores.

A missão e objetivos estão alinhados com as melhores práticas para o desenvolvimento de um repositório de dados sustentável, com um armazenamento seguro, preservação e acesso sustentados, fiável e aberto aos dados de investigação, de acordo com as orientações nacionais e internacionais e os princípios FAIR para dados de investigação.

Repositórios de dados: caso de estudo DataRepositóriUM

Aspetos de políticas e processos

O processo de publicação de *datasets* no DataRepositóriUM está definido pelo seu *workflow* de curadoria e termos de uso. A publicação pode ser feita de duas formas, definidas por níveis de permissão distintos concedidos aos utilizadores: publicação sem processo de revisão prévia (publicação em auto-depósito) e publicação após um processo de revisão por um curador da coleção (publicação com depósito mediado). Em ambos os casos, a equipa de suporte valida os novos *datasets* depositados.

O repositório está alinhado com a Política Institucional de Ciência Aberta e a Política de Preservação, sendo promovida a adoção de boas práticas para o depósito e publicação de dados de acordo com os princípios FAIR, através da disponibilização de materiais de apoio e ações de suporte e formação.

Aspetos técnicos da plataforma

O repositório é desenvolvido com base no *software* de código aberto Dataverse. Faz a atribuição de identificadores persistentes DOI (Digital Object Identifier) da agência DataCite, através do Serviço Nacional de Registo de Identificadores DOI da FCT-FCCN. É compatível com as diretrizes de interoperabilidade de metadados do OpenAIRE.

Agregação em diretórios e certificação

O DataRepositóriUM é um serviço institucional que está registado e aceite nos principais diretórios mundiais ou bases de dados europeias de referência, nomeadamente Re3data, FAIRsharing e OpenAIRE. Em junho de 2023 foi feita a submissão para obtenção da certificação CoreTrustSeal.

Links úteis

[Sobre](#)
[Termos de Uso](#)
[Apoio e formação](#)
[Workflow de curadoria](#)

Contactos

Universidade do Minho
Unidade de Serviços de Documentação
e Bibliotecas (USDB)
Campus de Gualtar, 4710-057 Braga, Portugal
datarepositorium@usdb.uminho.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



[Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

novembro 2023

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>



BioData.pt Data Management Portal

BioData.pt é a infraestrutura distribuída portuguesa para dados biológicos e o nó português do ELIXIR.

BioData.pt Data Management Portal

<https://dmportal.biodata.pt/>

O BioData.pt Data Management Portal surge da necessidade de garantir que os dados de investigação são publicados de forma FAIR (Findable, Accessible, Interoperable and Reusable).

O portal está aberto a todos os domínios de investigação nas áreas das Ciências da Vida e da Saúde.

Aspetos de políticas e processos

É dada permissão para depósito e curadoria a DataVerses (pastas) específicos de cada área científica após pedido por parte de um investigador. Não é feita curadoria ativa por parte da equipa do BioData.pt.

Os termos de serviço estão em desenvolvimento.

Aspetos técnicos da plataforma

Constitui um repositório digital baseado na aplicação web de código aberto Dataverse.

Permite a atribuição de um DOI, através do DataCite, aos *datasets* publicados.

Foram inseridos os vocabulários de metadados de base, nomeadamente "Citation metadata", "Geospatial metadata", "Life Sciences metadata", "Journal Metadata" e o vocabulário MIAPPE (Plantas).

Repositórios de dados: caso de estudo BioData.pt Data Management Portal

Agregação em diretórios e certificação

Repositório indexado em [FAIRsharing.org](https://fairsharing.org/):
<https://doi.org/10.25504/FAIRsharing.b2f583>.

Links úteis

Formulário de suporte disponível no *website*
[Dataverse Project User Guide](#)
[BioData.pt | ELIXIR PT Ethics and Legal Policy](#)

Contactos

dmportal@biodata.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

setembro 2024

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgd.i-rcaap.pt/grupos-de-trabalho>



GBIF.PT

O GBIF — Sistema Global de Informação sobre Biodiversidade — é uma rede internacional financiada por governos de todo o mundo, com o objetivo de dar acesso aberto a dados sobre toda a vida na Terra. Criado em 2001, o GBIF estabeleceu uma rede de países e organizações participantes que trabalham através de Nós Participantes e Afiliados. Fornece às instituições detentoras de dados em todo o mundo padrões comuns e ferramentas de código aberto que lhes permitem partilhar dados sobre a ocorrência de espécies.

O [Nó Português do GBIF \(GBIF.PT\)](https://dados.gbif.pt/) tem por missão promover a integração dos provedores de dados portugueses e recursos de informação sobre biodiversidade na rede GBIF, e a disponibilidade dos dados de biodiversidade para a investigação científica e usos pela sociedade.

Portal de Dados de Biodiversidade de Portugal

<https://dados.gbif.pt/>

Os dados mobilizados e publicados através do GBIF derivam de muitas fontes, incluindo espécimes de museus, alguns recolhidos há mais de um século, observações de inventários de campo, avaliações de impacto ambiental, registos fornecidos por ciência cidadã e amostras de ADN ambiental. O dados são usados em estudos dedicados aos impactos das alterações climáticas na biodiversidade, a propagação de espécies invasoras ou as prioridades para a segurança alimentar, a saúde pública, a conservação e a designação de áreas protegidas. Muitas destas análises não seriam possíveis sem o acesso aberto a dados abrangentes sobre biodiversidade.

Aspetos de políticas e processos

As instituições que desejam publicar dados de biodiversidade através do GBIF precisam de se registar e receber o endosso de um nó Participante. Durante o registo, os publicadores concordam, através do acordo de publicação de dados, em disponibilizar os seus dados universalmente, mantendo os seus direitos de propriedade intelectual.

Os conjuntos de dados devem adotar uma das seguintes licenças padrão: Creative Commons Zero (CC0), Creative Commons Attribution (CC-BY) ou Creative Commons Attribution Noncommercial (CC-BY-NC), embora a adoção de CC0 ou CC-BY seja altamente recomendada.

Os nós podem, adicionalmente, fornecer infraestruturas, formação e outros serviços para facilitar a publicação de dados, e ainda a coordenação nacional.

Repositórios de dados: caso de estudo

GBIF.PT

Aspetos técnicos da plataforma

O fluxo de dados envolve os seguintes passos: 1 - Registo do publicador; 2 - Mobilização dos dados; 3 - Seleção de padrão de dados (Darwin Core) e ferramentas (IPT); 4 - Preparação dos dados; 5 - Publicação dos dados; 6 - Descoberta dos dados.

A transformação de dados no padrão Darwin Core é uma etapa crítica para garantir a qualidade dos dados. Alguns dos termos do padrão de dados são obrigatórios, garantindo identificadores únicos e persistentes. É ainda possível o uso de extensões Darwin Core para suporte de tipologias de dados específicas, como, por exemplo, dados sobre ADN, características de espécies, dados em ficheiros multimédia.

A publicação de dados é feita através da ferramenta [Integrated Publishing Toolkit \(IPT\)](#). Esta ferramenta de código aberto permite fazer carregamento de dados e preencher os metadados no formato padrão Ecological Metadata Language (EML). No momento da publicação, a cada conjunto de dados é associado um DOI. O IPT é o ponto de acesso para o portal global GBIF.org realizar a ingestão dos dados, e sua integração na plataforma global.

O acesso aos dados faz-se através do portal global GBIF.org. A cada descarregamento de dados é associada uma citação recomendada, incluindo um DOI. O acesso pode ser feito também através de *software* (ex.: python, R). O Nó Português do GBIF implementou o [portal nacional](#), que acrescenta serviços ao portal global, para o contexto nacional. O portal nacional é baseado no Atlas of Living Australia (ALA), uma plataforma de dados de código aberto desenvolvida pelo nó australiano do GBIF e que é mantido no âmbito da comunidade Living Atlases.

Agregação em diretórios e certificação

O IPT do Nó Português do GBIF está registado enquanto repositório FAIRSharing. Os portais de dados GBIF.org e dados.gbif.pt estão ambos registados no EOSC Catalogue & Marketplace.

Links úteis

[Sobre](#)
[Acordo de utilização de dados](#)
[Acordo de publicação de dados](#)
[Formação](#)
[Infraestrutura](#)

Contactos

Nó Português do GBIF
Instituto Superior de Agronomia, Herbário
Tapada da Ajuda
1349-017 Lisboa, Portugal
node@gbif.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



[Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

setembro 2024

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.ccaap.pt/grupos-de-trabalho>



COI Catalogue Herbário da Universidade de Coimbra

O [Herbário da Universidade de Coimbra](#) é o maior herbário português, com cerca de 800.000 exemplares de plantas, fungos, algas e líquenes.

COI Catalogue <https://coicatalogue.uc.pt>

O espólio das coleções biológicas contém informação valiosa sobre biodiversidade. Durante séculos essa informação esteve pouco acessível, mas nas últimas décadas a informatização permitiu a sistematização desta informação e a sua divulgação de forma estruturada através da *internet*. O portal de dados do Herbário da Universidade de Coimbra disponibiliza a informação dos exemplares que se encontram digitalizados.

Aspetos de políticas e processos

Os dados são produzidos internamente como resultado do processo de digitalização das coleções. A curadoria é contínua e mantida pelo *staff*.

A disponibilização *online* é sincronizada diariamente a partir do sistema de gestão de coleções.

Aspetos técnicos da plataforma

A gestão de dados é feita com o *software* de gestão de coleções SPECIFY 6.

A maior parte da informação corresponde aos termos definidos no padrão DarwinCore, mas também existe informação em campos para os quais não há nenhum padrão.

O portal de dados foi desenvolvido localmente, tendo em conta vários aspectos UI/UX relevantes para os investigadores.

Repositórios de dados: caso de estudo COI Catalogue

A atribuição de identificadores únicos a cada registo está de acordo com as boas práticas da disciplina (CETAF Stable Identifiers).

Os dados podem ser exportados no formato CSV. São também disponibilizadas imagens de alta qualidade dos exemplares de herbário, com visualização lado a lado com a ficha de cada exemplar e funcionalidade de zoom.

Agregação em diretórios e certificação

Os dados são sincronizados mensalmente no GBIF, o maior agregador de dados de Biodiversidade: <https://www.gbif.org/dataset/a559e942-0fbe-4f09-93ca-28cf244ce2a0>.

Links úteis

[Filme institucional do Herbário da Universidade de Coimbra](#)

[EXPLORATOR: Plataforma de Ciência-Cidadã para transcrição/classificação de exemplares](#)

Contactos

coi@bot.uc.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

setembro 2024

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>



Dados.IPB

O Instituto Politécnico de Bragança (IPB) é uma Instituição de Ensino Superior. Integra seis escolas e oferece CTeSP, licenciaturas, mestrados e doutoramentos, pós-graduações e cursos de atualização e requalificação nos principais domínios.

O Dados.IPB foi operacionalizado pelos Serviços de Documentação e Bibliotecas do IPB e tem como objetivos armazenar, preservar, partilhar e gerir os dados de investigação gerados pela comunidade científica, aumentando assim a visibilidade dos dados de investigação do IPB.

Dados.IPB

<https://dados.ipb.pt/>

Em 2019, através dos Serviços de Documentação e Bibliotecas do IPB, foi disponibilizado o Dados.IPB para toda a comunidade académica.

O Dados.IPB é o repositório de dados de investigação institucional, que serve para armazenar, conservar e disponibilizar os dados de investigação gerados pelos centros de investigação do IPB e também pelos laboratórios colaborativos. Este repositório é apoiado pelos Serviços de Documentação e Bibliotecas do IPB.

Aquando da conceção do Dados.IPB, criaram-se *dataverses* (comunidades) seguindo a orgânica do IPB, ou seja, escolas e centros de Investigação e também os laboratórios colaborativos.

Possui as Normas da Comunidade Dados.IPB, os Termos Gerais de Uso e uma Política de Preservação do Dados.IPB.

Os dados aceites são multidisciplinares, com diferentes tipologias e formatos. O acesso é tão aberto quanto possível, tão fechado quanto necessário.

Os objetivos são: depositar ficheiros (*datasets*) em comunidades; preservar os *datasets* e os metadados; fazer cumprir os requisitos dos financiadores; utilizar identificadores persistentes como o DOI; facilitar a disponibilização de dados aos editores; permitir o acesso controlado a *datasets* que estejam em acesso restrito.

O Dados.IPB está alinhado com a Política de Ciência Aberta e Cidadã do IPB, a Política de Preservação e os princípios FAIR, e dá integridade à investigação efetuada no IPB, permitindo assim o desenvolvimento do conhecimento científico, reforçando e estimulando o conceito de Investigação e Inovação Responsável, para que seja possível alavancar a inovação e a reprodutibilidade da ciência.

Repositórios de dados: caso de estudo Dados.IPB

Aspetos de políticas e processos

O processo de publicação de *datasets* no Dados.IPB está definido pelo *workflow* de curadoria e termos de uso. A publicação é feita de duas formas, definidas por níveis de permissão distintos atribuídos aos utilizadores: publicação sem processo de revisão prévia (autoarquivo) e publicação após um processo de revisão efetuado por um curador. Em ambos os casos, a equipa dos Serviços de Documentação e Bibliotecas valida os novos *datasets* depositados.

O Dados.IPB está alinhado com a nova Política de Ciência Aberta e Cidadã do Instituto Politécnico de Bragança e a Política de Preservação, sendo promovida a adoção de boas práticas para o depósito e publicação de dados de acordo com os princípios FAIR, através do *e-mail* e formação pontual disponibilizada aos centros de investigação.

Aspetos técnicos da plataforma

O Dados.IPB é desenvolvido com o *software* de código aberto Dataverse. Atribui o DOI (Digital Object Identifier) da DataCite, através da FCT-FCCN. É compatível com as diretrizes de interoperabilidade de metadados do OpenAIRE e, portanto, possui API. Foi instalado o vocabulário controlado Life Science Metadata.

Agregação em diretórios e certificação

O Dados.IPB está registado nos principais diretórios mundiais, como o Re3data, o FAIRsharing e o OpenAIRE. Em junho de 2023 foi feita a submissão para obtenção da certificação CoreTrustSeal.

Links úteis

[Termos de uso](#)
[Normas da comunidade](#)
[Política de Preservação do Dados.IPB](#)
[Workflow dos dados](#)
[Guia de adesão ao Dados.IPB](#)
[Guia de curadoria do Dados.IPB](#)

Contactos

Serviços de Documentação
e Bibliotecas do IPB
Campus de Santa Apolónia
5300-253 Bragança, Portugal
dados@ipb.pt / clarisse@ipb.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

setembro 2024

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>



INESC TEC Research Data Repository (RDM INESC TEC)

O INESC TEC é uma associação privada sem fins lucrativos, com estatuto de utilidade pública, que se centra em atividades de investigação científica e desenvolvimento tecnológico, transferência de tecnologia, consultoria avançada e formação, e pré-incubação de novas empresas de base tecnológica.

Como instituição de interface, o INESC TEC junta a academia, as empresas, a administração pública e a sociedade.

INESC TEC Research Data Repository

<https://rdm.inesctec.pt/>

O RDM INESC TEC representa o compromisso institucional com os Dados Abertos. O INESC TEC foi rápido a adotar os princípios FAIR e o repositório foi lançado em 2017. Neste contexto, o INESC TEC apostou num serviço de apoio aos investigadores e o RDM suporta as atividades de gestão de dados de investigação como, por exemplo, no desenvolvimento de Planos de Gestão de Dados.

O repositório disponibiliza *datasets* relacionados com as linhas temáticas de investigação da instituição. São exemplo disso, a disponibilização de dados nos domínios da Inteligência Artificial, Bioengenharia, Robótica, Sistemas de Energia, Ciência da Computação e Engenharia no geral.

Aspetos de políticas e processos

O repositório disponibiliza *datasets* produzidos por investigadores do INESC TEC ou dos parceiros em projetos de investigação. O repositório é de acesso aberto e não é necessário registo para utilizar.

Todos os depósitos são mediados pelo *data steward* institucional e não podem ser feitos por outros utilizadores. *Datasets* com dados sensíveis requerem a intervenção do DPO (*Data Protection Officer*) e podem levar a medidas concretas de acordo com os requisitos do *datasets*.

Repositórios de dados: caso de estudo INESC TEC Research Data Repository

Aspetos técnicos da plataforma

A *software* é o CKAN e o DOI é obtido através do serviço DataCite Fabrica.

Um *template* elaborado pelo *data steward* é usado para recolher metadados. Há um conjunto mínimo de metadados a recolher para cada *dataset*, sendo que podem ser preenchidos campos opcionais (dependendo do contexto do *dataset*) e metadados mais específicos podem ser adicionados (*custom*). A edição de metadados que enriqueçam o *dataset* podem ser preenchidos após o depósito.

São feitos *backups* diários, com cópias adicionais periódicas em *tape*.

A licença atribuída, salvo casos excecionais, é a Creative Commons Attribution Share-Alike 4.0.

Não há limite de tamanho para o *dataset*, mas existe um limite de 4GB por ficheiro.

Se o *dataset* for de grande dimensão (TBs), opta-se por disponibilizar o *dataset* numa máquina virtual (open-datasets INESC TEC) acedida através do RDM INESC TEC (link como recurso).

Agregação em diretórios e certificação

O RDM INESC TEC está registado no re3data.org e fairsharing.org.

Links úteis

[Website do INESC TEC](#)

[Registo no re3data](#)

[Registo no FAIRsharing](#)

Contactos

joao.a.castro@inesctec.pt

ag@inesctec.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



Creative Commons Attribution 4.0 International License

novembro 2024

Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>

Repositório Iscte Partilha de Dados de Investigação

O Iscte - Instituto Universitário de Lisboa é uma instituição pública de ensino superior, fundada em 1972, que se destaca pela sua forte componente de investigação e inovação. Oferece uma ampla gama de cursos de licenciatura, mestrado e doutoramento nas áreas de ciências sociais, gestão, tecnologia e arquitetura.

O contributo que o Iscte pode dar à sociedade consiste em desenvolver, com elevados padrões de qualidade, a sua missão nestes três domínios: o ensino, em especial nos níveis pós-graduados, a investigação científica e a transferência de conhecimento para a sociedade. Desta definição da missão do Iscte enquanto instituto universitário decorrem as linhas de desenvolvimento estratégico fixadas no Plano de Estratégico.

Repositório Iscte - Partilha de Dados de Investigação

<https://repositorio.iscte-iul.pt/>

O repositório institucional do Iscte tem como objetivo preservar, divulgar e dar acesso à sua produção intelectual em formato digital. Ao reunir as publicações académicas e científicas, contribui para o aumento da sua visibilidade e impacto do trabalho de investigação a nível nacional e internacional. No final de 2022, foi criada uma comunidade com o objetivo de produzir um [catálogo institucional de registo dos dados de investigação](#) depositados.

Repositórios de dados: caso de estudo Repositório Iscte - Partilha de Dados de Investigação

Aspetos de políticas e processos

O Iscte publicou a sua [Política de Gestão e Partilha de Dados de Investigação](#) em novembro de 2023, na qual é requerido o depósito, na [comunidade Iscte do Repositório Zenodo](#), dos dados de investigação necessários à validação dos resultados de investigação científica produzida no Iscte.

Assim, o [fluxo de depósito](#) de dados de investigação no Iscte usa o repositório generalista Zenodo e as infraestruturas institucionais existentes (Ciência-IUL e Repositório Iscte), de forma a dar resposta aos requisitos dos financiadores de ciência e a garantir o acesso, a partilha, o armazenamento, a preservação e a reprodutibilidade dos dados gerados no âmbito dos projetos de I&D.

Aspetos técnicos da plataforma

O repositório é desenvolvido com base no *software* de código aberto DSpace. É compatível com as diretrizes de interoperabilidade de metadados do OpenAIRE. Faz a atribuição de identificadores persistentes Handle.

Agregação em diretórios e certificação

O Repositório Iscte está registado no Indexar, no FAIRsharing (através da Política de Gestão de Dados) e no OpenAIRE.

Links úteis

[Política de Gestão e Partilha de Dados de Investigação](#)
[Fluxo de depósito](#)
[Guia Online Gestão de Dados de Investigação](#)

Contactos

Iscte – Instituto Universitário de Lisboa
Serviços de Informação e Documentação – Biblioteca
administrador.repositorio@iscte-iul.pt

Recurso criado no âmbito do Grupo de Trabalho do Fórum de Gestão de Dados de Investigação 'Repositórios de Dados: Tecnologia, organização e certificação'



[Creative Commons Attribution 4.0 International License](#)

novembro 2024

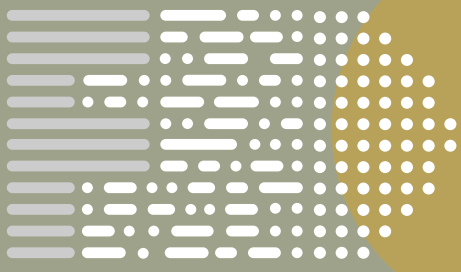
Grupo de Trabalho
Repositórios de Dados:
Tecnologia, organização e
certificação FAIR



Fórum de Gestão de Dados de Investigação

Grupo de Trabalho 'Repositórios de Dados:
Tecnologia, organização e certificação'

<https://forumgdi.rcaap.pt/grupos-de-trabalho>



FÓRUM
GESTÃO DE DADOS
DE INVESTIGAÇÃO

Repositórios de Dados
CASOS DE ESTUDO

A Instituição onde trabalha tem algum repositório de dados?

Entre em contacto connosco para se juntar a esta comunidade.

Contacto: forumgdi-repositorios@fccn.pt

Grupo de trabalho
"Repositórios de Dados"



Comunidade Zenodo
"Fórum GDI"

