# Research Data Management at KU Leuven: infrastructure and services

Leen Van Rentergem

ICTS KU Leuven

Welcome in Leuven
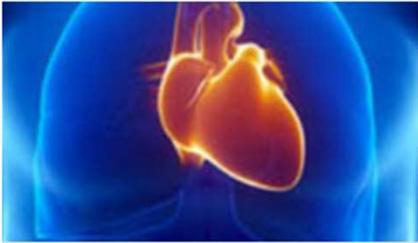
Human Health

Medical Technologies
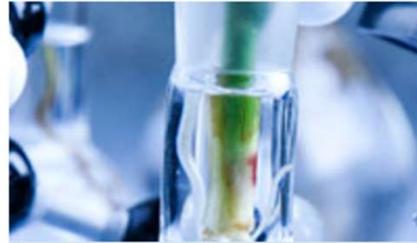
Bio-Sciences & Environment

Matter, Materials & Energy

Nature Unlimited

Manufacturing & ICT

Arts, Religion & Culture

Economy, Law & Society

Human Behavior

> 10.000 Peer Reviewed Publications

149 Spin-offs

130 m€ License Income

> 1850 faculty

KU LEUVEN

Challenge 1 : Why RDM

RDM is relevant for the researcher

RDM helps them…

- collect/create **high quality data** → high quality research
- work **more efficiently**
  - *e.g. develop consistent procedures early on, check that you fulfill legal and ethical obligations, improve impact and quality of grant applications, etc.*
- **Protect** their research data (e.g., security measures, data loss,…)
- Following their **values** as a researcher:
  - striving for **scientific integrity**
  - ensuring **scientific reproducibility**

Challenge 1 : Why RDM

RDM is relevant for institutes

RDM helps them...
- collect/create **high quality data** → high quality research
- **Protect** their research data (e.g., security measures, data loss,…)
- Track and promote **reuse of data**
- Following their **values** as a institution:
  - striving for **scientific integrity**
  - ensuring **scientific reproducibility**

KU Leuven data policy

KU Leuven considers research data as a valuable research output, an asset to KU Leuven and a critical contribution to the knowledge economy. A high standard of research data management is fundamental to both high quality research and academic integrity.

CORE PRINCIPLES OF RESEARCH DATA MANAGEMENT AT KU LEUVEN

1. research data must be stored and documented in a safe, secure & sustainable way - in this way data can be retrieved and accessed by KU Leuven when needed

WHAT DOES THIS IMPLY?

2. metadata of research data must be registered to ensure findability of the data

WHAT DOES THIS IMPLY?

3. deletion of data must be justified and documented; documentation relating to the data should not be deleted as this impedes the audit trial

WHAT DOES THIS IMPLY?

4. after the end of the research, relevant research data must be retained for a minimum of 10 years in a safe, secure & sustainable way for purposes of reproducibility, verification, and potential reuse; a best effort is expected to make relevant research data available for reuse

WHAT DOES THIS IMPLY?

5. all research is carried out in accordance with and taking into account existing contractual agreements, legislation, regulations, guidelines or valorization potential

WHAT DOES THIS IMPLY?

KU LEUVEN

# Challenge 2: Research ecosystem is complex, research data even more

Starting project		Active project				Published Project		Long-term (from during to years beyond the project)	
Data management plan description	(Legal) compliance	Data Collection	Data Storage	Metadata description	Data Processing	Data Publishing	Metadata Publishing	Data Archiving	Data Preservation

RDM infrastructure and tools at KU Leuven
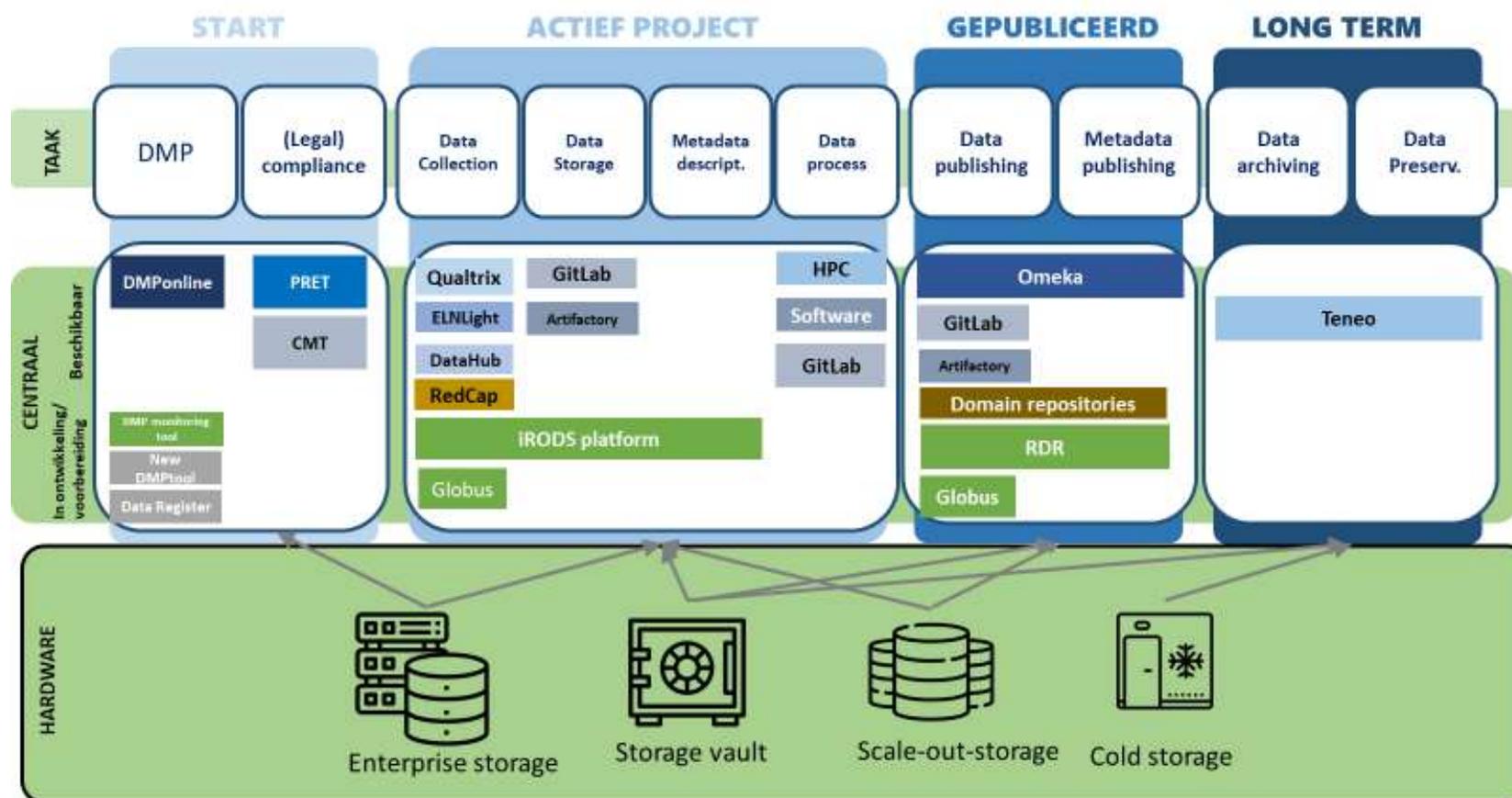
Data flow between data storage systems

ACTIVE DATA

| Data Collection | Data Storage | Metadata descript. | Data process |

PUBLISH DATA

| Data publishing | Metadata publishing |

LONG TERM STORAGE

| Data archiving | Data Preserv. |

Other Storage
- GVS
- OneDrive
- Shared drives
- HPC

iRODS
KU LEUVEN

Regional/domain specific rep.

RDR
KU LEUVEN
The Dataverse Project

Teneo (preservation)

Cold storage (archiving)
iRODS
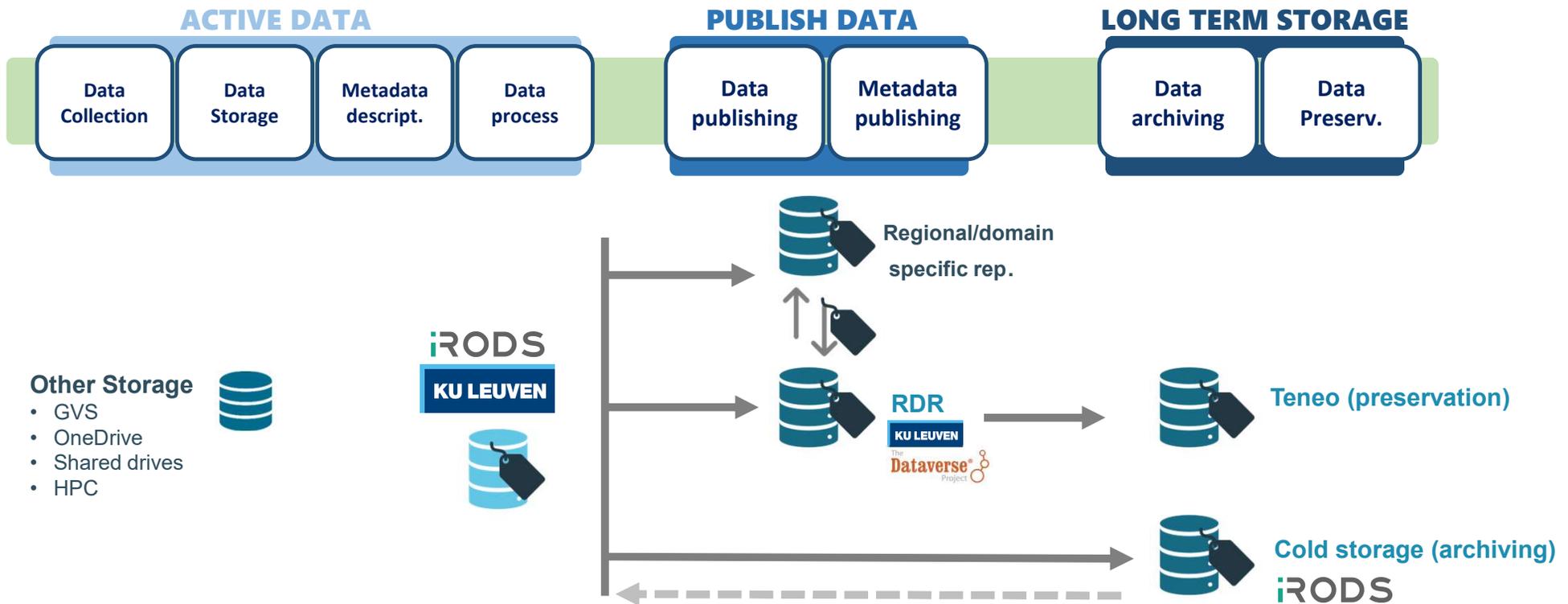
Challenge 3: Move in a Connected world
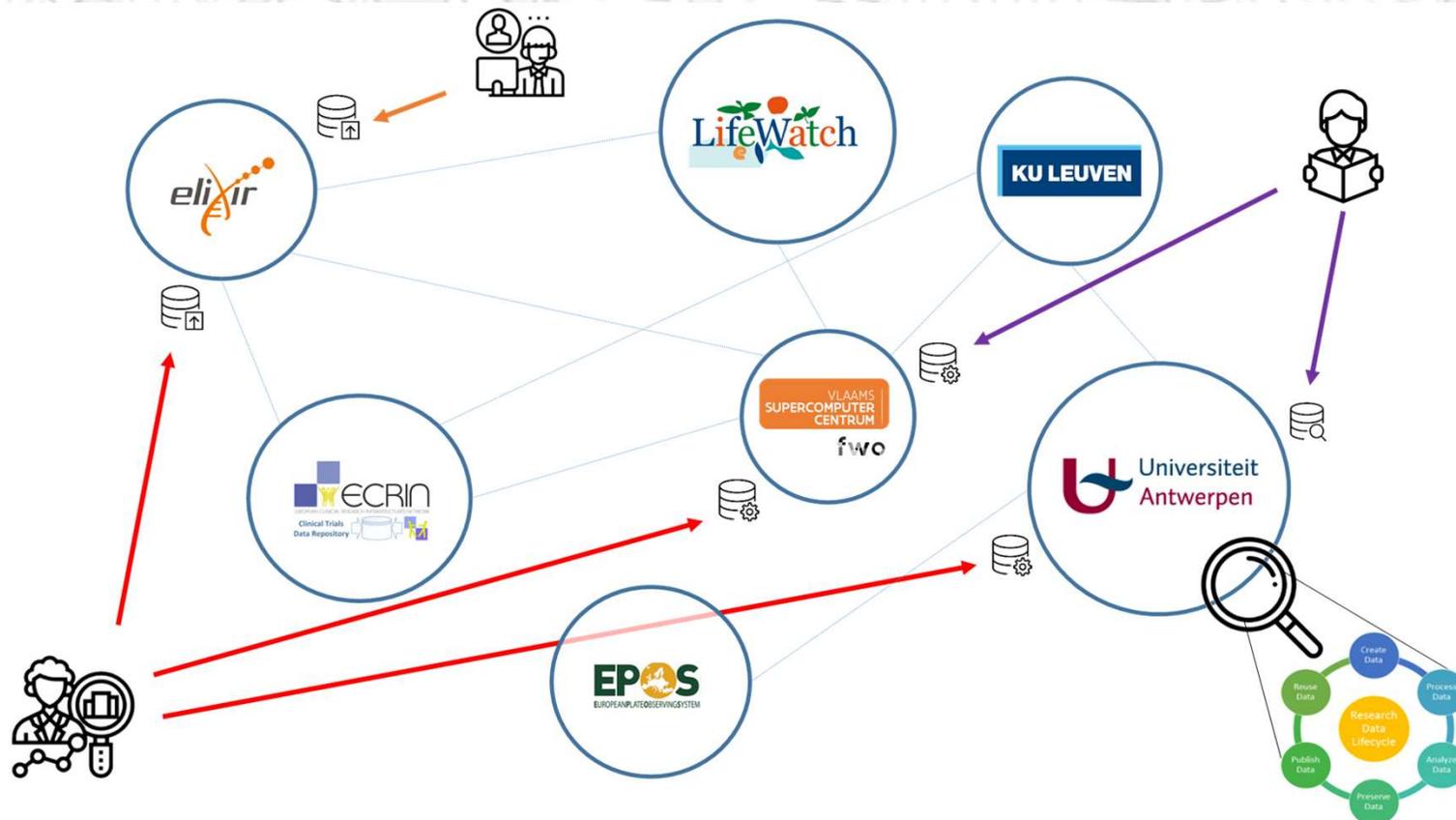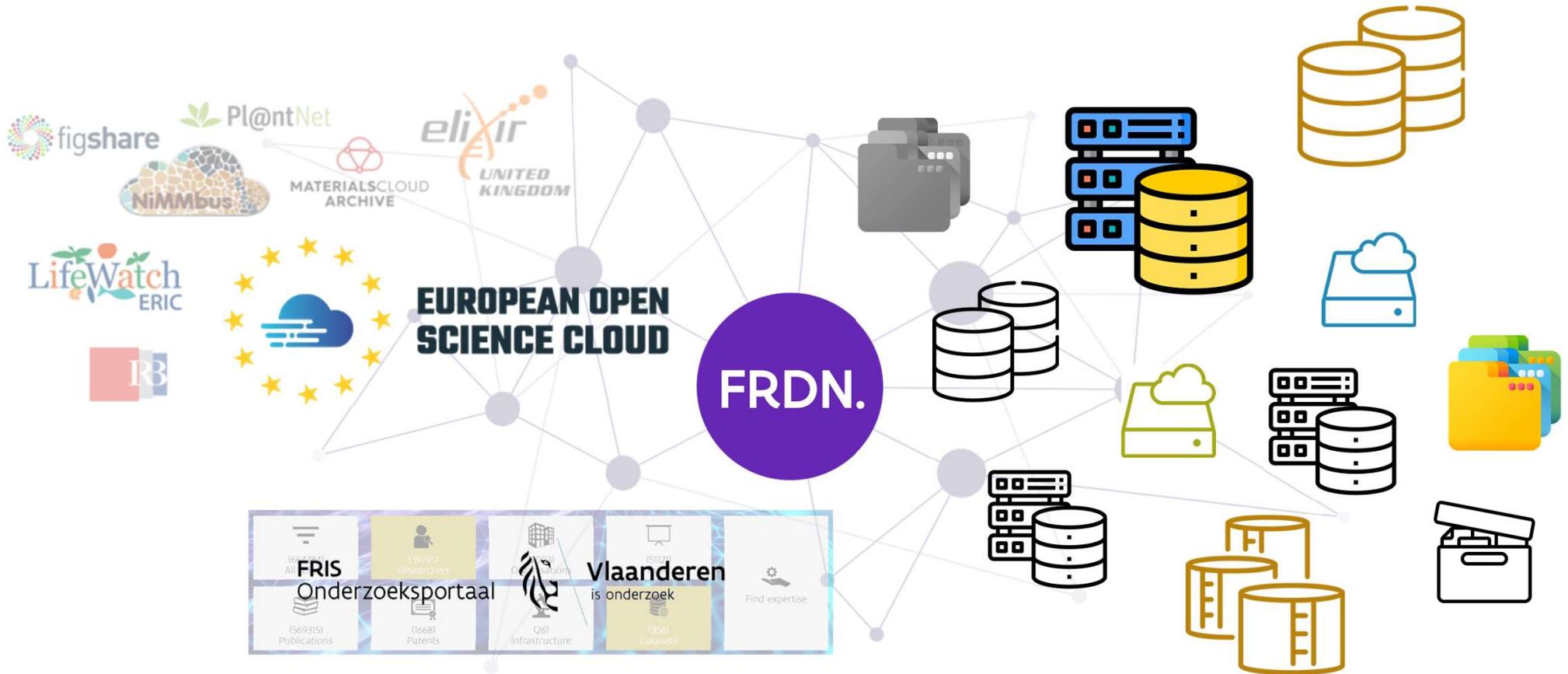
Best practice guidance for research data management

Search RDM website

[   ] [Search]

RDM HELPDESK

Funder requirements

data management requirements for various funders, with practical tips

Write a DMP

practical guidance on how to write a data management plan

Data & standards

what are research data, file formats and standards

Legal and ethical compliance

data privacy, intellectual property rights and ethics with solutions

Documentation and metadata

guidance for annotating data and generating and exporting metadata

Storage

central solutions to store research data and collaborate with partners during research

Preservation

which research data to keep where and for how long

Data sharing

data repository solutions to archive research data and make them available for reuse

FAIR data

making research data findable, accessible, interoperable and reusable

Responsibilities

for managing data and keeping good oversight

Challenge 4: "FAIR"ification
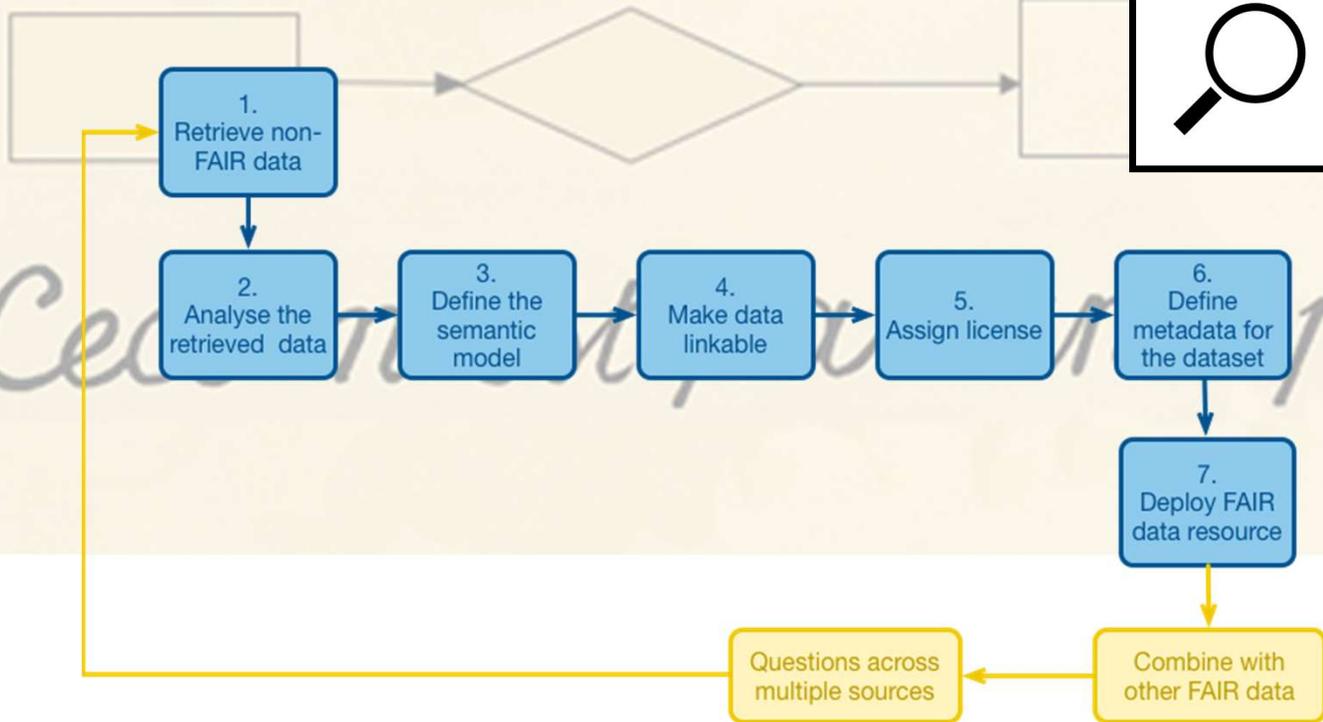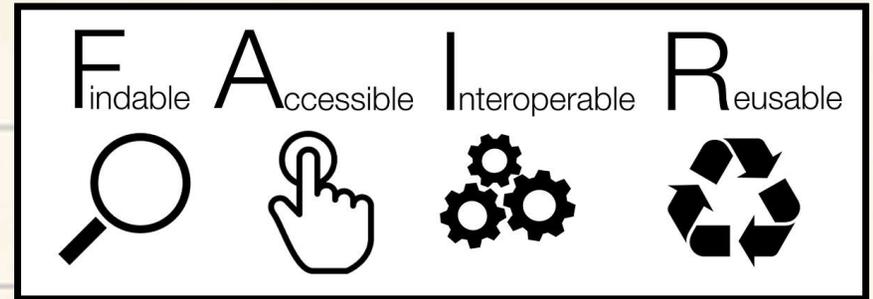
Help and encourage researchers
to apply <u>good data management practices</u>

Right
FROM THE
✓**Start**

Regional/domain specific rep.

Long term storage

iRODS KU LEUVEN

RDR KU LEUVEN
The Dataverse Project

Teneo (preservation)

Cold storage (archiving)
iRODS

Active data

Published data
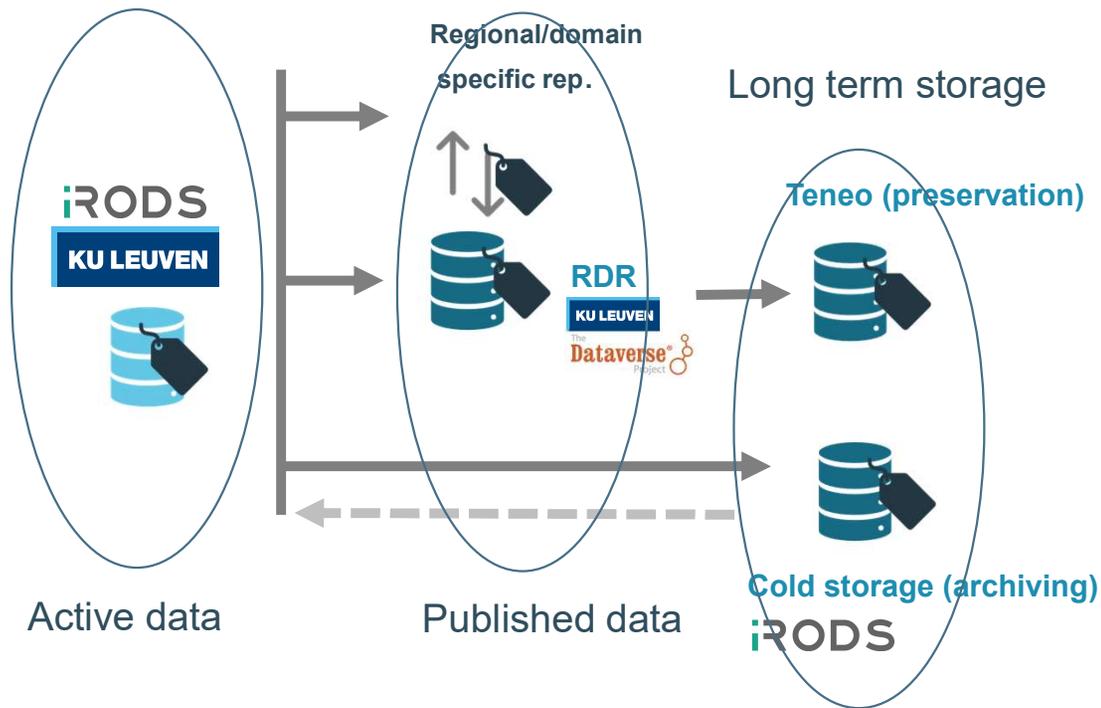
💡 **Give context to data**

⚙️ **Automate data workflows**

🔍 **Search data**

Share data

🔒 **Secure data**

KU LEUVEN

Active data platform (iRODS)

Central Storage

- High Available
- Secure

Multiple clients/interfaces

- GUI, CLI, WebDav, NFSRods, Python, Go, etc
- Linux, MacOS and Windows

iRODS

Metadata

- Rich metadata templates
- Management interface:
  - Upload/create/modify/share
  - Enforce standardization
  - Combine and apply

Workflow automation

- System wide/project rules:
  - e.g: netCDF metadata extraction)
- Delegation of the
  implementation/deployment of rules

Challenge 5: The Cost

Cost / Volume

Volume Demand
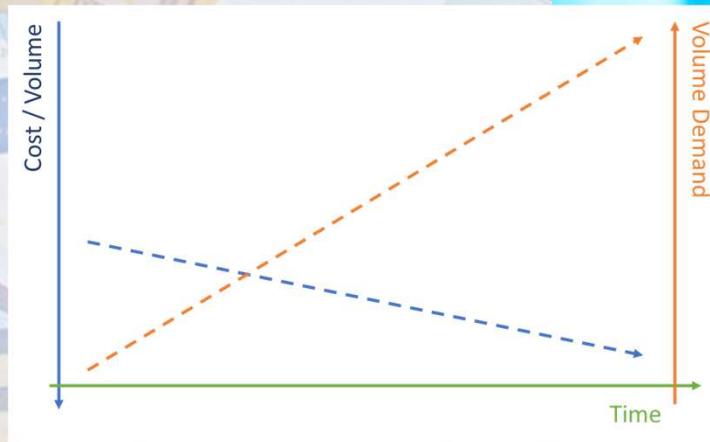
Time

Data life time

Funding period

Funding period

Data Life time

Sponsoring

DEPARTEMENT ECONOMIE WETENSCHAP & INNOVATIE

Vlaamse overheid

€

KU LEUVEN

Tiering

Tier 1	Tier 2	Tier 3	Tier 4
Mission Critical	Hot	Warm	Cold

&
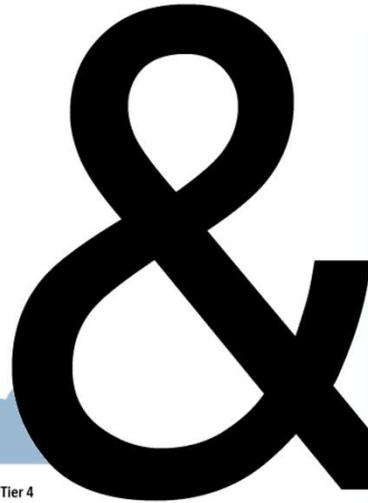
Use the right solution
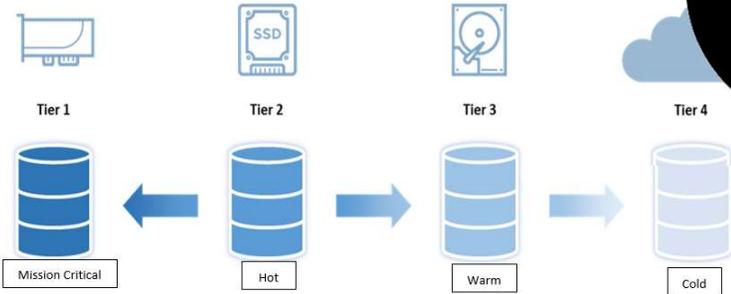
KU LEUVEN

Storage Guide

KU Leuven Home >

This Storage Guide helps you choose central IT solutions for data storage and management. Based on the needs you indicate in the left column, the possible solutions are shown. Click on a solution to see more details. From the details you can click even further to the relevant service sheet in the ICTS Service Catalogue.

For more information please contact the ICTS Service Desk or the RDM Support Desk

Clear all filters

What is the classification of your data with regard to confidentiality?
- Not confidential
- Confidential
- Strictly confidential

Do you want to share data with colleagues?
- No
- Yes, but only within the university
- Yes, with persons outside the university

What type of data do you want to store?
- Research data
- Other

The ability to add metadata is important to me.
- Yes
- Less important

How much storage space do you think you will need?
- Less than 1TB
- Between 1TB and 5 TB
- More than 5TB

To what extent are your data reproducible?
- Easy
- Difficult
- Not

iRODS
Solution for storing active research data, with corresponding management tools for (meta)data.

Large Volume Storage
Solution for storing large amounts of research data in a cost-efficient manner.

Shared network drive
A shared file system available to all staff via PCs and laptops managed by ICTS KU Leuven or your local KU Leuven IT department.

SharePoint on premise-site
'On premise' Microsoft solution for storing, managing and sharing documents and files, hosted in the ICTS data center.

SharePoint online-site
Microsoft cloud solution to store, manage and share documents and files.

Teams-site
Microsoft cloud solution for communication and collaboration.

Personal Network Drive
A personal file system available to most staff via PCs and laptops managed by ICTS KU Leuven or your local KU Leuven IT department.

OneDrive for Business cloud storage
Personal cloud storage via KU Leuven Microsoft 365.

Enterprise Box cloud storage

Local storage on managed PC/laptop

KU LEUVEN

Challenge 6: The Expertise

RDM Network @ KU Leuven

RDM Competence Centre

RDM Policy initiatives

- **International** (e.g. LERU)
- **National** (e.g. FOSB)
  ↓
- **KU Leuven context:**
  - Open Science Task Force
  - RDM steering committee
  - Underlying WGs:
    - Policy
    - Infrastructure
    - Advice&Training

RDM Support Staff
Central & Embedded

Departement
Faculty
Libraries
Local services (e.g., IT)
Central Services
RDM Network KU Leuven Libraries
…

RDM Support Desk

RESEARCHERS

Strategic/Policy

Operation / Support

Challenge 7: Researcher-first

Machine-actionable
Data management Plan
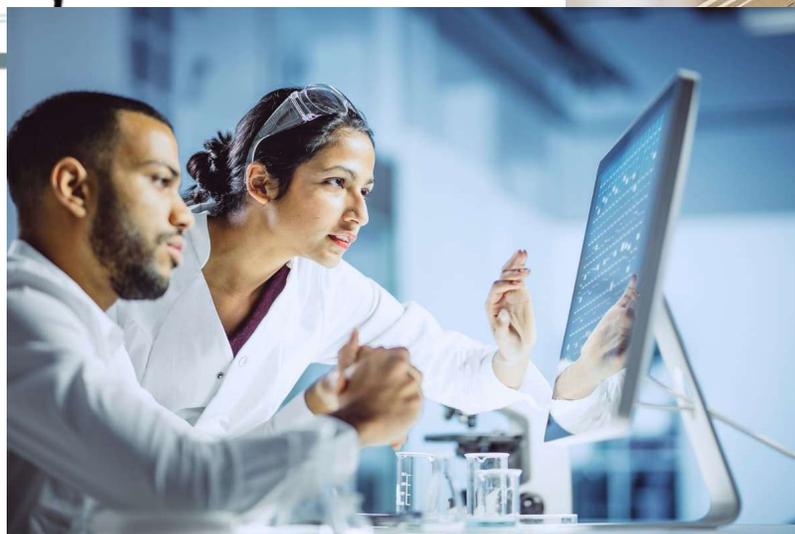
Data Steward

"I ensure data usage
complies with our
policies."

Dashboard

Collections

Advanced search

Your groups

Trash

Metadata templates

🏠 kuleuven_tier1_pilot » home » vsc33436 » irods-portal-python

irods-portal-python

Add collection   Choose file   No file chosen   Upload

Content   Meta data   Permissions

other   project   another schema

Area	Engineering
Project name	Mango: acti
Description	An iRODS P and vanilla J

Add element

colour
- red
- blue
- green
- yellow

- Required   ↑ ↓ Delete

Add element

Length
- 1m
- 1cm
- 1mm

☑ Required   ↑ ↓ Delete

Add element

KU LEUVEN

The greater the obstacle, the more glory in overcoming it.

~ Moliere

AZ QUOTES