

# Research Data Management at INESC TEC using Dendro, CKAN and EUDAT

**Cristina Ribeiro**, João Rocha da Silva, João Aguiar Castro, Yulia Karimova,  
Nelson Pereira, Gabriel David

INESC TEC—Faculdade de Engenharia da Universidade do Porto



UNIÃO EUROPEIA  
Fundo Europeu  
de Desenvolvimento Regional



# TAIL\_Team@INESC\_TEC



Cristina  
Ribeiro



João Rocha  
da Silva



João Aguiar  
Castro



Yulia  
Karimova



Nelson  
Pereira



Carla  
Lopes



Artur  
Rocha



João  
C. Lopes



Gabriel  
David



William  
Fukunaga



Bruno  
Marques



# 10 Questions

- ❑ 1: What is Research Data Management?
- ❑ 2: Why RDM?
- ❑ 3: Do we have the tools and processes?
- ❑ 4: What do we need technologies for?
- ❑ 5: How is RDM at INESC TEC?
- ❑ 6: Where are the traps?
- ❑ 7: Why link to EUDAT?
- ❑ 8: Does one size fit all?
- ❑ 9: What have we achieved?
- ❑ 10: What is missing?



# 1: What is Research Data Management?

- Research data: originated in research
- ... but also other sources
- RDM: a researcher-centric approach



## 2: Why RDM?

- Funding bodies are requesting it: EU, FCT
- Open Science
- Rewards for researchers: new outputs, more impact

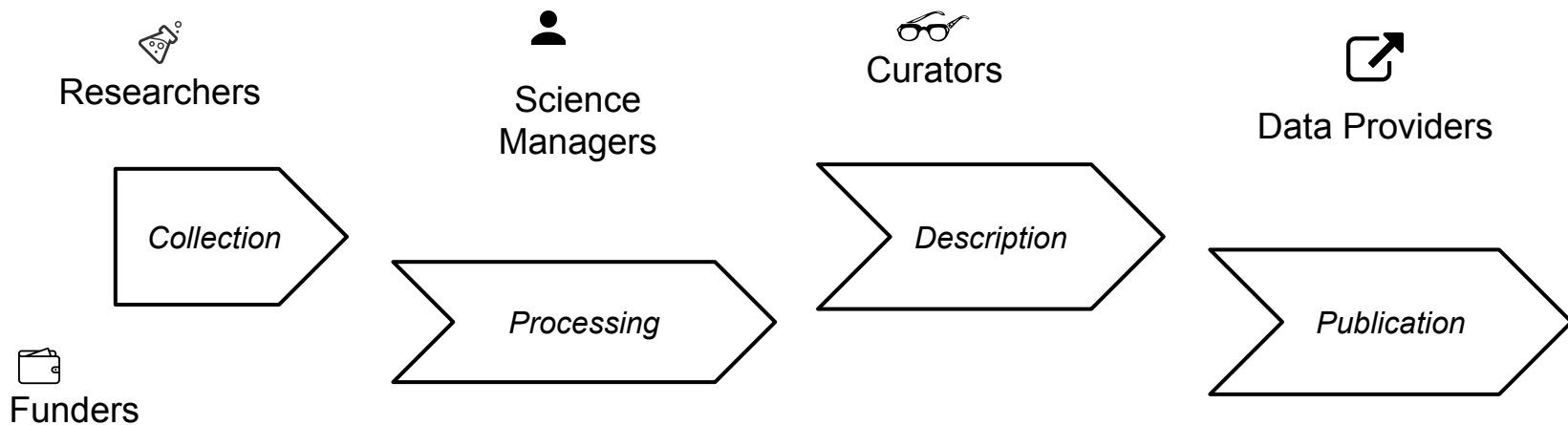
### 3: Do we have the tools and processes?

- May look like: institutional repositories are mature
- But data requires different processes
- Open Knowledge Foundation:
  - advocacy, technology and training...
- Tools for researchers: not yet there



## 4: Technologies for what?

- Data collection and organisation
- Metadata model design
- Metadata creation
- Data deposit
- Data publication



Developers

[dendro.fe.up.pt](http://dendro.fe.up.pt)










Institutions






# 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

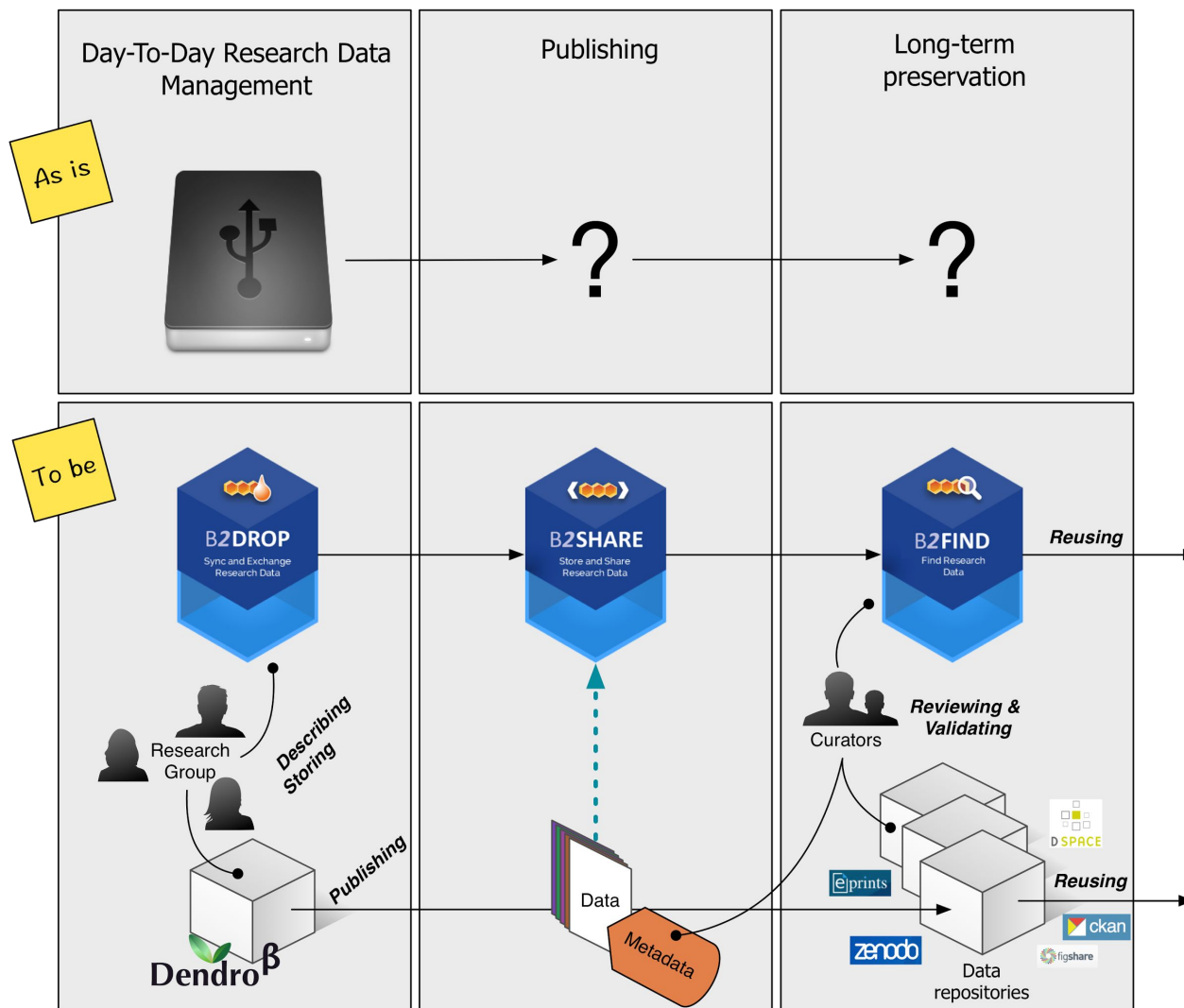
★ 

## 5: What RDM at INESC TEC?

- Legacy data
- Multiple workflows
- CKAN repository
- Deposit on EUDAT B2SHARE
- Data publication
- ...researchers at the center of the stage

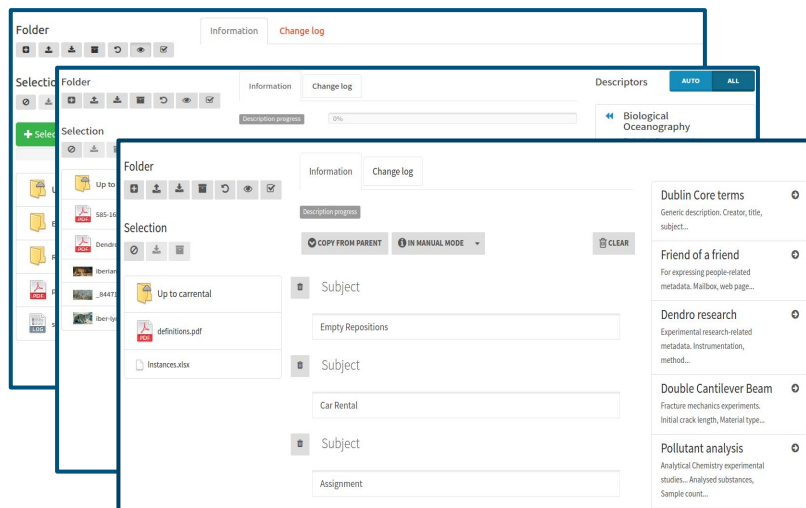
## 6: Why link to EUDAT?

- European “Common Data Infrastructure”
- 10-year horizon
- Services for RDM at several levels
- Persistent identifiers, dissemination
- Preservation



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

# 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	
<a href="#">files.zip</a>	10 Jul 2016	2.1 MB	<a href="#">Download</a>
<a href="#">metadata.rdf</a>	10 Jul 2016	1.9 kB	<a href="#">Download</a>
<a href="#">metadata.txt</a>	10 Jul 2016	1.7 kB	<a href="#">Download</a>

### Export

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

### Metadata

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

Rate this document:



(Not yet reviewed)

[Report abuse](#)

Powered by **INVENIO**

## 7: Where are the traps?

- Project structure (TAIL)
- Software sustainability
- Repository platforms: not mature (e.g. CKAN)
- EUDAT B2SHARE: also a project
- ...

## 8: Does one size fit all?

- No, but...
- Many common problems in diverse domains
- Find the common requirements and solutions
- Metadata is where we specialize
- ...also looking at roadmap infrastructures



<http://www.windscanner.eu>



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



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



## g: What have we achieved?

- Dendro development
- CKAN development contributions
- Ontologies for specialized metadata
- Data deposit at INESC TEC and B2SHARE
- Roadmap links: [windscanner.eu](http://windscanner.eu), PORBIOTA (Lifewatch), C4G (EPOS)
- International links: EUDAT, Research Data Alliance

# 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		

## 10: What is missing?

- Connect the dots: DMP, repositories, data reuse
- Identifiers: DOI
- More automated metadata
- Metadata exported as Linked Open Data
- Link to aggregators
- The connection to a national network (forthcoming GT-PNCA recommendations)

## Funções Concentradas

- Coordenação Nacional & Articulação Internacional
- Normalização & Certificação
- Monitorização & Estatísticas

## Plataformas de Acesso e Preservação

- Repositórios de acesso (Ex: RNOD, RCAAP, Dados.gov)
- Repositórios de preservação (Ex: RODA, RCAAP)
- Ferramentas de apoio, tais como planificadores, validadores, visualizadores, manipuladores de dados, plataformas e-Science.
- Formação & Suporte

## Infraestruturas partilhadas (*e-Infrastructure commons*)

- AAI (Ciência ID, Edugain) & Identificadores permanentes (DOI).
- Armazenamento (CEPH)
- Computação (INCD/EGI, LCA/PRACE, etc.)
- Rede (RCTS, GÉANT, rSPtic) & Centros de dados (Nacionais/Internacionais, Setoriais/ Institucionais)

Fonte:  
GT-PNCA-IP  
(preliminary)

# TAIL Team@INESC\_TEC



Cristina  
Ribeiro



João Rocha  
da Silva



João Aguiar  
Castro



Yulia  
Karimova



Nelson  
Pereira



Carla  
Lopes



Artur  
Rocha



João  
C. Lopes



Gabriel  
David



William  
Fukunaga



Bruno  
Marques





# Visit us

Publications and updates

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

INESC TEC data repository

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

Dendro is *open source* on GitHub!

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