FAIRVault

Nicky Daniels, Jacob Laureyns, Kevin Leonard, Myriam Mertens, Philippe Moens, Ine Moerman, Jone Paesmans, Jef Peeters, Sybren Slegers, Thomas Van de Velde

Abstract

The FAIRVault project is an interuniversity collaboration between four Flemish universities (Ghent University, Hasselt University, University of Antwerp, and Vrije Universiteit Brussel) to develop a generic, FAIR-enabling solution for archiving research data (in particular sensitive data) for which no suitable repository exists.

Why the FAIRVault:

Researchers and institutions are expected to:

Preserve for a specified retention period & **provide access** to finished research data for validation & reuse

*"As **open** as possible, as **closed** as necessary"*

Protect data according to legal and ethical frameworks

Existing repositories do not always meet our needs:

- Insufficient access control mechanisms for **sensitive data**
  - **Variability** in data sensitivity
- Compliance with applicable **legislation/regulations**?
- Not always trusted by researchers
- Constraints of **scope** & collection policies
- Data beyond institution's control
- Sometimes **dissemination** rather than preservation focus

Current research data 'archiving' practices are often suboptimal

- Use of expensive, 'active data' storage types
- Lack of **metadata** & documentation
- Risks of data alteration
- No clear disposal process
- No (proper) **access request** & reuse management

Different institutions have shared needs, but different requirements:

- Benefit from shared infrastructure and expertise
- Highly sensitive data may require local storage
- Data curation at institutional level

FAIRVault in a nutshell:

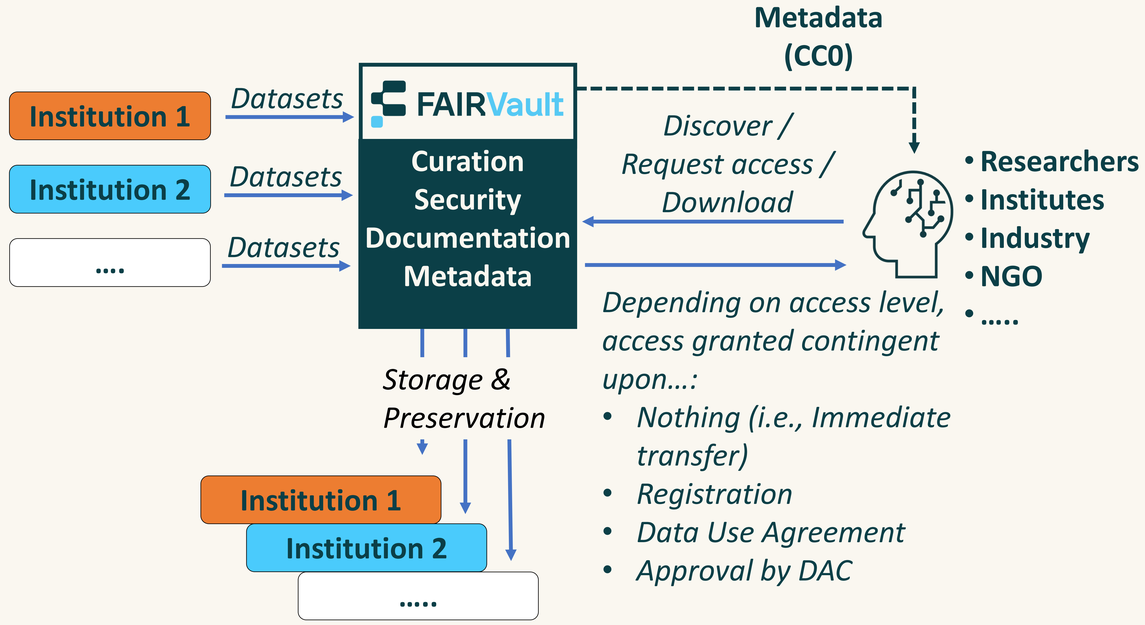
Secure environment
Create a trusted solution for researchers, including those dealing with sensitive data

Institutional control
Custom storage, centrally controlled access to institutional research data assets

Collaborative
Multi-institutional setup to address shared needs
Minimise redundancy and optimise resources – exchange knowledge

Open-source
Making use of Dataverse, engaging in an already vibrant community with longevity and stability

The concept:



Choosing a Data Repository Software

Feature Requirements:

- Customizable, multi-tenant storage solutions
  - S3, Globus endpoint, trusted remote storage
  - Configurable on institutional- or dataset-level
- Graded sensitivity levels
  - File-level access configurability
- Customizable data curation workflows
- Community and sub-community collections
- Integrations with long-term preservation systems
- Active open-source community
  - Users and Developers

Software Evaluation Process:

- Original long list of 14 software choices
- Shortened to two main options:
  - Dataverse and InvenioRDM
- Dataverse selected as option for future testing and development

The Dataverse® Project