

The background features a complex network diagram with numerous nodes and connecting lines. The nodes are represented by circles in light blue, yellow, and red. Interspersed among the nodes are binary digits '0' and '1' in various colors (blue, yellow, red, grey). The overall aesthetic is technical and data-oriented.

ExPaNDS WP3 Status Update ALBA

*18th of June 2021
Daniel Salvat*

Prerequisites

- Data Infrastructure → Yes
 - Facility Proposal Info → Yes
 - Standard Data → no at this point (?)
 - PID Issuer → WiP
 - Data Policy → Yes, since 2017. May need review.
 - Resources
 - 1 Software Engineer (ExPaNDS)
 - Controls, IT Systems, MIS and Scientific Data Management Sections involved.
-

Choosing a Catalogue

- ICAT in a Production instance
 - Although not yet externally accessible.
 - 1 Beamline fully integrated (Non-Crystalline Diffraction)
 - 1 Beamline to be integrated (lean approach – file/folder parsing) before summer break.
 - 1 Beamline to be integrated before end of the year.
 - Smooth integration for the rest of the beamlines.
 - Requires coordinated effort with Controls and IT Systems (Priorities...)
-

Minimum Product

- File Ingest → Yes
 - Authenticated Data Access → Yes
 - Domain agnostic metadata
 - No standard metadata from all beamline, except for the Proposal (which comes from the User Office)
 - PaN API
 - Implementation of the OAI-PMH API, but not tested (won't until ICAT is externally accessible)
 - Stability
 - Lack of usage to be able to measure at this point.
-

EOSC Integration

- No federated access to data yet (no open data to expose at this moment)
 - Still working on our catalogue and Data Policy compliance.
 - OAI-PMH implemented
 - Yes, but will not be tested/used until Data Catalogue is accessible.
 - Harvesting by EOSC services
 - No. Dependency on ICAT and OAI-PMH.
 - EOSC Services access.
 - Not for the moment.
 - EOSC Statistics
 - Not for the moment.
-

PaN Specific implementation

- Domain metadata ingest
 - Integrating 1 beamline at a time, adapting metadata ingestion for each one of them.
 - Raw and derived data catalogued.
 - Only Raw is planned. Still need to measure what would suppose to store also derived data in terms of cost.
 - Data provenance.
 - Not at the moment. Will be considered.
 - Calibration data
 - Depending on the Beamline.
 - Link to External PID
 - No.
 - Link to Publications
 - Not yet.
 - eLogBook Data.
 - Not yet integrated with Data Catalogue (exploring ICAT's solution)
 - Currently using PSI's solution but not integrated.
-



Thank you

*18th of June 2021
Daniel Salvat*