

Research Data Management at the University of Porto using Dendro, CKAN and EUDAT

Cristina Ribeiro, João Rocha da Silva, João Aguiar Castro, Yulia Karimova,
Eugénia Matos Fernandes

INESC TEC—Faculdade de Engenharia da Universidade do Porto



TAIL Team



Cristina
Ribeiro



João Rocha
da Silva



João Aguiar
Castro



Ricardo
Amorim



Yulia
Karimova



Carla
Lopes



Artur
Rocha



João
C. Lopes



Gabriel
David



Tito
Vieira



Eugénia
Fernandes



Ângela
Lomba



João
Honrado



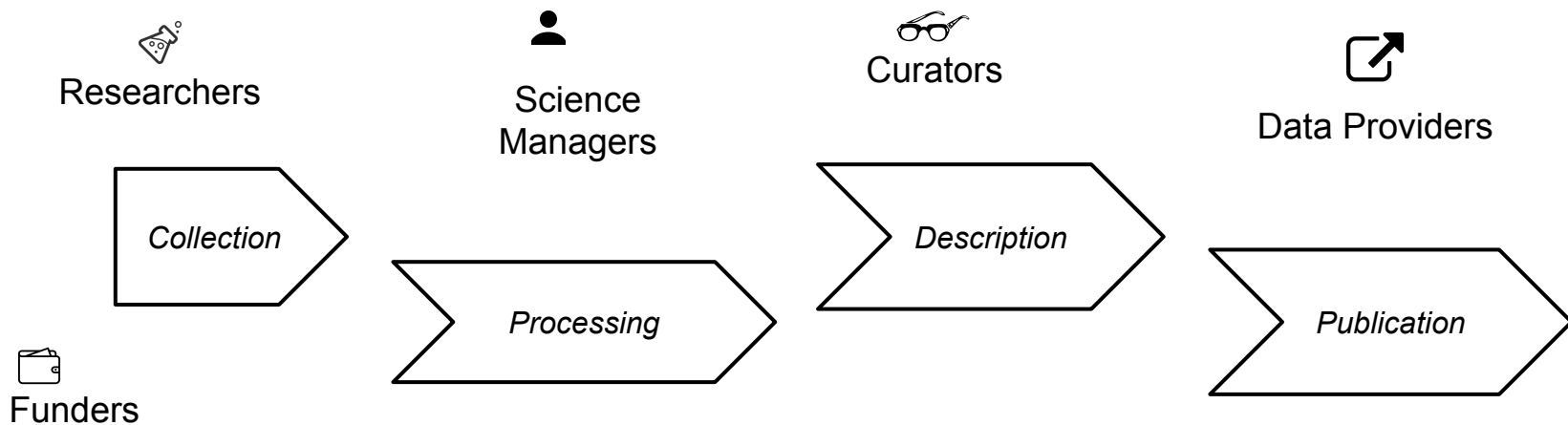
Contents

- ❑ Research data management
- ❑ Researchers, tasks and workflows
- ❑ Dendro: ontology-based organisation & description
- ❑ LabTablet: data collection on the spot
- ❑ Metadata and curator workflows in the long tail
- ❑ EUDAT pilot: *DataPublication@U.Porto*
- ❑ Long-term repositories: generic and disciplinary



Research Data Management

- Motivation:
 - Data management required by funding agencies
 - Data publication: credit to researchers
 - Data have value: experiments are expensive
- Challenges:
 - Data include complex objects
 - Publication process and venues are not clear yet
 - Researchers not familiar with data description



Developers

dendro.fe.up.pt



Institutions

RDM approach at U.Porto

- Focus on Researchers
 - They generate requirements
 - They know data and create descriptions
- Focus on Tools
 - For curators
 - For researchers
- Processes for researchers and curators

Dendro

Folder



Selection



- Up to biodata
- meta
- scala_2014.jpg
- observations.xls
- ontTraffic.pdf

Information

Change log

Description progress

0%

✓ SAVE

↶ UNDO

✓ COPY FROM PARENT

ⓘ IN MANUAL MODE

🗑️ CLEAR



Description

B

I

H



🔍 PREVIEW

Studies of the endangered species of felid living mainly in the Iberian Peninsula in southwestern Europe. They are categorized as endangered by many institutions, including the International Union for the Conservation of Nature



Coverage



Descriptor Selection

ALL

AUTO

Find descriptors

Type to find descriptors...

🔍 DEPICTION



A depiction of some thing. 🧑🧑

🔍 DESCRIPTION



An account of the resource. 🧑🧑

🔍 DATE



A point or period of time associated with an event in the lifecycle of the resource. 🧑🧑

🔍 COVERAGE



The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant. 🧑🧑

🔍 SUBJECT

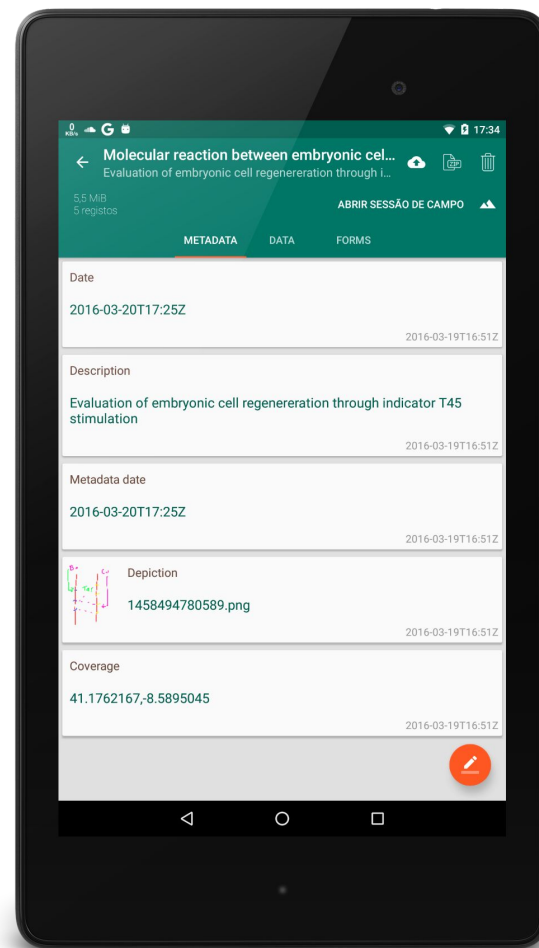


Dendro platform

- Data organisation and description
 - Early: from data creation
 - Detailed, domain-specific description
- Ontologies: loaded for each domain
 - Metadata is ready to become Linked Open Data
 - Preservation-friendly metadata models
- Integration with long-term repositories

LabTablet

- “Automatic” metadata
- Metadata from sensors
- Ontologies: from Dendro
- Annotations in lab notebooks: metadata



Dendro + LabTablet

Folder



Selection



- Up to biodata
- meta
- scala_2014.jpg
- observations.xls
- ontTraffic.pdf

Information

Change log

Description progress

0%

✓ SAVE

↶ UNDO

✓ COPY FROM PARENT

ℹ IN MANUAL MODE



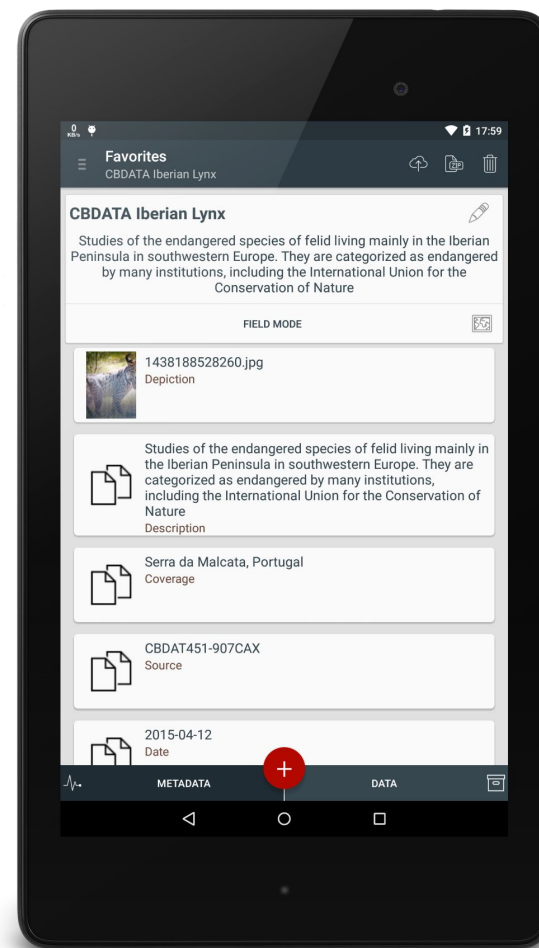
Description

B I H PREVIEW

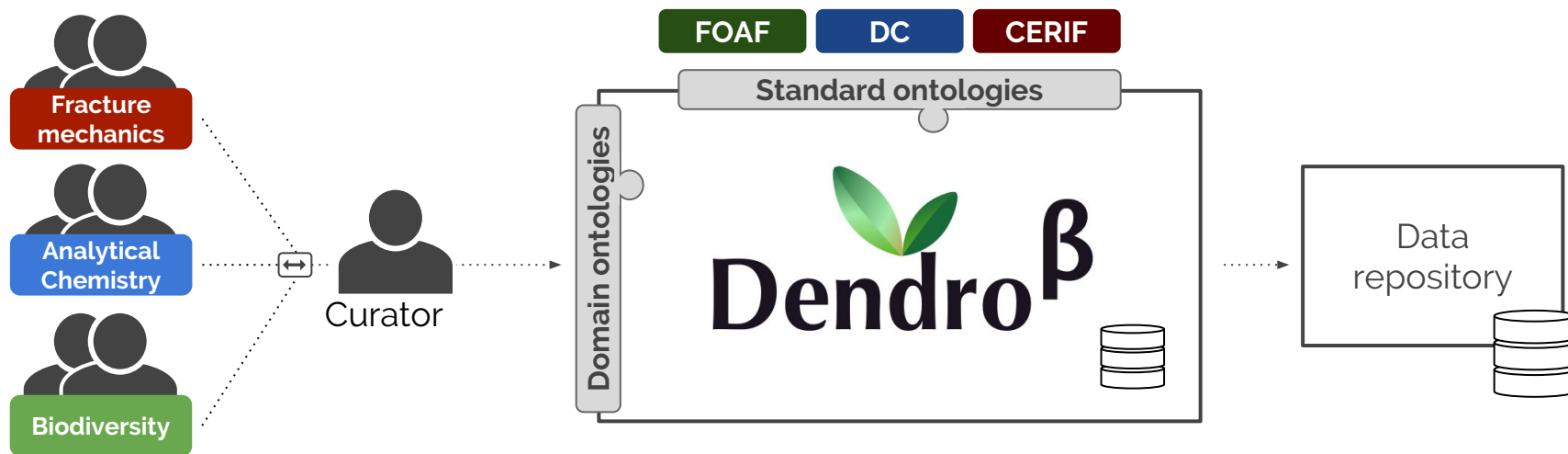
Studies of the endangered species of felid living mainly in the Iberian Peninsula in southwestern Europe. They are categorized as endangered by many institutions, including the International Union for the Conservation of Nature



Coverage



Generic and domain-specific ontologies

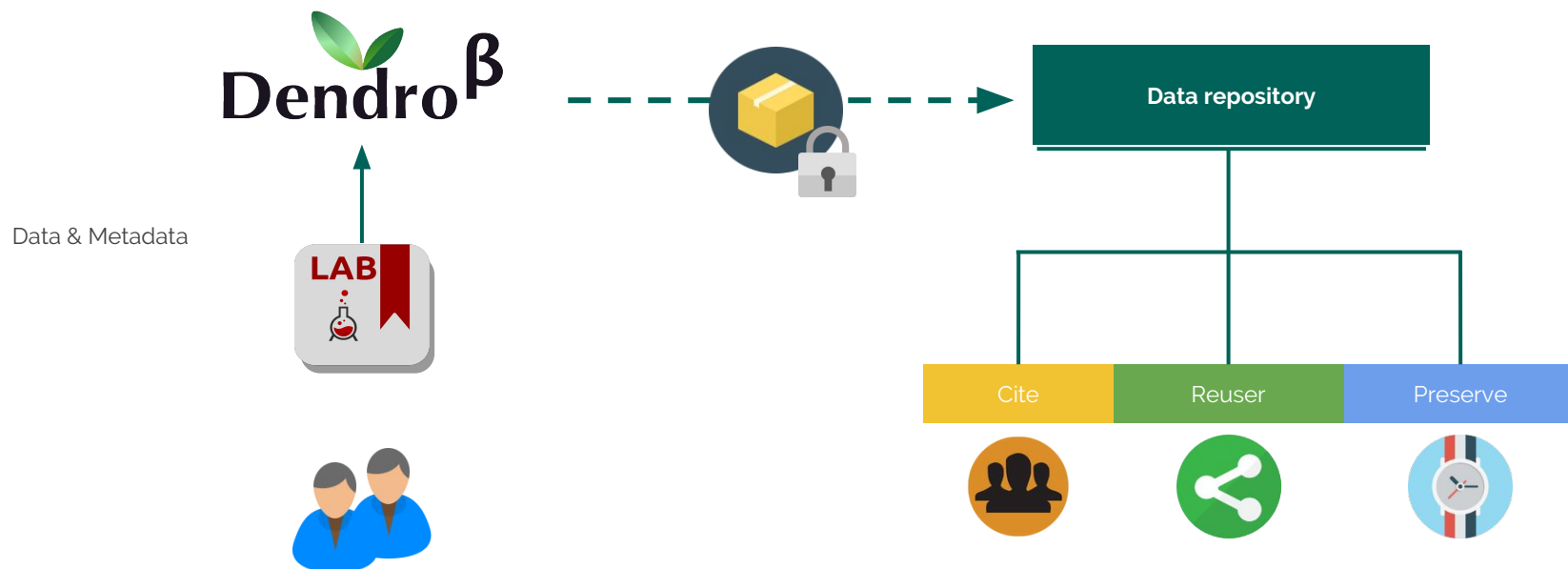


- Reuse metadata models: generic ones
- Domain ontologies: collaboration between curators and researchers

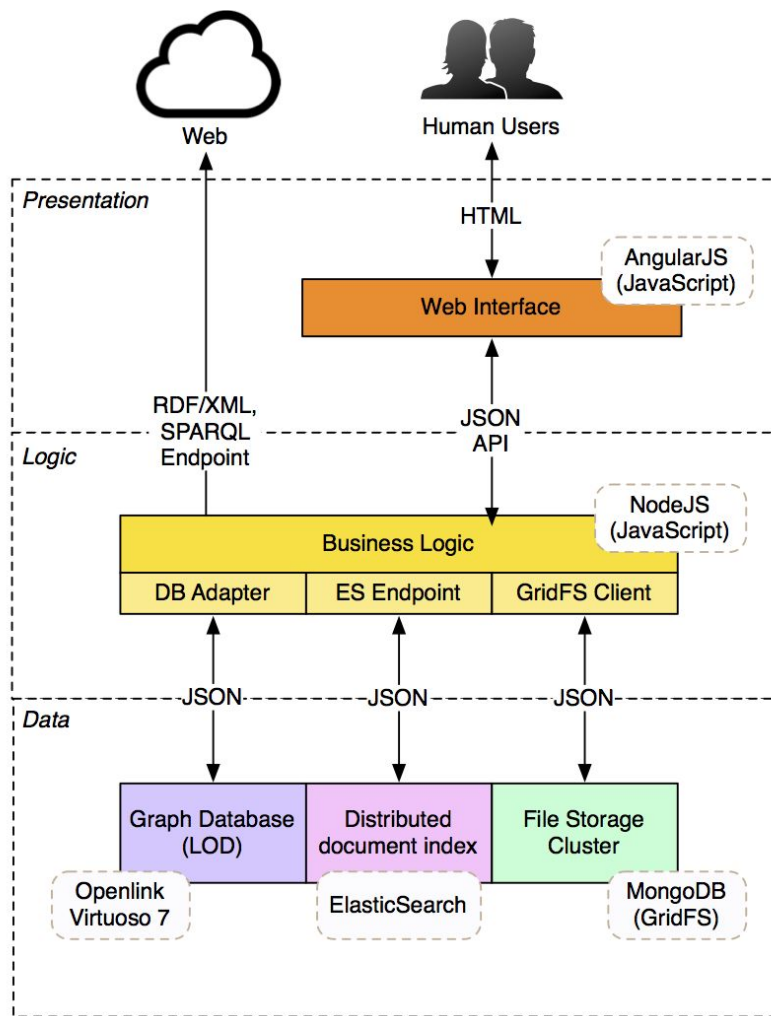
Domains and types of data

Domain	Data Type	
Fracture Mechanics	Experimental	
Hydrogen Production		
Analytical Chemistry	Experimental	Observational
Gravimetry		
Biological Oceanography		
Biodiversity	Observational	
Social and Behavioural Sciences		
Cutting and Packing	Simulational	
Vehicle Simulation		
Computational Fluid Dynamics		

Repositories and long-term preservation



Architecture





- Modules for several steps of the RDM workflow:
 - Management and communication in teams
 - Platform for cloud computing
 - Metadata harvesting and indexing
 - Long-time preservation

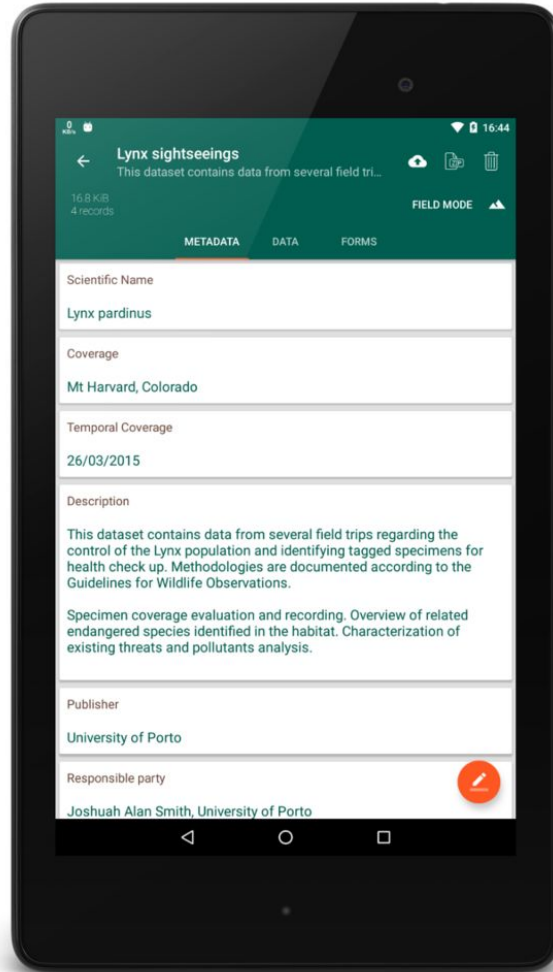




- Collaboration RDM@FEUP - EUDAT
 - **DataPublication@UPorto:**
*multi-disciplinary data description
and deposit linking the Dendro
staging platform with the EUDAT
European Infrastructure*

<https://eudat.eu/communities/datapublication-uporto>





✓ SAVE
↶ UNDO
✓ COPY FROM PARENT
ℹ IN MANUAL MODE ▾
🗑 CLEAR

Temporal Coverage Added

26/03/2015

Scientific Name Added

Lynx pardinus

Coverage Added

Mt Harvard, Colorado

⏪ Biodiversity evolution studies

For INSPIRE-represented observational data for biodiversity. Reference system identifier, Metadata point of contact...

⌚ CONFORMITY DATE ▾

This citation shall include the date (date of publication, date of last revision or creation) of the implementing rules adopted under Article 7(1) of Directive 2007/2/EC of the specification.

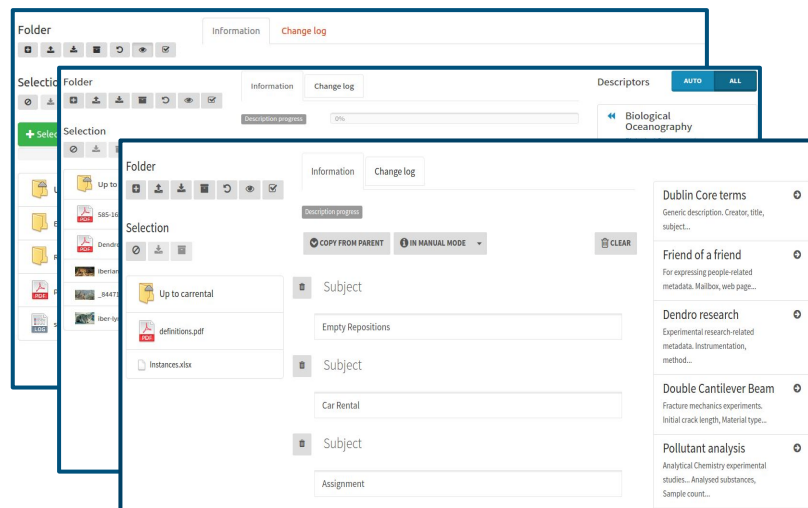
⌚ CONFORMITY DATE TYPE ▾

The citation of product specification shall include the data type of the implementing rules adopted under Article 7(1) of Directive 2007/2/EC or of the specification. The value domain of this metadata element a code list: - Date of creation; - Date of last revision; - Date of publication.

⌚ CONFORMITY SPECIFICATION ▾

This is the degree of conformity of the resource to the implementing rules adopted under Article 7(1) of Directive 2007/2/EC or other specification.

Dendro & EUDAT: exporting data



Vehicle reservation assignment in car rental

Beatriz Oliveira

University of Porto

Abstract: Data for the vehicle-reservation assignment problem in a car rental company, including vehicle and reservation information, car transfers time and cost (collected for Portugal), and upgrading and downgrading policies. Data used in Oliveira, B.B., Carravilla, M.A., Oliveira, J.F. and Toledo, F.M.B., 'A relax-and-fix-based algorithm for the vehicle-reservation assignment problem in a car rental company', European Journal of Operational Research, Volume 237, Issue 2, 1 September 2014, Pages 729-737, ISSN 0377-2217, <http://dx.doi.org/10.1016/j.ejor.2014.02.018>.

Files			
Name	Date	Size	
files.zip	10 Jul 2016	2.1 MB	Download
metadata.rdf	10 Jul 2016	1.9 kB	Download
metadata.txt	10 Jul 2016	1.7 kB	Download

Export

Export as [BibTeX](#), [MARC](#), [MARCXML](#), [DC](#), [EndNote](#), [NLM](#), [RefWorks](#)

Metadata

PID:	http://hdl.handle.net/11304/12dc-e69e-6355-4353-b021-8887fed6528c
Publication:	University of Porto
Licence:	Attribution-ShareAlike 4.0 International
Uploaded by:	dendrordm@gmail.com
Domain:	generic
Checksum:	8c5azedgbd664ca14121a957d9764a046aaeaba8cb615151225doadb4e128f73

Rate this document:



(Not yet reviewed)

[Report abuse](#)

Powered by **INVENIO**

Dendro & EUDAT: exporting data

The screenshot shows a web form for creating a project. It includes fields for 'Handle', 'Title', and 'Description'. Below these is a 'Privacy' section with three radio button options. A red rectangular box highlights the 'Privacy' section, and a red line connects it to a callout box on the right containing the text 'Access Levels'.

Handle *

Handle

A unique handle in the system for your project. Example: "gravimetry01"

Title *

Title

Your project's title. Example: "Gravimetry run campaign over the Azores"

Description *

Project description...

Your project's description, maybe the abstract of its public grant application? Example: "Results of the third Gravimetry run campaign over the Azores"

Privacy

- ☒ Public: The project's metadata will be exposed and the project made publicly available, as read-only.
- ☐ Private: Your project won't have any exposure.
- ☐ Metadata Only: The project's metadata will be exposed, but access to the project will be required to you.

Access Levels

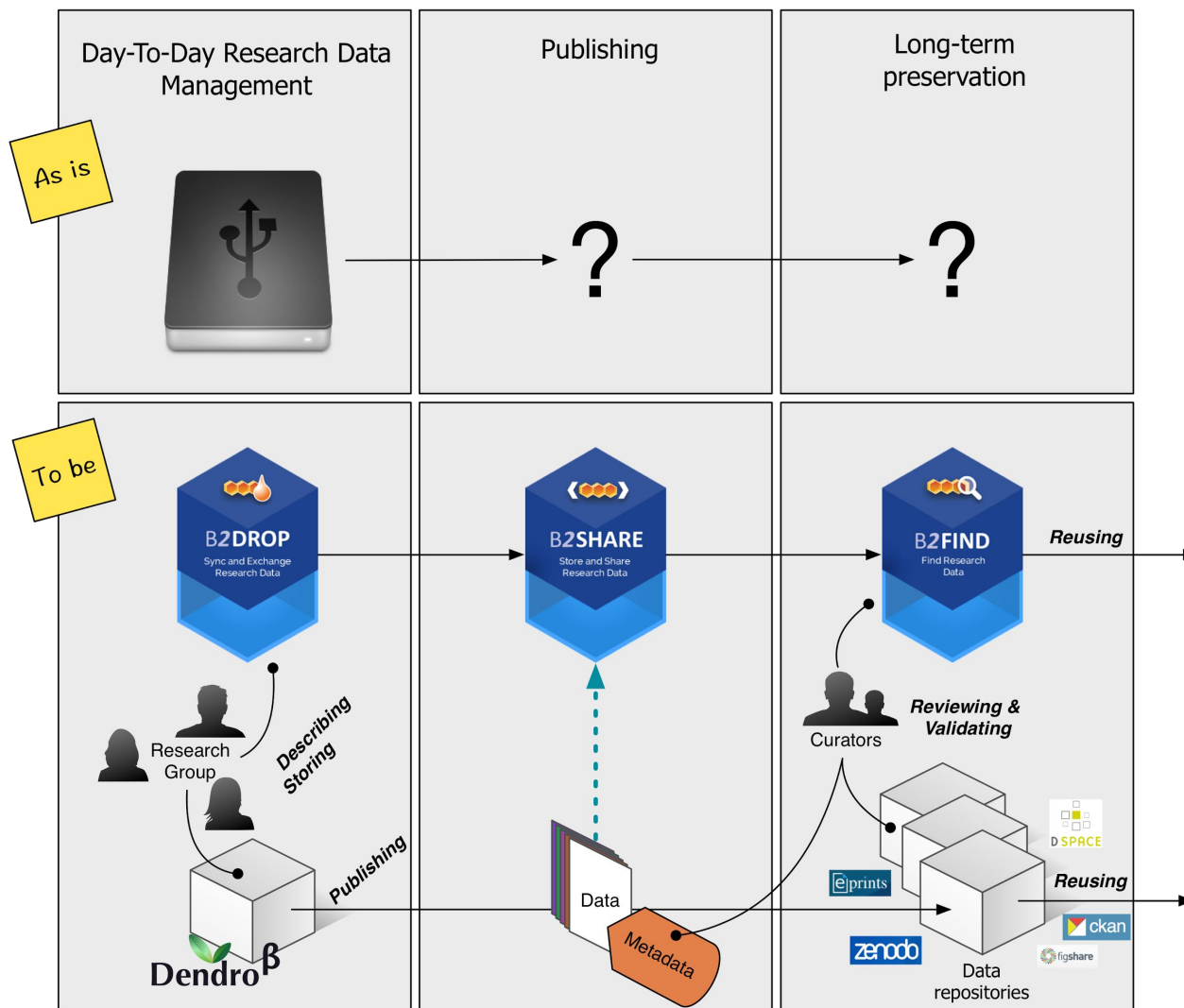
CREATE

Dendro

&

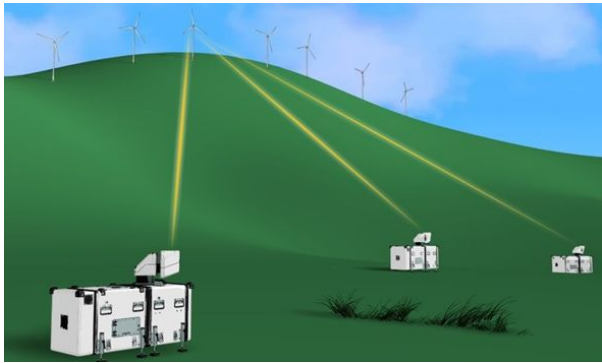
EUDAT

- Early data management practices
 - From project day 1
 - Involve researchers
- Collaborative environment
 - Data peer review
 - Guidelines from senior researchers
- Persistent identifiers
- Data preservation
 - Compliance with existing standards
- Dissemination
 - Automatic harvesting
 - Dissemination protocols



To know more about EUDAT and Dendro, please visit <https://eudat.eu/> and <http://dendro.fe.up.pt>

Alignment with european e-infrastructures



<http://www.windscanner.eu>



<https://www.epos-ip.org/>



<http://www.lifewatch.com/>

TAIL Team



Cristina
Ribeiro



João Rocha
da Silva



João Aguiar
Castro



Ricardo
Amorim



Yulia
Karimova



Carla
Lopes



Artur
Rocha



João
C. Lopes



Gabriel
David



Tito
Vieira



Eugénia
Fernandes



Ângela
Lomba



João
Honrado



Related publications

- ❑ TPDL 2016
 - ❑ **Usage-driven Dublin Core descriptor selection—A case study using the Dendro platform for research dataset description**
- ❑ UAIS, Springer 2016
 - ❑ **A comparison of research data management platforms: architecture, flexible metadata and interoperability**
- ❑ iPRES 2014
 - ❑ **The Dendro research data management platform: applying ontologies to long-term preservation in a collaborative environment**
- ❑ Digital Libraries 2014
 - ❑ **Creating lightweight ontologies for dataset description: Practical applications in a cross-domain research data management workflow**



Visit us

Publications and updates

<http://dendro.fe.up.pt/>

Dendro is *open source* on GitHub!

<https://github.com/feup-infolab-rdm>