

# WP 5.2 – Open Science and (FAIR) Data

EURO-LABS Third Annual Meeting  
30/10/2023

A. Lemasson (GANIL), M. Jouvin (IJC Lab),  
C. Hornung (GSI), A. Matta (LPC-Caen)



*This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101057511.*

- Task 1 : Data Management Plan (A. Lemasson -- GANIL)
- Task 2 : openNP Research Data catalog (A. Matta -- LPC Caen)
- Task 3 : New Services for improved access to data (and shared services)
  - 3.1 : Authentication and Authorization Infrastructure (M. Jouvin -- IJC Lab)
  - 3.2 : Prototype of data access platform (C. Hornung -- GSI)

- **Task 1 : Data Management Plan (A. Lemasson -- GANIL)**
- Task 2 : openNP Research Data catalog (A. Matta -- LPC Caen)
- Task 3 : New Services for improved access to data (and shared services)
  - 3.1 : Authentication and Authorization Infrastructure (M. Jouvin -- IJC Lab)
  - 3.2 : Prototype of data access platform (C. Hornung -- GSI)

## Deliverable D5.7

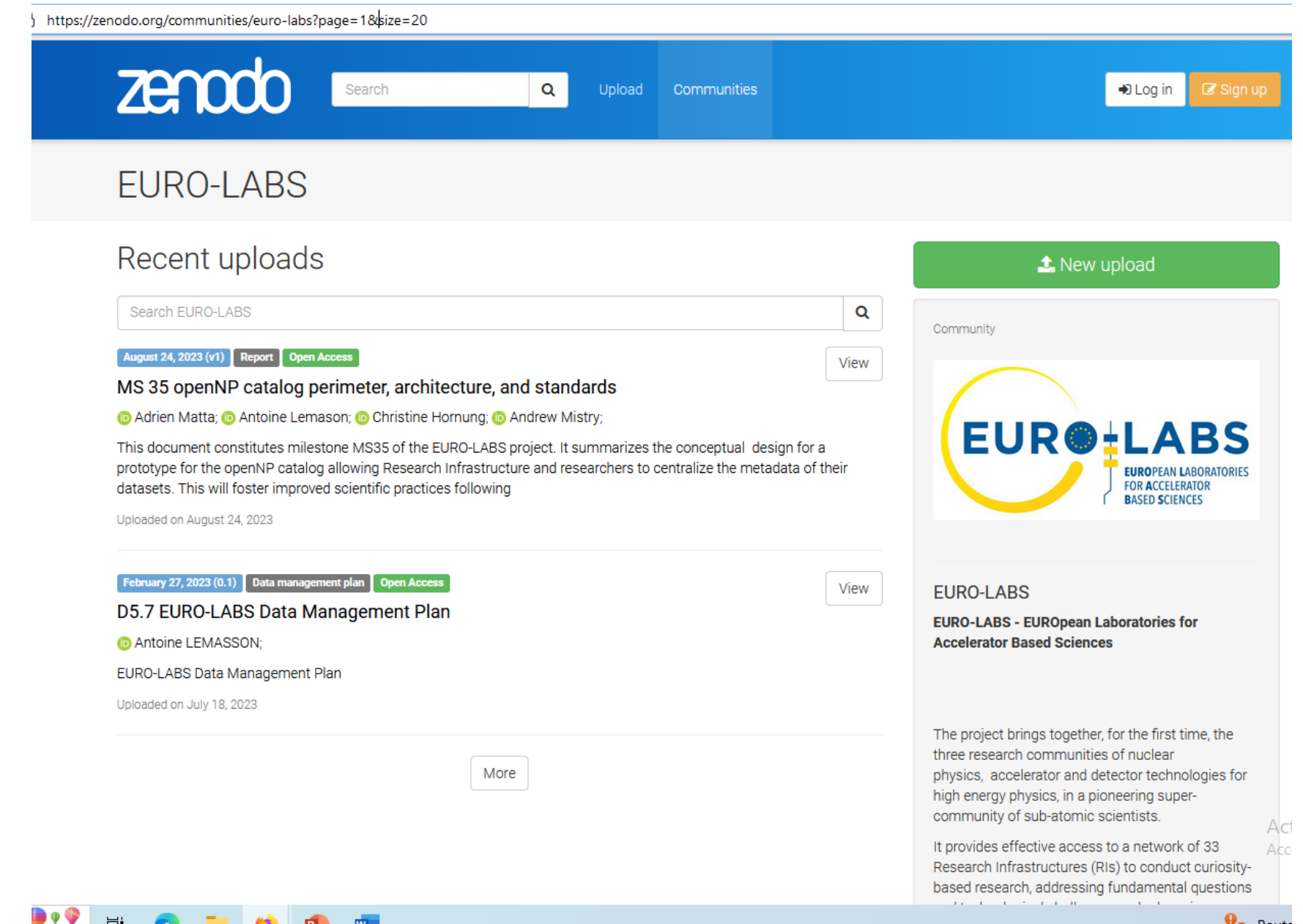
Released on zenodo on Feb 27th 2023 :

Revision on August 2024 -> D5.8

<https://doi.org/10.5281/zenodo.8160229>

- 1) Definition of data produced within the WP of the EURO-LABS project and their management
- 2) Encourage -> **Require** FAIR (Findable, Accessible, Interoperable Reusable) data pinciple by depositing EURO-LABS productions (reports, slides, codes, ...) on zenodo in the EURO-LABS community page to obtain persistent identifier : <https://zenodo.org/communities/euro-labs>
- 3) Data from Research Infrstructures do not fall under the EURO-LABS DMP

DOI 10.5281/zenodo.8160229



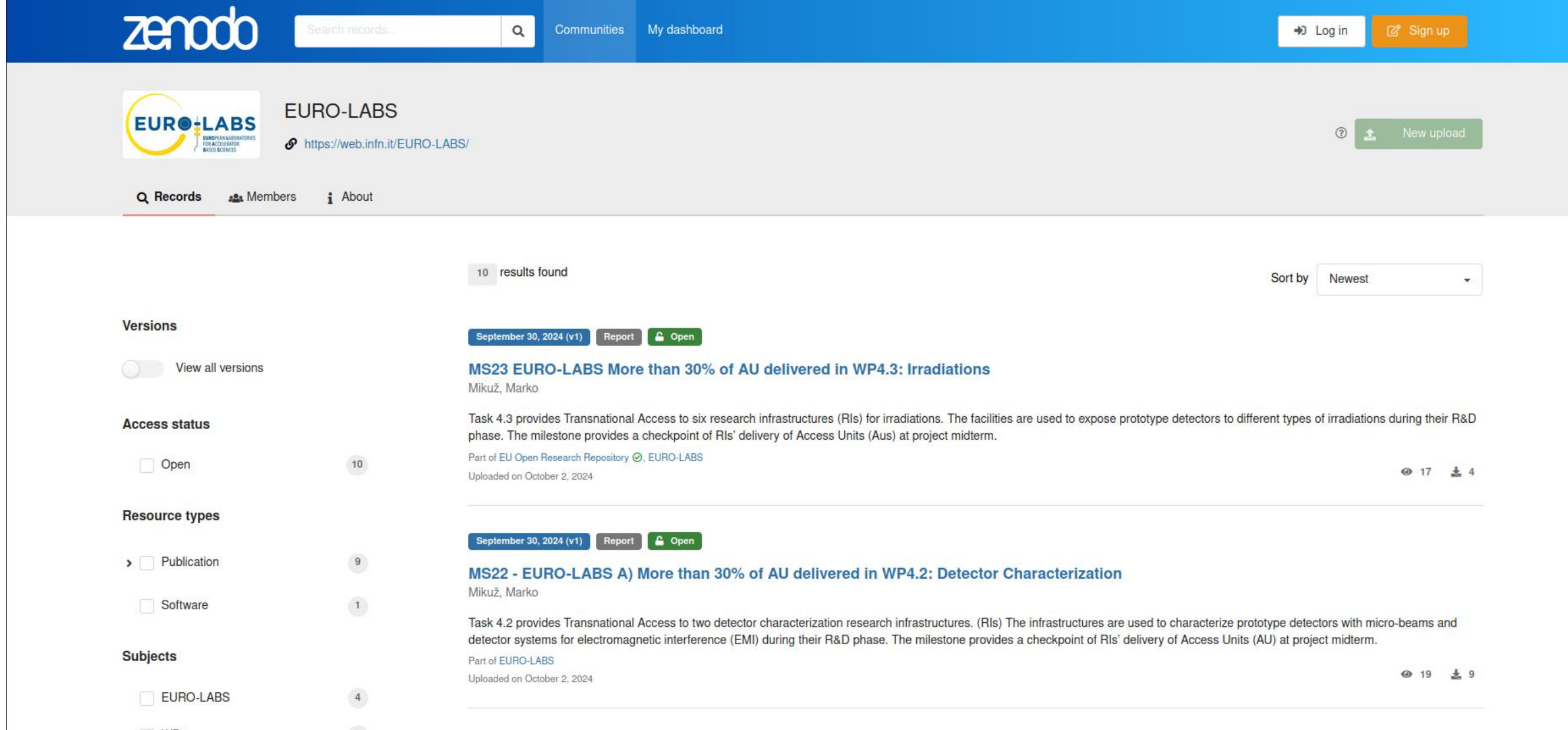
The screenshot shows the Zenodo website interface for the EURO-LABS community. The header includes the Zenodo logo, a search bar, and navigation links for 'Upload', 'Communities', 'Log in', and 'Sign up'. The main content area is titled 'EURO-LABS' and features a 'Recent uploads' section. This section lists two uploads: 'MS 35 openNP catalog perimeter, architecture, and standards' (uploaded August 24, 2023) and 'D5.7 EURO-LABS Data Management Plan' (uploaded February 27, 2023). A 'New upload' button is visible in the top right. On the right side, there is a community profile for EURO-LABS, including its logo and a brief description of the project's goals.

## FAIR data on production from EURO-LABS

Deposit of the projects related data on the EURO-LABS zenodo repository :

- Milestones, Deliverables (Barbara)
- Slides / Posters
- Source Codes Releases (Geoff, ...)
- Lectures of EURO-LABS Schools ....

<https://zenodo.org/communities/euro-labs/>



The screenshot shows the Zenodo repository page for the EURO-LABS community. The page features a search bar, navigation links for 'Communities' and 'My dashboard', and a 'Log in' / 'Sign up' button. The main content area displays a list of records, with two records visible: 'MS23 EURO-LABS More than 30% of AU delivered in WP4.3: Irradiations' and 'MS22 - EURO-LABS A) More than 30% of AU delivered in WP4.2: Detector Characterization'. Each record includes a date (September 30, 2024), a version number (v1), and buttons for 'Report' and 'Open'. The page also shows a sidebar with filters for 'Versions', 'Access status', 'Resource types', and 'Subjects'.

- Task 1 : Data Management Plan (A. Lemasson -- GANIL)
- Task 2 : openNP Research Data catalog (A. Matta -- LPC Caen)
- Task 3 : New Services for improved access to data (and shared services)
  - 3.1 : Authentication and Authorization Infrastructure (M. Jouvin -- IJC Lab)
  - 3.2 : Prototype of data access platform (C. Hornung -- GSI)



## openNP is a catalog of

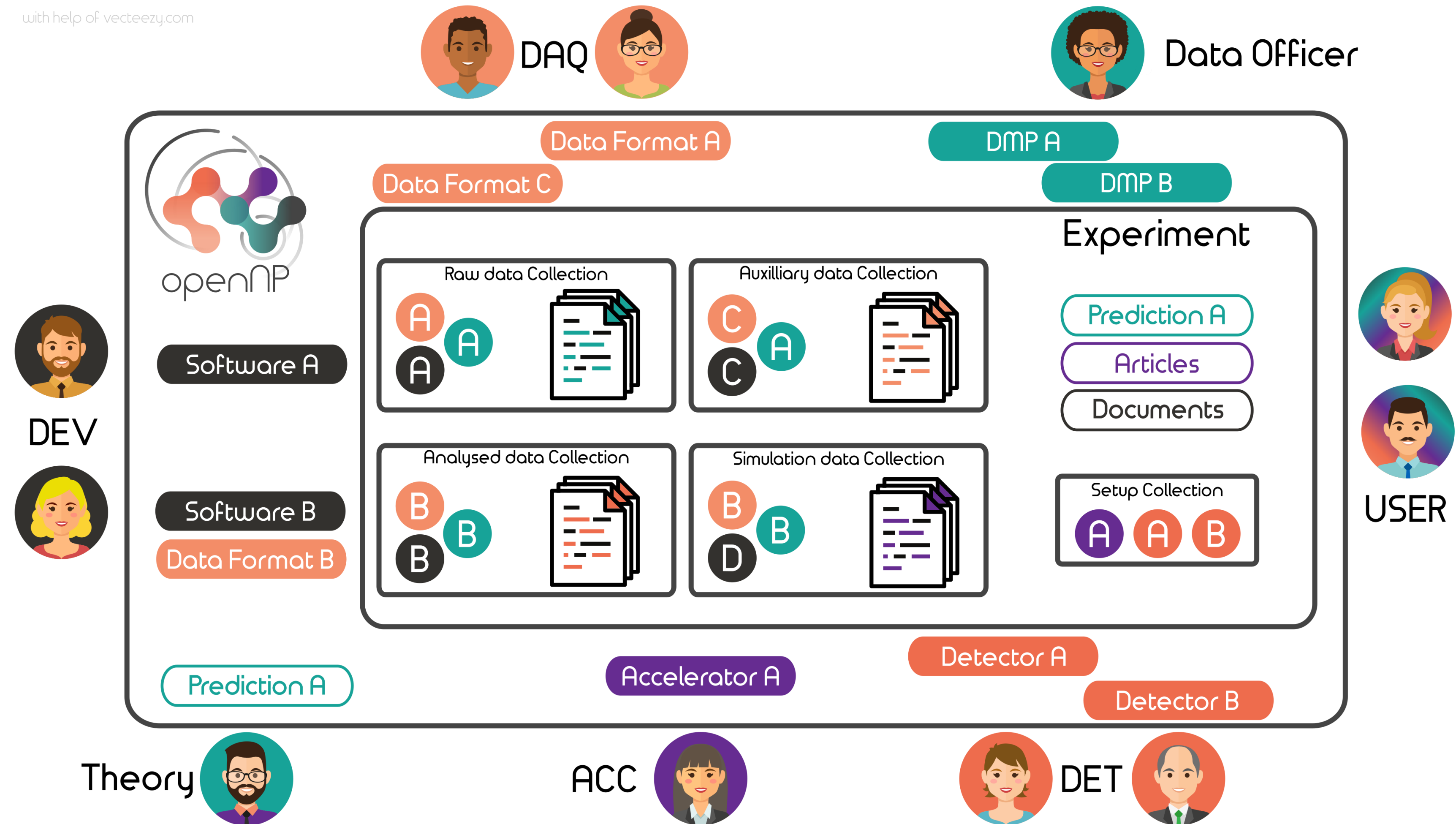
- experimental datasets from RI
- auxiliary Data (elog, ...)
- DMP
- DAQ softwares (data format, readout, ...)
- Analysis Softwares
- Experimental setup descriptions
- Analysed data

## Involving

- Researchers,
- Data Officers,
- IT Departments,
- ...

Do not intend to store the large experimental datasets

with help of vecteezy.com





## Milestone MS35 openNP catalog perimeter, architecture, and standards

Released on zenodo Aug 24th 2023: <https://doi.org/10.5281/zenodo.8279798>

DOI 10.5281/zenodo.8279798

### 1) Definition of the data type and metadata fields

Beams, Accelerator, Experimental Setup, DataSet, Softwares, ... => Subject to evolution

Synergies with ongoing activities on metadata standardization at GSI and PUNCH4NFDI project

(C. Hornung, A. Mistry *et al.* GSI)

➔ Initiated new Project NAPMIX with OSCARS Funding

### 2) 08/2023 => openNP catalog perimeter => Initially Dataverse was considered

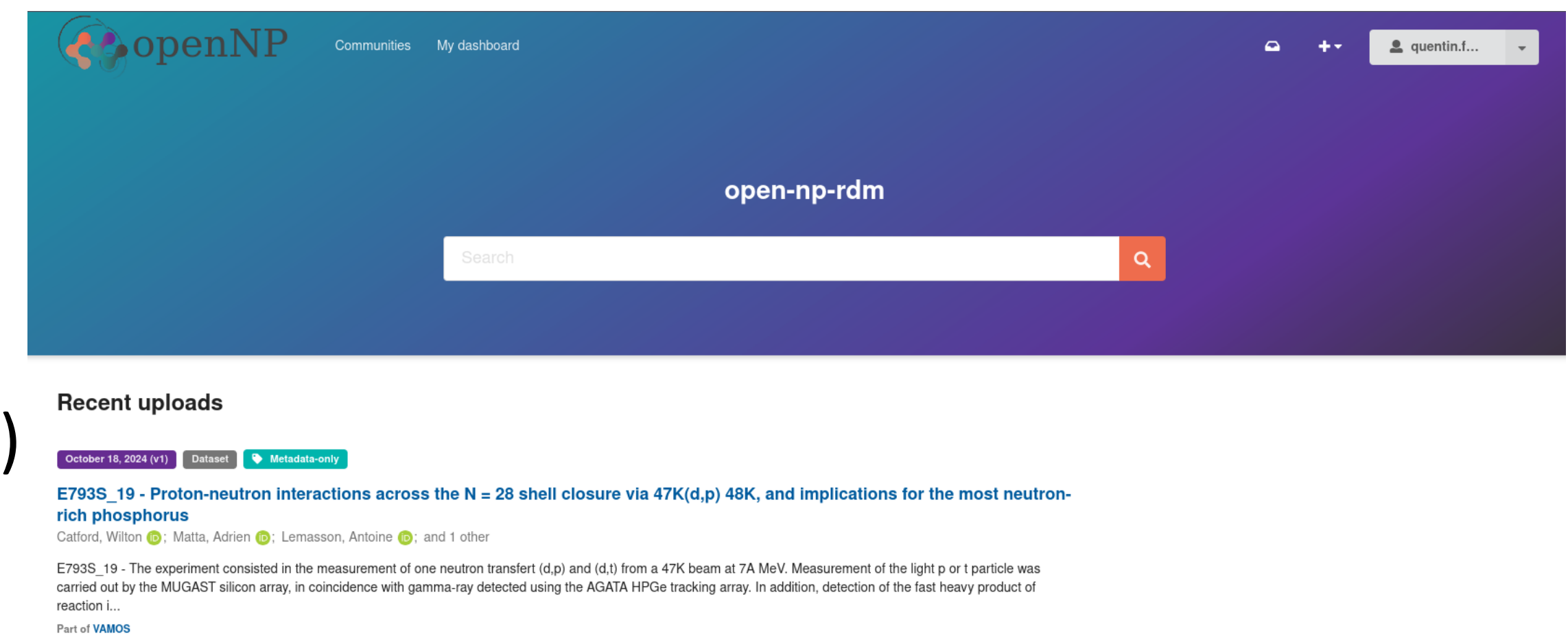
### 3) Since July 2024 : Q. Fable in charge of the developpement at GANIL

- Evaluaton of the different plateform (invenio-rdm, Dataverse)
- Chosen plateform : Invenio-rdm
- Prototype locally

=> Poster of Quentin Fable

### 4) Prototype release expected S1 2025

- Integration of specific vocabularies (NAPMIX)
- Prototype tests and adjustment with the help of RI
- Deliverable is final service in M36





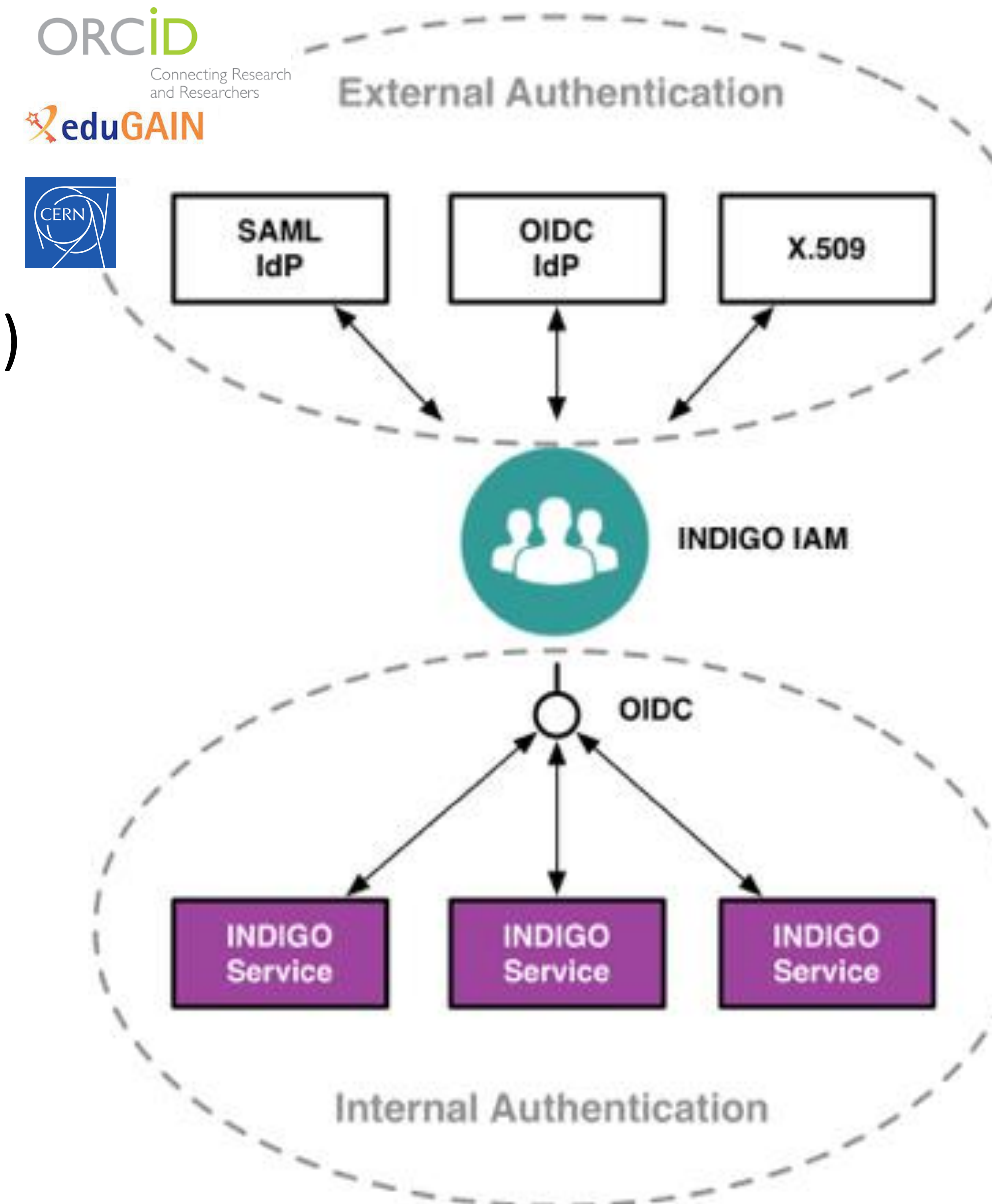
- Task 1 : Data Management Plan (A. Lemasson -- GANIL)
- Task 2 : openNP Research Data catalog (A. Matta -- LPC Caen)
- Task 3 : New Services for improved access to data (and shared services)
  - 3.1 : Authentication and Authorization Infrastructure (M. Jouvin -- IJC Lab)
  - 3.2 : Prototype of data access platform (C. Hornung -- GSI)

## Task 3.1 (M. Jouvin IJCLab)

### Common identification and authentication service

- Identification using existing federations (eduGAIN, ...) using single sign on (SSO) or ORCID
- Easy registration of new services (compared to registration directly to the federation)
- Management of users based on groups to handle authorizations levels on the service side
- Benefiting from the experience gained from ESCAPE

=> INDIGO IAM



Service was released in June 2023 (you can register now)

<https://iam-eurolabs.ijclab.in2p3.fr>

Already used by

- Theory4 Exp virtual access theory platform (s) (Manuela Rodriguez-Gallardo)
- Grafana service identification at GANIL

=> presently > 200 users

If such a service may seem useful to you, please contact us

Large developpement in the framework of Theory 4 EXP

## What is needed now ?

- > build an operation committee with members from the various labs to manage the service (approval of registration, group management), define policies so far mostly handled by :  
M. Rodriguez Gallardo, M. Jouvin and A. Lemasson




INDIGO - DataCloud

Welcome to **indigo-dc**

Sign in with

eduGAIN 

Local credentials

- Task 1 : Data Management Plan (A. Lemasson -- GANIL)
- Task 2 : openNP Research Data catalog (A. Matta -- LPC Caen)
- Task 3 : New Services for improved access to data (and shared services)
  - 3.1 : Authentication and Authorization Infrastructure (M. Jouvin -- IJC Lab)
  - 3.2 : Prototype of data access platform (C. Hornung -- GSI)

## Task 3.2 (C. Hornung, GSI)

Prototype of data access platform

### Test of different data access platform prototypes (Data Lakes)

Different data lake prototypes are being developed/tested at the GSI.

- For the CERN data lake, the GSI contributes two Rucio Storage Elements (RSEs). This data lake was build and tested within the WP2 in the ESCAPE project.
- Another data lake that span three participating sites, GSI, Karlsruhe Institute of technology (KIT) and the university of Groningen (RUG) was also build and tested within the same ESCAPE project.
- A further prototype based on XRootD is also under investigation within the activities of the PUNCH4NFDI project.

### Update Since SAM in Krakow :

- The GSI-FRS group has prepared a server to test distributing data to external participants.
- Collaboration with the GSI-IT group has started to access to Lustre (GSI/FAIR cluster filesystem) via XRootD and a web single page app (SPA).
- Authentication will be done using Keycloak and a GSI Web Login, and either the GSI Linux account or a granting to specific experiments by those not with a GSI Linux account. Need to be have linked capabilities with IAM-EUROLABS

The results from these prototypes combined with the other related activities like AAI and caching will lead to a federated infrastructure concept that will be implemented and used for FAIR and could be implemented in other Research Infrastructures.

**Deliverable is a prototype of platform by M36.**

**Nic Hubbard (GSI) started this year**

Second Annual Meeting - EURO-LABS – A. Lemasson

# EURO-LABS Advanced Training: Open Science and Data Management

Local organizers:

Andrew Mistry (GSI/FAIR), Christine Hornung (GSI), Gerhard Bura (HGS-HIRe)

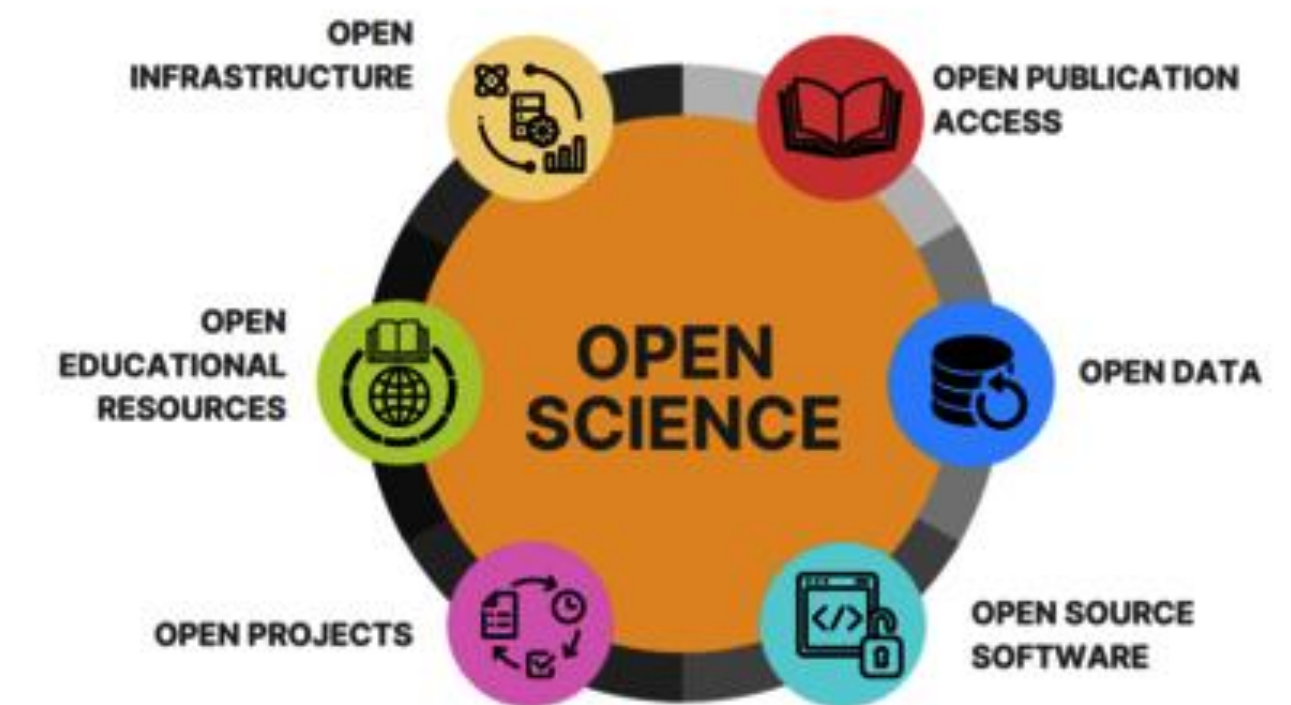
## Speakers :

- Antoine Lemasson (CNRS - GANIL) – **Introduction to Open Science**
- Florian Uhlig (GSI) – **Tools for sustainable programming**
- Özlem Özkan (Helmholtz Metadata Collaboration) – **Metadata for beginners**
- Andrew Mistry (GSI) – **European infrastructures for Open Science**
- Harry Enke (AIP) and Elena Sacchi (AIP) – **Astroparticle/PUNCH4NFDI**
- Clemens Lange (PSI) – **Open Science in HEP**
- Kathrin Göbel (GSI) – **Open technology transfer**
- Adrien Matta (LPC Caen - CNRS) and Jérémie Dudouet (IP2I Lyon - CNRS)  
– **Hands-on data challenge**
- Christoph Scheidenberger (GSI) and Christine Hornung (GSI)  
– **Excursion to visit the GSI Helmholtzzentrum für  
Schwerionenforschung and FAIR**

Dates : 24. - 29.11.2024

=> 18 confirmed participants

<https://indico.gsi.de/event/19808/>



**G S I**

**HGS-HIRe for FAIR**  
Helmholtz Graduate School for Hadron and Ion Research



Photo: Walter P. Lhotzky

- **Task 1 : Data Management Plan (A. Lemasson -- GANIL)**  
=> **Delivered 02/2023 / Updated 08/2024**  
=> Encourage deposing EURO-LABS output to zenodo (in addition to EURO-LABS website)  
=> Encourage to establish Data Management Plan for data produced at all RI  
=> If needed please contact us for discussion about DMP at your infrastructure
- **Task 2 : openNP Research Data catalog (A. Matta -- LPC Caen)**  
=> **MS 35 Delivered**  
=> Deliverable : Catalog to be released by M36
- **Task 3 : New services for improved access to data (and shared services)**
  - 3.1 Authentication and Authorization Infrastructure (M. Jouvin -- IJC Lab)  
=> **Service delivered, already used within EURO-LABS task 2.4**  
=> Policies and Management to be constructed with the different facilities
  - 3.2 Prototype of data access platform (C. Hornung -- GSI)  
=> Deliverable : Prototype of data platform to be released by M36

Research Data Management is not a threat but an opportunity for RI and researchers.  
A lot of tools already exists, that needs to be made available to the community.  
Chosing the right tools and initiate change of habits !