#### ExPaNDS European Open Science Cloud Photon and Neutron Data Services

# **TCB** meeting

Paul Millar (DESY) Technical Coordinator

2023-01-23



# Agenda

- Introduction (10')
  - Setting the scene (goal of this meeting)
  - Reviewing the activity inventory along with possible future work.
- Brainstorming session on future work (30')
- Discussion on available funding sources (10')
- Wrap-up and next steps (10')





#### Introduction

• The Technical Coordination Board is ... "responsible for [...] analysing the market offering of both commercial and open source solutions that can represent opportunities or **new technical collaboration for the project**."

ExPaNDS is now ending, with only the final deliverable due.

• ExPaNDS has achieved considerable work, but there is more needed to **fully realise** the promise of FAIR data.

Communication is very important – something I suggest we continue through LEAPS WG3 , along with LENS facility's participation.

• The goal of this meeting is to start a document, in which we highlight useful future work and possible ways to fund those developments.



#### **Output inventory**





## **WP2: Outcomes summary**

- (FAIR Data) Policy Framework
- (FAIR) Data Management / PIDs
- (FAIR) Data Management Plans: DMPs
- (FAIR) Self assessment





## **WP3: Outcomes summary**

- Ontologies: PaNET, NeXus
- OAI-PMH interfaces (SciCat and ICAT)
- PaN federated search (support, deployment)
- Ontology API (supporting PaNET expansion in PaN search API)
- Landing page optimisation for Google Dataset Search





## **WP4: Outcomes summary**

- Reference datasets
- Packaged workflows: JupyterHub or containers
- Testing and validation framework (CI/CD)
- Sharing experience with federated authentication (i.e., UmbrellaID)
- VISA deployments





#### **Possible developments**

(just some ideas to kick-start the discussion)





# **WP2: Possible developments**

- (FAIR Data) Policy Framework
  - Update, based on feedback from facilities
  - Align with related policy frameworks
- (FAIR) Data Management / PIDs
  - Share deployment/adoption stories between RIs  $\rightarrow$  Maybe update document
  - Build collaborations to support concrete elements (e.g., Instrument PID adoption).
- (FAIR) Data Management Plans: DMPs
  - Adoption stories, use-cases, circulate benefits.
  - Improve automatisation and integration with RIs.
- (FAIR) Self assessment
  - Simultaneous, periodical, cross-RI surveys assessments.
  - Build feedback loops to update assessment.
  - Explore FAIR assessment tools; e.g., F-UJI





# **WP3: Possible developments**

• Ontologies: PaNET, NeXus

Seek adoption (e.g., user-office using PaNET). Feedback experiences (missing terms in PaNET)

• OAI-PMH interfaces (SciCat and ICAT)

Investigate providing more PaN-specific metadata (e.g., PaNET terms), allowing metadata harvesters to include PaN specific attributes in queries.

• PaN search API (support, deployment)

Deploy support; Build or extend tools to integrate with PaN Search API.

- Ontology API (supporting PaNET expansion in PaN search API) Investigate other aspects of searching that benefit from expansion
- Landing page optimisation for Google Dataset Search Increase adoption. Check what other Schema.org metadata may be provided
  - Sign-posting in landing pages





# **WP4: Possible developments**

• Reference datasets:

Provide additional reference datasets; register DOIs for these datasets; update software and tutorials to take advantage of these datasets.

• Packaged workflows: JupyterHub or containers

Move this operation from heroic to routine; investigate workflow catalogues (e.g., galaxy); more work on standardising platforms; investigate using EOSC resources.

• Testing and validation framework (CI/CD)

Increase adoption at different sites; investigate using this when deploying production software.

- Sharing experience with federated authentication (i.e., UmbrellaID) Increase adoption; Establish a group to share experiences.
- VISA deployments





#### OK, let's start talking ...

See "live" document:

https://docs.google.com/document/d/1jLrCUTOgKDk4EE2KNjWR5jzo-uZanX88lZRD 8ENZrz4/edit?usp=sharing





#### **Possible funding**

#### (just some ideas to kick-start the discussion)





# **Possible funding**

- Direct funding from RIs
  - Ad hoc collaborations, perhaps coordinated through LEAPS+LENS.
  - Majid's work on establishing two MoUs between RIs seems relevant.
- Horizon Europe calls:
  - INFRA-EOSC-2023-01-01: cascading (2 rounds) of funding for small projects (€100k-€250k; ~100 projects). See Sophie's talk tomorrow.
  - INFRA-EOSC-2023-01-02: software/analysis development
  - INFRA-EOSC-2023-01-03: active DMPs and PID-graph.
  - Any others?
- National funding;
  - NFDI (DAPHNE4NFDI), Helmholtz (HMC)





#### OK, let's start talking ...

See "live" document:

https://docs.google.com/document/d/1jLrCUTOgKDk4EE2KNjWR5jzo-uZanX88lZRD 8ENZrz4/edit?usp=sharing





#### Wrapping up





## Next steps

- The Google document is currently open for people to contribute.
- 2023-02-03 (Friday, in a couple of weeks) I'll close the document for direct editing.
  - You'll still be able to comment with just the link.
  - Just let me know if you would like to edit the document directly.
- We'll continue working on the document: identifying possible future work, who would do that work and how they would be funded.
- We anticipate information from this document will be folded into the final report.
- Hopefully this will trigger future collaborations.





#### Thanks!





## **Overview of the technical WPs**



WP2: FAIR data principles



WP3: FAIR data catalogues



#### WP4: FAIR data analysis





#### WP2: FAIR data principals









#### **Steps towards FAIR Photon and Neutron Facilities**

Slide from **Brian Matthews** 

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857641

21

## WP2: FAIR data policy framework





22

................

Results

Subset of processed

data and other

outcomes

## **WP2: Persistent Identifiers (PIDs)**







# WP2: FAIR data management



NO. BORNER AND A DESCRIPTION

C. E. et al. to prove the

#### **WP2: Data Management Plans**



## **WP2: FAIR self-assessment**



D2.6 Self-evaluation Photon and Neutron RIs for FAIR data certification

Settings	Value
Document Identifier:	D2.6
Project Title:	ExPaNDS
Work Package:	WP2
Work Package Lead	UKRI
Document Author(s):	Simon Lambert (UKRI), Abigail McBirnie (UKRI), Brian Matthews (UKRI), Stephen Collins (DLS), Maik Fielder (HZDR), Brighte Gagey (SOLEL), Heike Görzig (HZB), Oliver Knodel (HZDR), George Kourousias (Elettra), Anders Markvardsen (ISIS), Carlo Minotti (PSI), Majid Ounsy (SOLEL), Krisztian Pozsa (PSI), Silvan Schön (DESY), Nicolas Soler (ALBA), Darren Spruce (MAX IV), Andrei Vukolov (Eletra)

- How do we know that the policies, standards and tools we have put in place will mean we achieve data which people can find, access and use?
- FAIR Assessment
  - Test your data to see if it satisfies the FAIR Principles
- But we want to be sure that every experiment results in FAIR data
  - Test your process to see if it includes the right things to make the data FAIR
- A self assessment method for facilities to ask the questions
   29 questions to test the FAIR-Ness of the facility science cloud Photon



#### WP3: FAIR data catalogues







## WP3: PaNET ontology



<ul> <li>IRI: http://purl.org/pan-science/PaNET/PaNET</li> <li>Source <ul> <li>https://en.wikipedia.org/wiki/Small-angle</li> </ul> </li> <li>has super-classes <ul> <li>c h m o 0000182 <sup>c</sup></li> <li>diffraction <sup>c</sup></li> <li>low modes scattering <sup>c</sup></li> <li>q133900 <sup>c</sup></li> </ul> </li> <li>has sub-classes <ul> <li>anomalous small angle x-ray scattering angle scattering <sup>c</sup></li> <li>grazing incident scattering <sup>c</sup></li> <li>inelastic small angle scattering <sup>c</sup></li> <li>small</li> </ul> </li> </ul>	k to <u>ToC</u> or <u>Class ToC</u>
Source https://en.wikipedia.org/wiki/Small-angle has super-classes <u>c h m o 0000182 <sup>c</sup>, diffraction <sup>c</sup>, low mo scattering <sup>c</sup>, <u>q133900 <sup>c</sup></u> has sub-classes <u>anomalous small angle x-ray scattering</u> <u>angle scattering <sup>c</sup>, <u>grazing inciden</u> <u>scattering <sup>c</sup>, inelastic small angle scattering</u></u></u>	01124
has super-classes <u>c h m o 0000182</u> <sup>c</sup> , diffraction <sup>c</sup> , low mo <u>scattering</u> <sup>c</sup> , <u>q133900</u> <sup>c</sup> has sub-classes <u>anomalous small angle x-ray scattering</u> <u>angle scattering</u> <sup>c</sup> , <u>grazing inciden</u> <u>scattering</u> <sup>c</sup> , <u>inelastic small angle scattering</u> <u>angle inelastic scattering</u> <sup>c</sup> , <u>small</u>	_scattering
anomalous small angle x-ray scattering angle scattering <sup>c</sup> , <u>grazing</u> inciden scattering <sup>c</sup> , <u>inelastic</u> small angle scat	mentum transfer
scattering <sup>c</sup> , small angle x-ray scatter small angle scattering <sup>c</sup> , ultra small ang	y <sup>c</sup> , <u>diffuse small</u> ce small angle ceringng <sup>c</sup> , <u>small</u> angle neutron ing <sup>c</sup> , <u>spin echo</u> e scattering <sup>c</sup>







#### WP3: OAI-PMH



Slide thanks to Carlo Minotti

E x P a N European Open Science C

and Neutron Data Service





0

## WP3: PaN Search API



Slide thanks to Carlo Minotti



\*\*\* \* \* \*\_\*

# WP3: Google dataset search



Slide thanks to Carlo Minotti





# WP3: Ontology API



#### WP4: FAIR data analysis







#### **WP4: nine reference datasets**





# WP4: Working with workflows

• Identified five workflows to act as exemplars:

Serial crystallography, THz Spectroscopy, Full-field Tomography, Ptycho-tomography, Small-Angle Neutron Scattering.

Packaging workflows to allow execution of workflows outside of facility

Using containers or Jupyter notebooks, as appropriate

- Demonstrating execution of these workflows on multiple infrastructures
- Build a testing and validation framework for CI/CD





## WP4: the practical bits....

- On-boarding of (remote) analysis services within EOSC Marketplace
- Helping facilities adopt federated authentication (e.g. Umbrella)
- Work focusing on VISA platform
  - Multiple facilities agreed on adopting VISA.
  - Sharing use-cases with VISA development team.
  - Sharing expertise, within the project, in VISA deployment
  - Site-local VISA integration effort.











37

## Conclusions

- Significant body of work
- Contributions at several levels:
  - Policy
  - Guidelines
  - Ontologies
  - Software development
  - EOSC onboarding
  - Improving analysis workflows
  - Service integration

#### Strong impact and sustained contribution



