



# Making Research Data Findable with B2FIND

Anna-Lena Flügel (DKRZ/EUDAT)

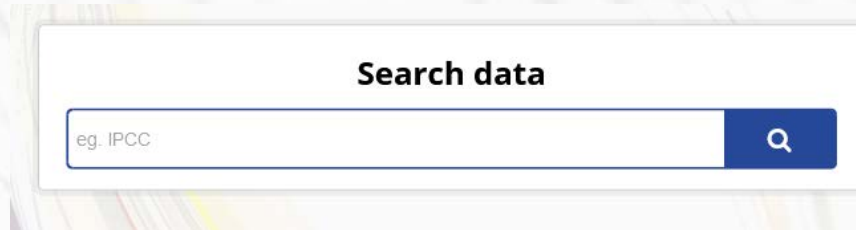
NFDi4ing Conference, 27.10.2022

CC-BY-4.0




The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207

# B2FIND Overview



Search data

<http://b2find.eudat.eu/>

- simple and user-friendly **discovery portal** for **research data**, allows text- and faceted search **across scientific disciplines**
- up to now **62** repositories on productive machine, several more on test-instance; **>1 Mio metadata records of research datasets**
- enables findability of data stored in **EUDAT data centers**, other **research infrastructures** or **data providers**, covering divergent research areas
- was the central indexing tool for **EOSC-hub**; registered discovery service in **EOSC**, currently part of **DICE** project

# Project Summary



**Consortium: 24 Partners**  
**Coordinator: CINECA**



**Grant Agreement: 101017207**  
**Call: INFRAEOSC-07-2020**



**30 months**  
*(1 Jan 2021 – 30 Jun 2023)*



**Enable a European storage  
and data management infrastructure for EOSC**  
**Provide generic services to store, find, access,  
and process data in a consistent and persistent way**



**Budget: € 6.997.706,66**



**18 providers from 11 European countries are offering  
14 state-of-the-art data management services  
together with more than 50 PB of storage capacity**

- ▶ **14 services** covering the whole research data management lifecycle
- ▶ **>50 PB** of storage offered with the services
- ▶ Standard requests to be served within **2 weeks**
- ▶ Support can be provided for customised solutions
- ▶ Request the services from the EOSC marketplace [HERE](#)
- ▶ Ask for further information by email [HERE](#)



## Data Discovery

Making research data findable  
Data source harvestable and to make data widely discoverable



## Data Repository

Long-term data resources preservation and sharing  
To store non-active FAIR data



## Policies based Data Archive

Mid-long term data resources, accessible from computing facilities to store non-active research data  
Value-added services (integrity checks, replications, ...) for long-term, stable data archives.



## Data Archive

Mid term storage resources after/between projects, Accessible from computing facilities to store non-active research data, bitwise preservation



## Personal and/or Project workspace

Mid-term storage during project,  
Good connection to computing facilities,  
To store active resource data

# Metadata Ingestion

## supported channels

OAI – PMH  
CSW  
Rest API  
[SparQL]

## Harvesting

## supported metadata

Dublin Core  
Datacite  
OpenAire  
**EUDAT Core**  
ISO 19139/19139, FGDC  
community specific



## Search

free text search  
faceted search  
e.g. spatial/temporal  
coverage  
**narrow down results**

## Mapping

**technical** format conversion  
**semantic** „readers“ for standards, mapping  
to target schema; including  
vocabularies, validation  
**specific** according to community needs

# B2FIND as Catalogue and Aggregator

## Ingestion of Metadata

- from EUDAT's B2SHARE
- from research communities worldwide across all academic disciplines



## Exposure of Metadata

- OAI-PMH extension for CKAN
- B2FIND is a metadata provider for OpenAIRE and EOSC



## Useful B2FIND Features for Data Searchers

- \* Search across disciplines in one easy-to-use portal
- \* Easy access to source dataset (if openly accessible) and/or its original metadata
- \* Filter options:
  - \* spatial coverage (easily selectable on a map)
  - \* temporal coverage
  - \* measuring instrument, eg. beamline, float sensor
  - \* search by research discipline

## Useful B2FIND Features for Data Providers

DATA CATALOGUE

REPOSITORIES

COMMUNITIES

FOR PROVIDERS

FOR USERS

ABOUT

- \* Possibility to display single repositories as well as overarching projects or research infrastructures („Communities“), e.g. PaNOSC
- \* Could be used to display NFDi4ing as a Community, with belonging individual data repositories under its umbrella
- \* Useful feature for creating Community-wide search spaces
- \* Metadata record links back to original source and full metadata



Home / Communities / PaNOSC

**PaNOSC**  
The PaNOSC project, Photon and Neutron Open Science Cloud, brings together six strategic European research infrastructures (ESRF, CERIC-ERIC, ELI Delivery Consortium, the...)

Followers: **0** | Datasets: **2.7k**

Spatial Coverage:

Temporal Coverage:

Publication Year:

Repositories:

Filter:  9-1

- ESRF 2332
- ESS 456
- CERIC 7

Communities:

Filter:  9-1

PaNOSC 2785

Datasets About

Search datasets...

**2,795 datasets found** Order by: Relevance

**Degradation of green historical pigment Verdigris**  
Preliminary study in order to establish the suitability of the instrumental method

**2022\_SciComPty\_paper**  
Data and metadata used in a 2022 manuscript on the SciComPty ptychography software framework

**manuscript\_dataset\_sample\_prep\_XRF2022**  
FAIR and Open dataset for the manuscript...

**Crystallisation lab-Marta**  
20195588-2

**PyPore3D**  
FAIR and Open data repository for the raw data of the manuscript : PyPore3D: An open source software tool for imaging data processing and analysis of porous and multiphase...

**manuscript\_dataset\_for\_Compressive\_Sensing\_for\_Dynamic\_XRF\_Scanning**  
OpenFAIR Dataset of Synchrotron Low Energy XRF and STXM files used in a manuscript on "Compressive Sensing for Dynamic XRF Scanning" Synchrotron Low Energy XRF and STXM Dataset...

**COVID-19 Multiscale Imaging of COVID-19 Lung Injury Patterns Bridges the Gap ...**  
This is an addition to Fast-track Proposal #20202210 (SISSIAFM), please process the proposals combined. COVID-19 related project in conjunction with structural lung...

**Sample Data from V20**  
This data was collected as part of BrightnESS, funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant...

**Sample Data from V20**  
This data was collected as part of BrightnESS, funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant...

Home / Communities / Nordic Archaeology

**Nordic Archaeology**  
This Community consists of records from Danish and Norwegian Sites and Monuments Data Providers. Askeladden is a Norwegian database system for managing cultural heritage...

Followers: **0** | Datasets: **398.1k**

Spatial Coverage:

Temporal Coverage:

Publication Year:

Repositories:

Filter:  9-1

- SLKS 206278
- Askeladden 181636

Communities:

Filter:  9-1

Nordic Archaeology 398114

Datasets About

Search datasets...

**398,114 datasets found** Order by: Relevance

**130812-354 Viskum Høje**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: Mulig del af jysk...

**090702-2 Tante Dorthes have**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: Skår af glaseret...

**010510-57 Kassemose**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: Fund af...

**010510-56 Kassemose**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: to...

**090704-198 St. Rise**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: Grube med skår af...

**200103-271 Fuglebæk**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: bjergartsekse...

**010109-218 Hyrdehøj**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: Fund af tre...

**160406-277 Addit**  
This record describes ancient sites and monuments as well archaeological excavations undertaken by Danish museums. Excerpt of the Danish description of events: 2020-03-05:...

Let's take a look!

\* <https://b2find.eudat.eu>



\* Do you have any questions?





## JOIN OUR COMMUNITY



DICEosc



/company/diceosc

# More Information

## **B2FIND**

<https://b2find.eudat.eu>

## **B2FIND Guidelines for Data Providers**

<https://b2find.eudat.eu/site/forproviders/>

## **B2FIND in GitHub**

<https://github.com/EUDAT-B2FIND/md-ingestion>

## **Contact for Metadata Integration in B2FIND**

DICE

<https://www.dice-eosc.eu/contact>

EUDAT RT

<https://eudat.eu/contact-support-request>