

NRP and its Service Ecosystem

National EOOSC CZ Conference 2024, 2024-11-26

Michal Růžička, ruzicka@ics.muni.cz

Masaryk University | CERIT-SC



Co-funded by
the European Union



MINISTRY OF EDUCATION,
YOUTH AND SPORTS

About Me

- Leader of EOSC CZ open working group Core Services.
 - <https://www.eosc.cz/en/working-groups/core-services>
- Leader of end-user-services oriented work packages in the NRP project.
 - <https://www.eosc.cz/en/projects/national-repository-platform-for-research-data-os-i-nrp/national-repository-platform>
- Open Science support team member at CERIT-SC | ICS MU
 - Focus on data management and FAIR data support.
 - Including sensitive data (SensitiveCloud).
- Good coffee and tea lover.

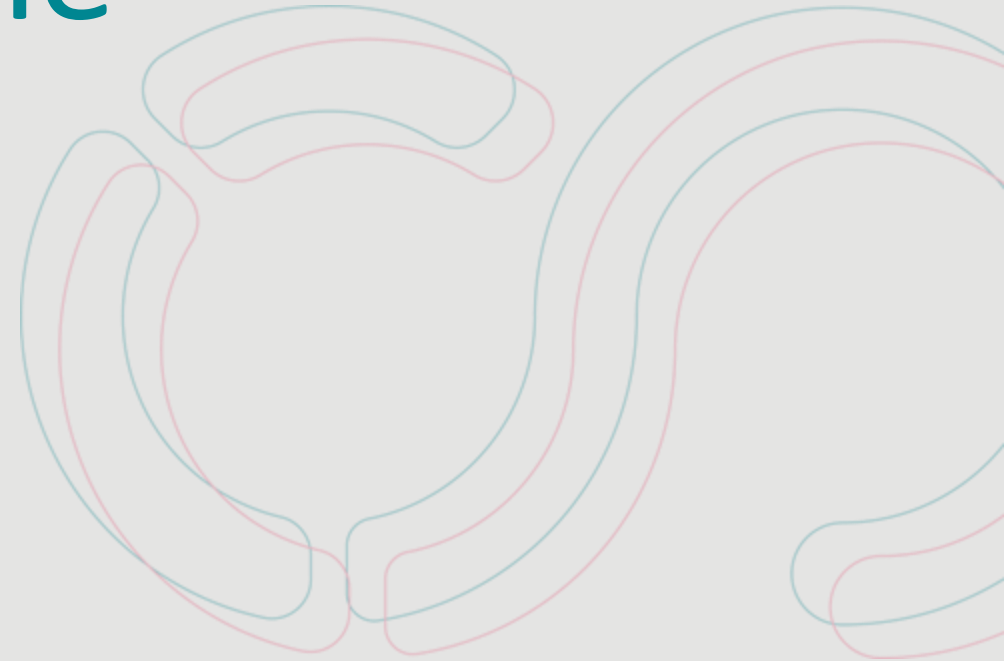
26 November 2024



Outline

- Research **Data Life-Cycle**
- National Repository Platform
(**NRP**) Project
 - **Focusing**
- **NRP Services**
 - High-Level **Overview**
 - Some in More **Details**

Research Data Life-Cycle



Research Data Life-Cycle

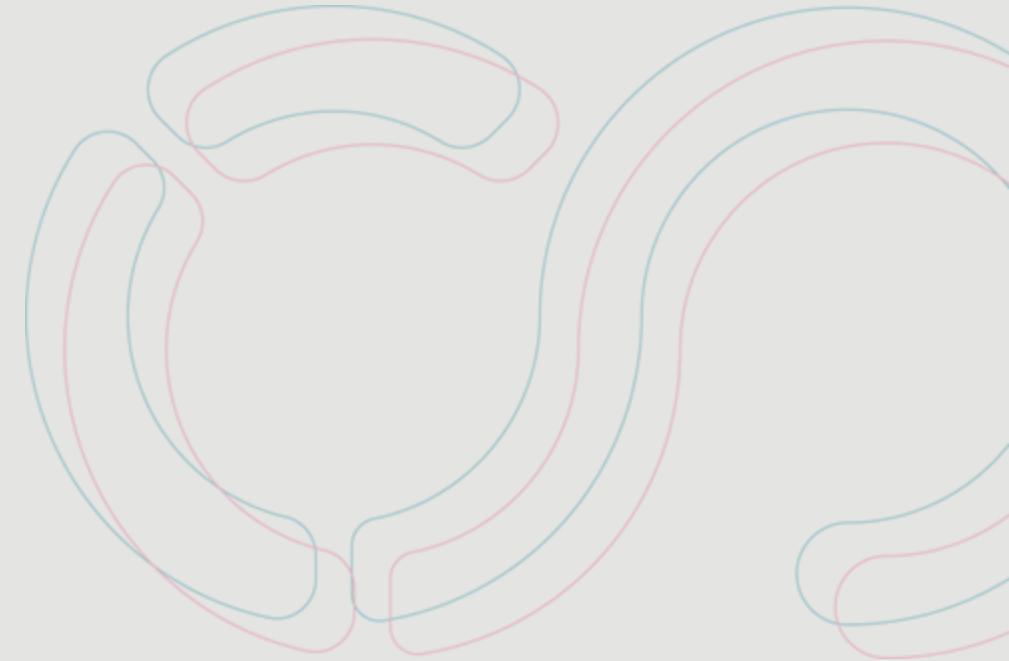


Source: ELIXIR RDMkit, <https://rdmkit.elixir-europe.org/>

- What **data** do you (re)use?
 - Including **licenses** that allow you to do so.
- What **data** do you generate and how?
- **Where** do you store, backup, and long-term archive data?
- How do you **permanently** and **uniquely identify** it?
- How do you **process** it?
- How do you **analyze** it?
- Where do you **publish** and **share** it?
- Who **pays** for all of this?
- **What** are the data **really about**?
- **What** are the data **suitable** (or **unsuitable**) for?
- **Who can reuse** the data?
- What specific data support your results?
- How can they be used to **repeat** your **experiments**, etc.?

National Repository Platform (NRP) Project

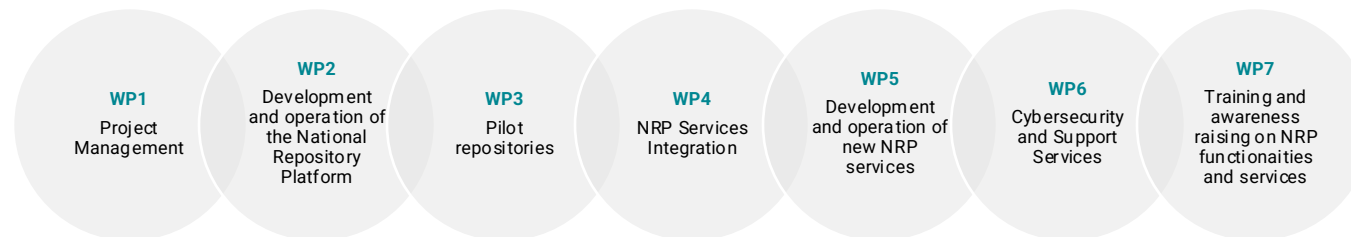
Focusing



NRP for End-Users

- The main objective of NRP:
 - provide various research data-related services for NRP users:
 - academic institutions and
 - individual scientists and researchers.
- What NRP will provide
 - to enhance their abilities to fulfil FAIR principles and
 - data management requirements
 - through the life cycle of research data in projects.

Project NRP

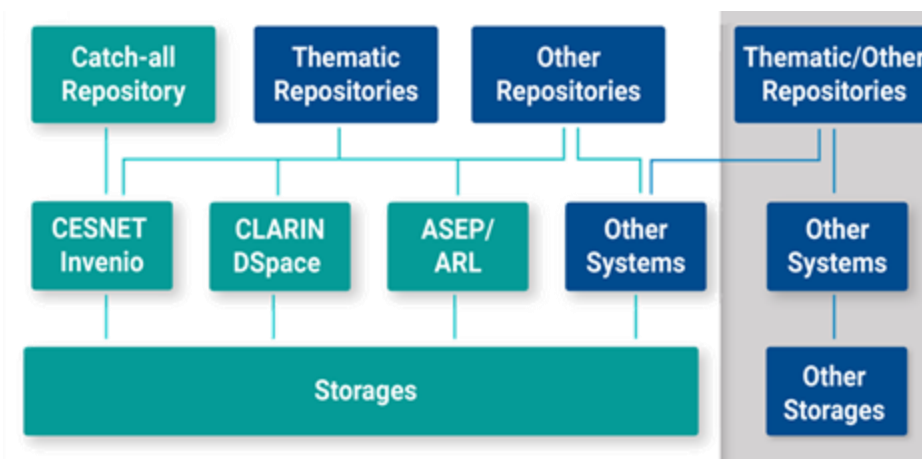


NATIONAL REPOSITORY PLATFORM FOR RESEARCH DATA

MAIN BENEFICIARY: **CESNET** ➔ Project consortium of 11 partners

Project objective:

- Design, implement and pilot the National Repository Platform and related key services
- Provide training for users, data experts and other staff of research organizations
- Create a cybersecurity background, including full compliance with relevant regulations



Project NRP: Services

WP1
Project
Management

WP2
Development
and operation
of the National
Repository
Platform

WP3
Pilot
repositories

WP4
NRP Services
Integration

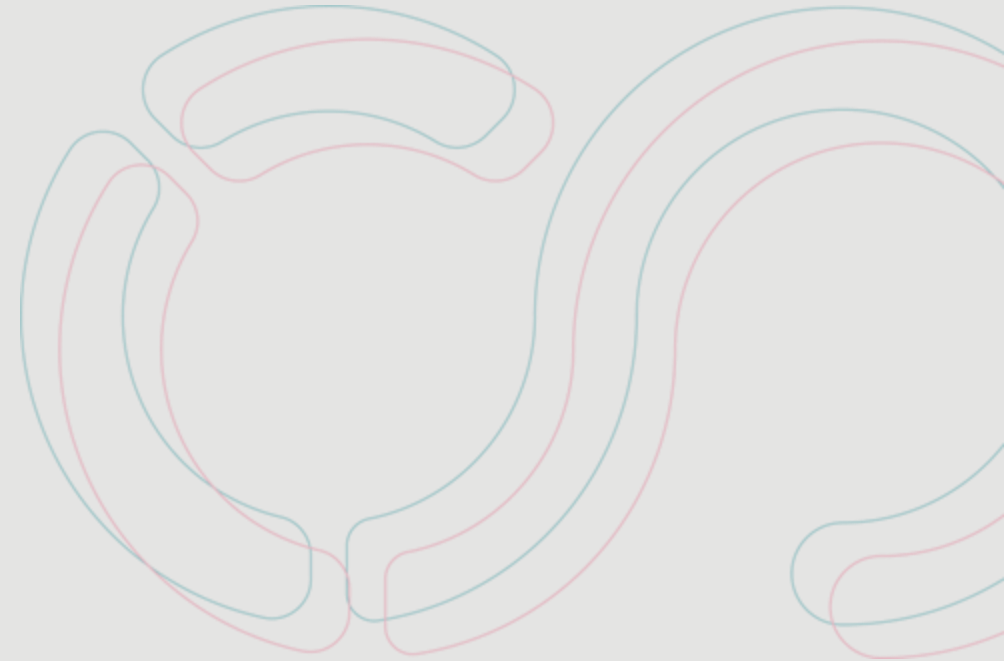
WP5
Development
and operation
of new NRP
services

WP6
Cybersecurity
and Support
Services

WP7
Training and
awareness
raising on NRP
functionaities
and services

NRP Services

High-Level Overview



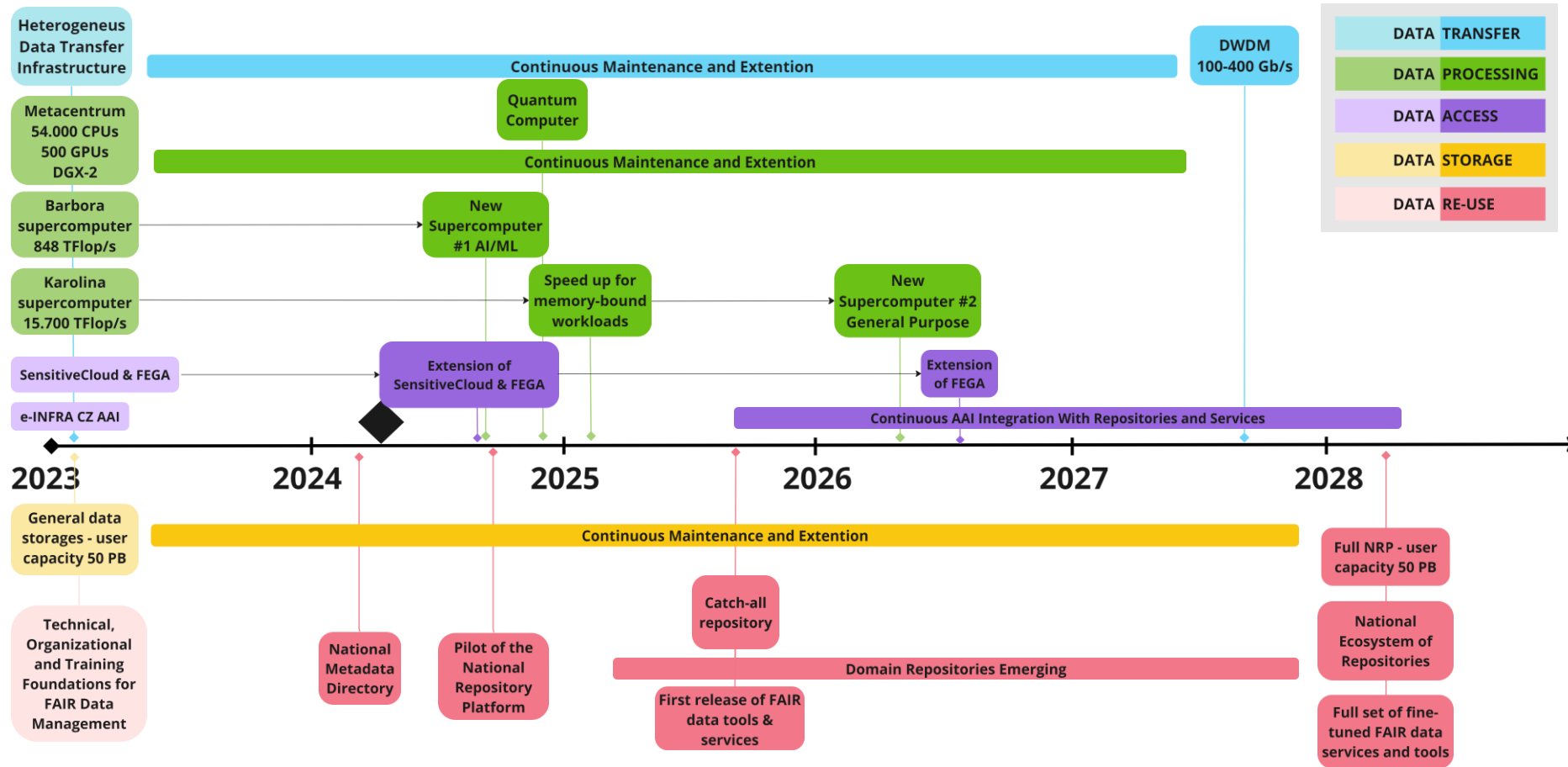
NRP Services – High-Level Overview

- **WP4: Integration of NRP Services**
 - **T4.1 Managing metadata profiles in NRP**
 - Tool for tracking metadata profile compatibility, updates etc. in automatic manner.
 - **T4.2 Support work with licenses**
 - License chooser tool, automatization over repositories etc.
- **T4.3 Data access control**
 - Authentication and authorization all over there.

NRP Services – High-Level Overview

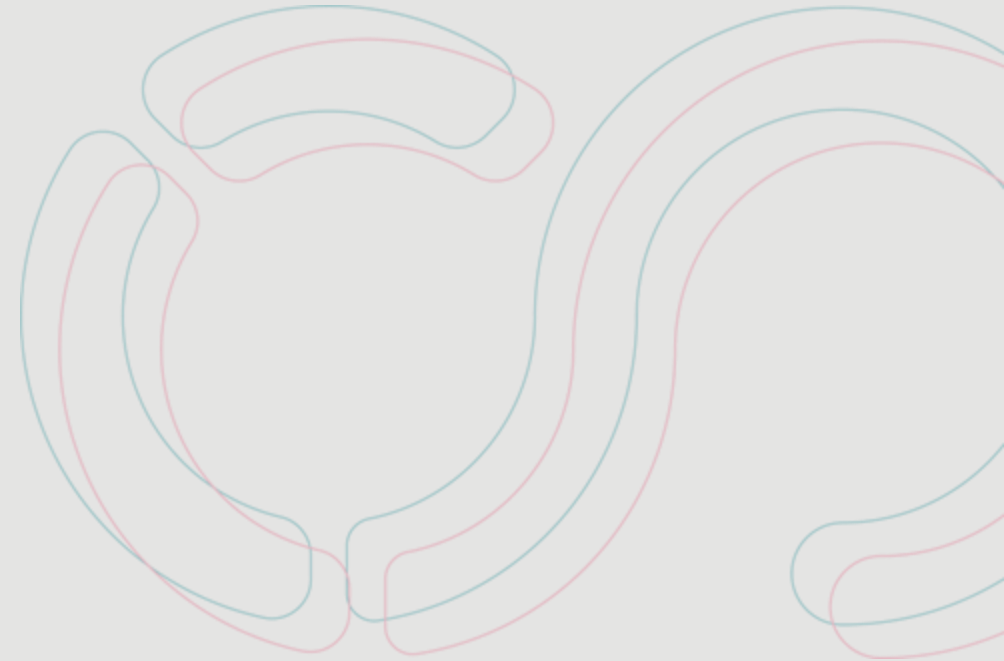
- **WP5: Development and Operation of New NRP Services**
 - **T5.1 Support for FAIRification of research data**
 - Tools for FAIRification support, FIP Wizard profiles for different research areas, for example.
 - **T5.2 Support for data management planning**
 - Integration of DMP-supporting tool like Data Stewardship Wizard.
 - **T5.3 Automation of data and metadata collection**
 - Use cases of connection of research instruments with NRP to get data and metadata from them.
 - **T5.4 Integration of computational workflows**
 - Use cases of using NRP as data source or target for scientific computation in national e-infrastructure.
 - **T5.5 Advanced Services Roadmap**
 - Roadmap of NRP services for end users.

NRP Services Timeline in Wider Context



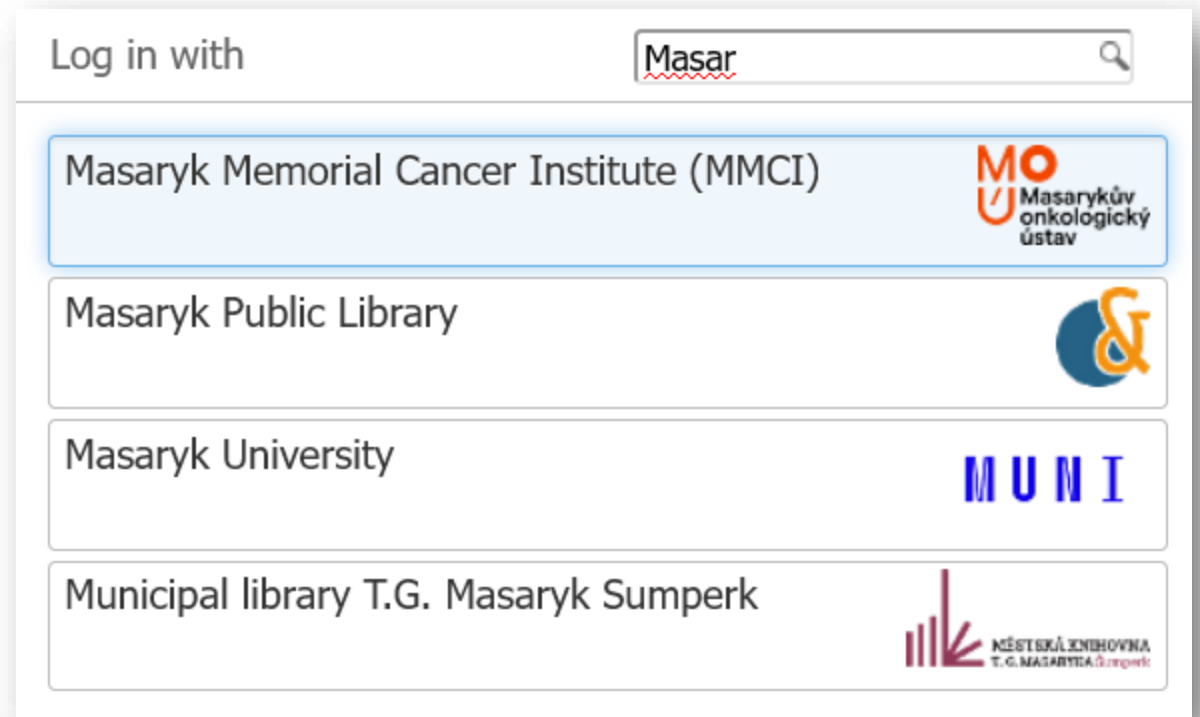
NRP Services

Some in More Details

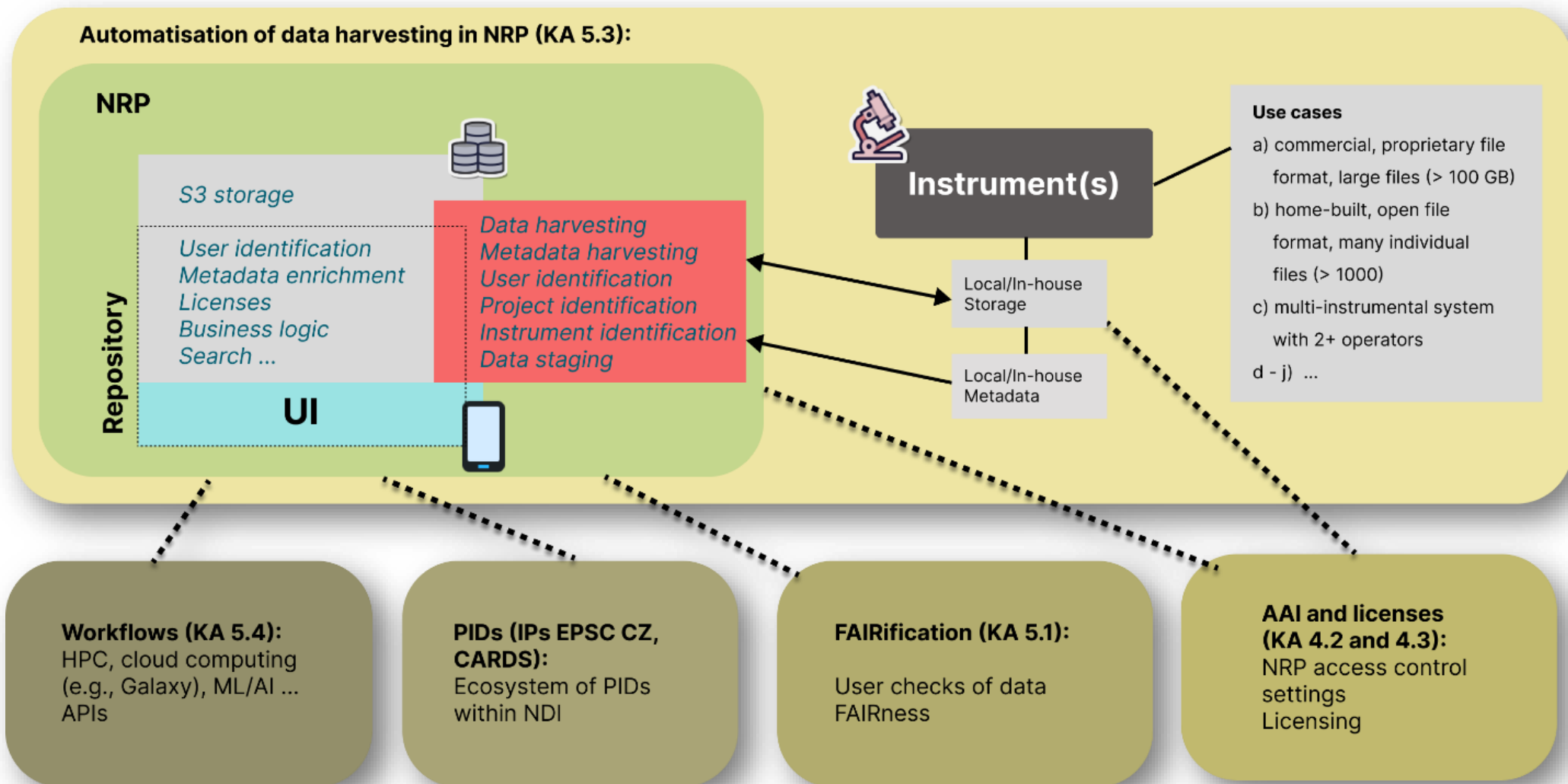


Data Access Control

- FAIR Data, “As open as possible, as closed as necessary”.
 - We need precise access control to data, metadata, services.
- Federative system – use your institutional identity, your well-known authentication web page.
 - Effective collaboration across institutions, individuals, ...

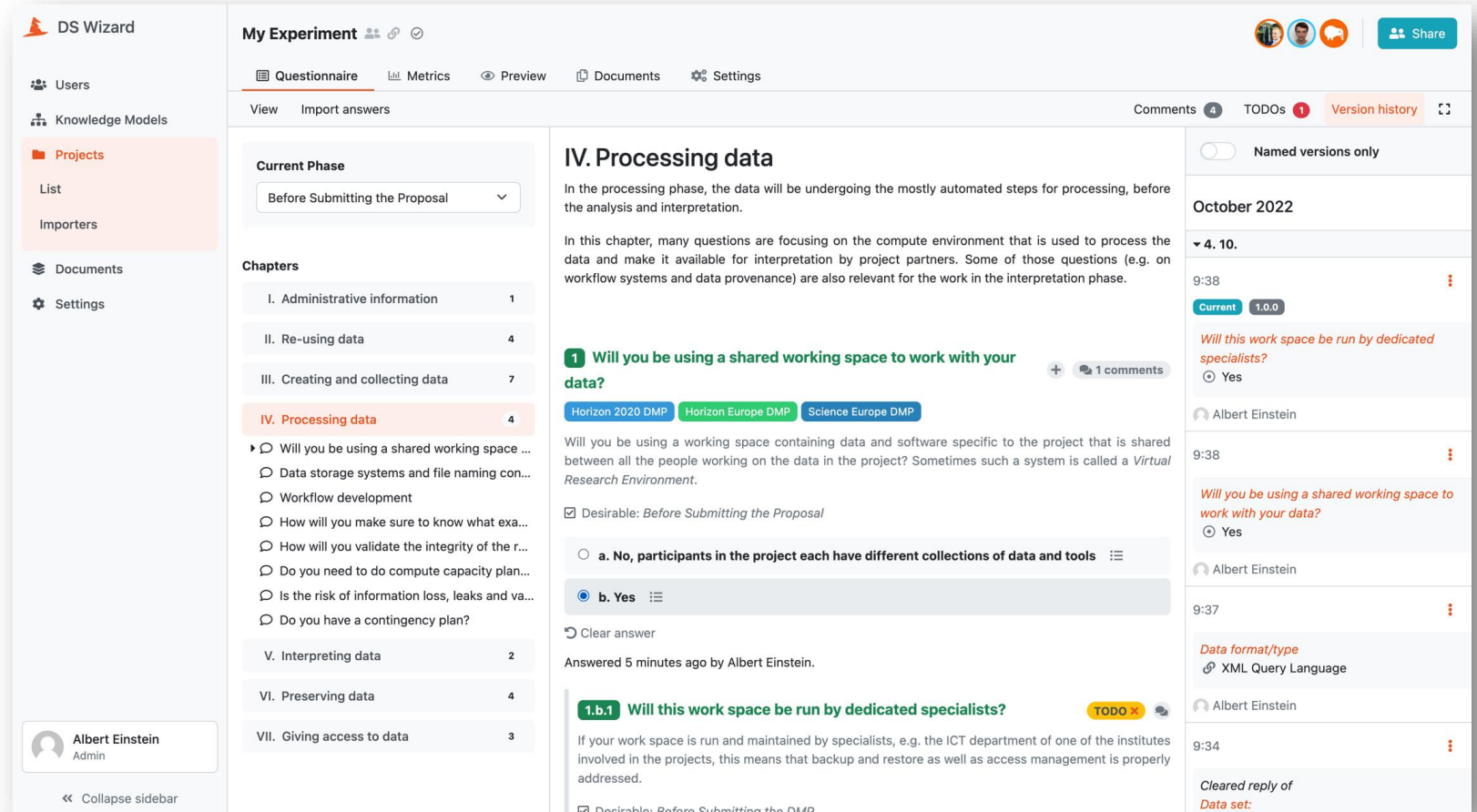


Automation of Data and Metadata Collection



Support for Data Management Planning

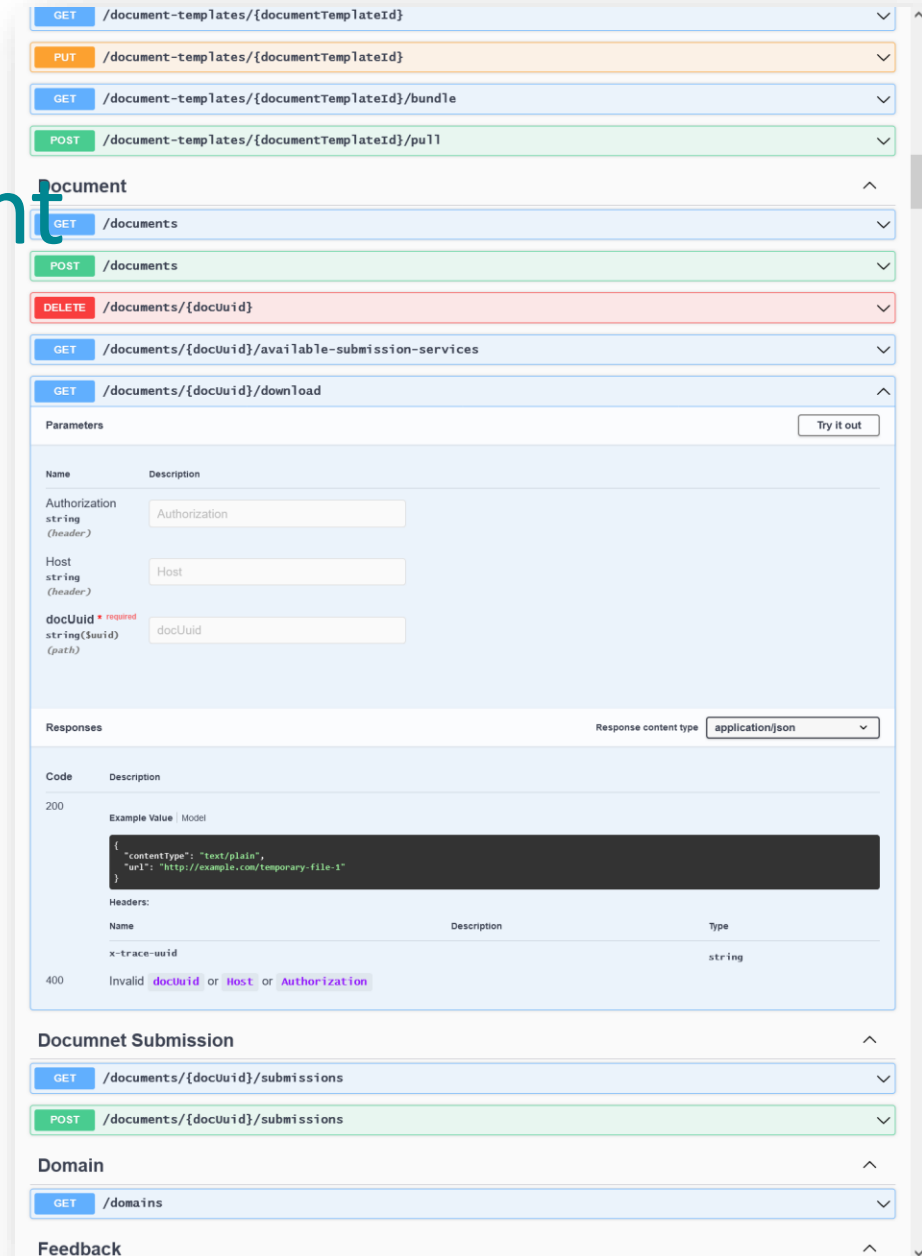
- We expect integration of tools like Data Stewardship Wizard (DSW) directly to the platform.
 - <https://ds-wizard.org/>
- Integration allows effective re-use of available (meta)data.



The screenshot displays the DS Wizard interface for a 'My Experiment'. The left sidebar shows navigation options: Users, Knowledge Models, Projects (List, Importers), Documents, and Settings. The main content area is titled 'My Experiment' and includes tabs for Questionnaire, Metrics, Preview, Documents, and Settings. The 'Questionnaire' tab is active, showing a 'Current Phase' dropdown set to 'Before Submitting the Proposal'. Below this is a 'Chapters' list with items like 'I. Administrative information', 'II. Re-using data', 'III. Creating and collecting data', 'IV. Processing data' (highlighted), 'V. Interpreting data', 'VI. Preserving data', and 'VII. Giving access to data'. The 'IV. Processing data' section contains a question: 'Will you be using a shared working space to work with your data?'. The question has two options: 'a. No, participants in the project each have different collections of data and tools' and 'b. Yes'. A comment by Albert Einstein is visible, stating 'Answered 5 minutes ago by Albert Einstein.' The right sidebar shows a 'Comments' section with a 'Version history' tab, displaying a list of versions for 'October 2022', including '4. 10.' and '1.0.0'.

Support for Data Management Planning

- Automate as much as possible.
 - API and machine-readability for data management plans (DMPs).
- As much as possible read from metadata of the datasets, from infrastructure configuration, ...



The screenshot displays an API documentation interface for document management. It features a list of endpoints with their respective HTTP methods and paths, such as `GET /document-templates/{documentTemplateId}` and `POST /documents`. The interface is organized into sections: Document, Document Submission, Domain, and Feedback. The `Document` section is expanded, showing a table of parameters for the `GET /documents/{docUuid}/download` endpoint. The parameters table includes:

Name	Description
Authorization string (header)	Authorization
Host string (header)	Host
docUuid * required string(docUuid) (path)	docUuid

Below the parameters table, the `Responses` section shows a table of response codes and descriptions. The `200` response is detailed with an example JSON body and a table of headers:

Name	Description	Type
x-trace-uuid		string

The `400` response is described as "Invalid docUuid or Host or Authorization".

Support Work with Licenses

- Templates of **deposition licenses** governing the rules for upload of datasets to repositories.
- **License chooser** for users' datasets on upload.
- Framework for dataset's **license-based access control** to dataset with limited access.
- **Integration and machine actionability.**
 - Connection to data management, access control, ...

Choose a License

Answer the questions or use the search to find the license you want

↻ Start again ← | | →



Is your data within the scope of copyright and related rights?

Yes No

Search for a license...



Public Domain Dedication (CC Zero)

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

Publicly Available   **OPEN DATA**

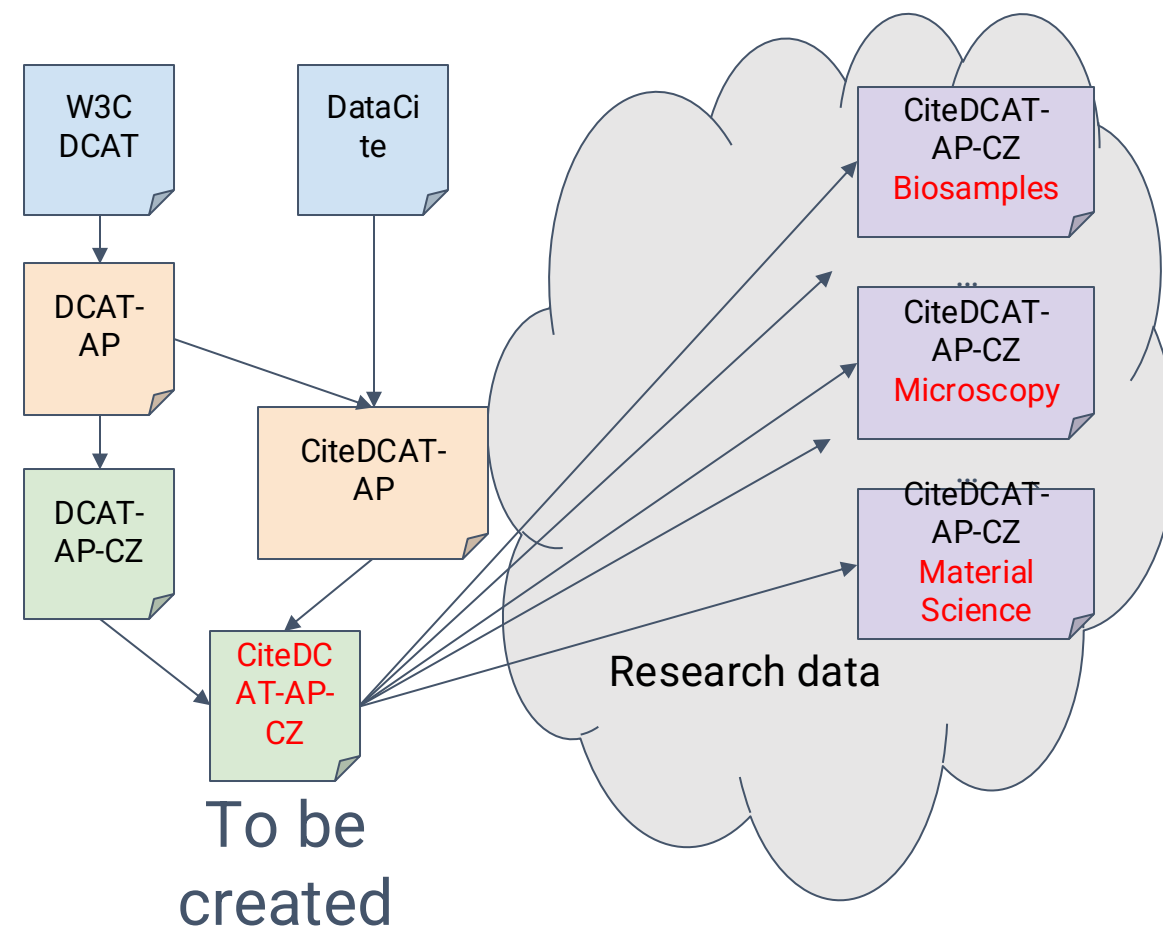
Creative Commons Attribution (CC-BY)

This is the standard creative commons license that gives others maximum freedom to do what they want with your work.

Publicly Available   **OPEN DATA**

Managing Metadata Profiles in NRP

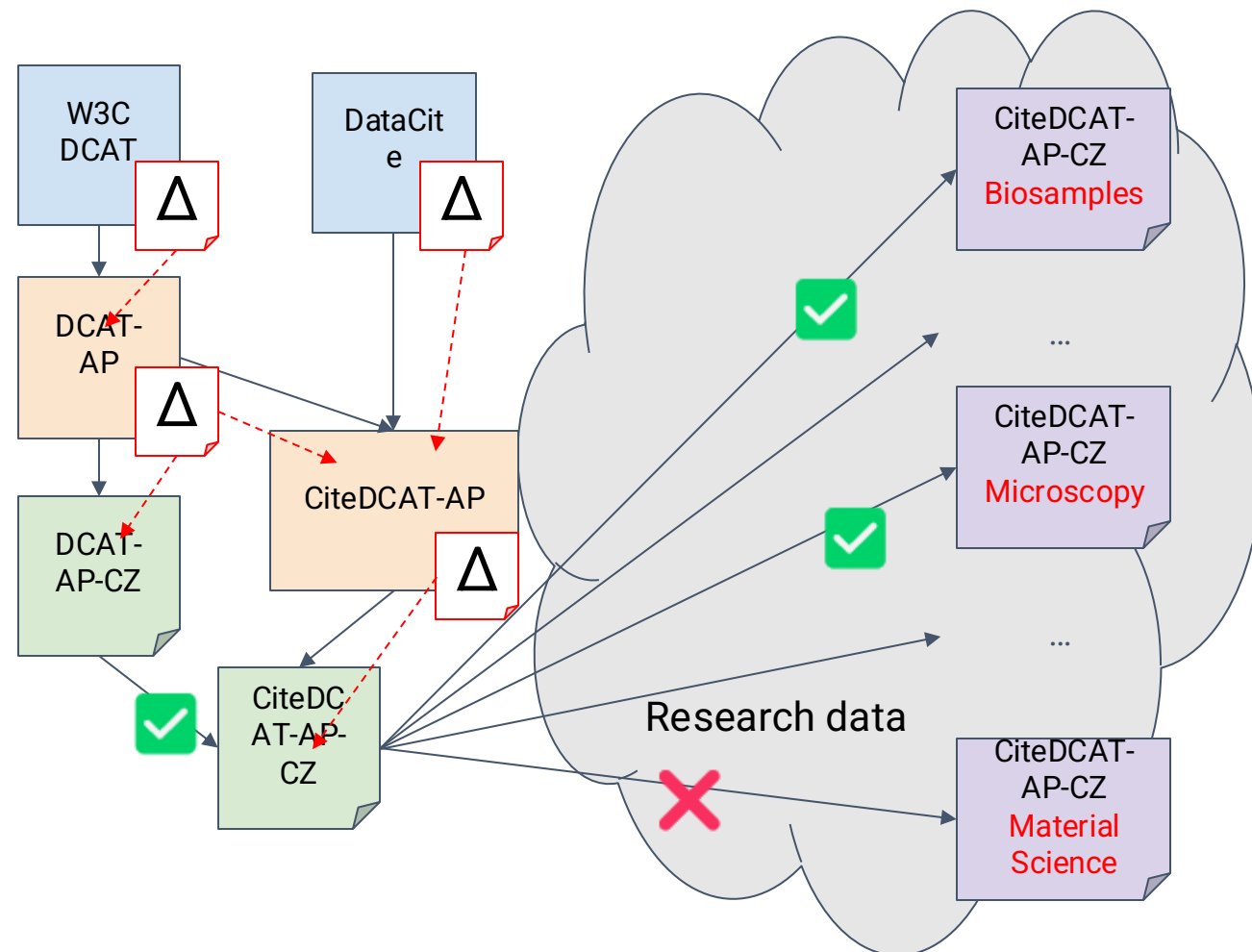
- Research data metadata profiles:
 - What happens, when
 - DCAT v2 → DCAT v3?
 - DCAT-AP v2.1.1 → DCAT-AP 3.0.1?
 - DataCite 4.4 → DataCite 4.5?
- We want changes to be propagated automatically.



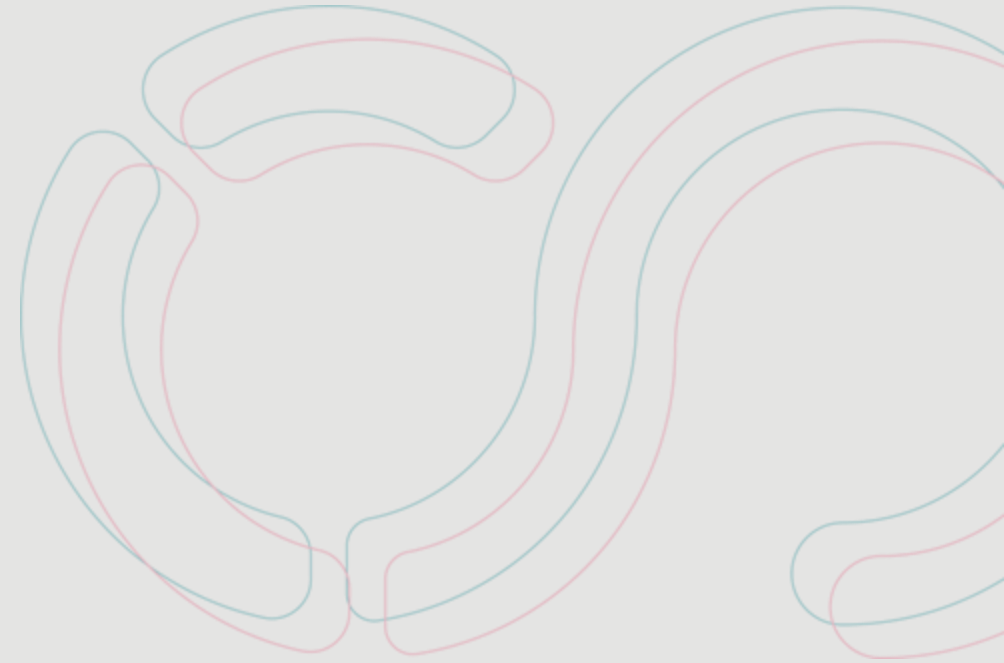
Managing Metadata Profiles in NRP

- But also

- profile compliance validation (✓ ✗),
- description of changes in specifications (Δ),
- change propagation mechanism (→),
- implementation in tools.



Summary



Summary



Source: ELIXIR RDMkit, <https://rdmkit.elixir-europe.org/>

- The main objective of the NRP is to provide research data-related services to its users.
- It focuses on
 - support of FAIR principles and requirements on data management in research projects,
 - security and access control,
 - machine actionability,
 - integrations with other systems.

E: ruzicka@ics.muni.cz

E: info@eosc.cz

W: www.eosc.cz/en



Co-funded by
the European Union

