

Periodic progress report

March 2021 to February 2023

Document Control Information

Settings	Value
Document Identifier:	D1.9
Project Title:	ExPaNDS
Work Package:	WP1
Document Author(s):	Sophie Servan (DESY), Alun Ashton (PSI), Anton Barty (DESY), Isabelle Boscaro-Clarke (Diamond), Thibaud Cayla (Diamond), Patrick Fuhrmann (DESY), Oliver Knodel (HZDR), Uwe Konrad (HZDR), Juliane Marauska (DESY), Brian Matthews (UKRI), Abigail McBirnie (UKRI), Paul Millar (DESY), Carlo Minotti (PSI), Majid Ounsy (SOLEIL), Kat Roarty (Diamond), Silvan Schoen (DESY), Nicolas Soler (ALBA), Ana Valcarcel-Orti (Soleil)
Responsible Partner:	DESY
Doc. Issue:	1.0
Dissemination level:	Public
Date:	03/05/2023

Abstract

This document presents the progress of the ExPaNDS project between March 2021 and February 2023. It reproduces the explanation of the work carried out by our partners as provided to the European Commission in the second and last periodic report of the project.

Licence

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.



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

Document Log

Issue	Date	Changes
1.0	03/05/2023	First version

Glossary

Term	Definition in the context of this document
DMP	Data Management Plan
DoA	Description of Action (also known as Description of Work, DoW)
EOSC	European Open Science Cloud ¹
ELN	Electronic Lab Notebook
LEAPS	League of European Accelerator-based Photon Sources ²
LENS	League of advanced European Neutron Sources ³
NeXus	NeXus is a common data format for neutron, x-ray and muon science ⁴
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting ⁵
PaN	Photon and Neutron, the community of photon and neutron sources
PaNET	Photon and Neutron Experimental Techniques ontology ⁶
PID	Persistent (and global) IDentifier
POPD	Protection of Personal Data
RDM	Research Data Management
RIs	Research Infrastructures

¹ <https://eosco-portal.eu/>

² <https://leaps-initiative.eu/>

³ <https://lens-initiative.org/>

⁴ <https://www.nexusformat.org/>

⁵ <https://www.openarchives.org/pmh/>

⁶ <http://purl.org/pan-science/PaNET>



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

Executive Summary

ExPaNDS is a Research and Innovation Action funded project under topic (b) of INFRAEOSC-05-2018-1019. Topic (b) particularly aims to support the gradual alignment of policies and practices of thematic initiatives to EOSC standards and to enable EOSC-relevant, non-commercial services to be accessed through the EOSC portal.

ExPaNDS emerged as a project in response to this call, with 10 national Photon and Neutron research institutions and EGI, collaborating on a cross-disciplinary cloud infrastructure, to enhance the accessibility and reusability of their valuable data products.

The collaboration aimed to streamline the workflows of our resource-intensive infrastructures by ensuring that data is Findable, Accessible, Interoperable, and Reusable (FAIR). In alignment with the objectives of the call, the project contributes to the EOSC by offering PaN-related data, services, and training resources to bolster the FAIRness of PaN data. This integration positions the EOSC as a comprehensive, centralised hub for obtaining these valuable resources.

The Photon and Neutron community encompasses over 40,000 scientists who utilise the various European synchrotrons, free electron lasers and neutron sources for diverse scientific areas. These disciplines span a broad array, from protein research, virology, and biomedicine to technical fields such as battery development, transportation, and supercomputer chip research, as well as studies in cultural heritage.

As historically there is little interaction between the above mentioned sciences and as the primary focus of our large facilities is the efficient use of their devices, long-term policies to handle the valuable data collected during expensive beam times were not consistently defined and often relied on undocumented agreements between the primary investigator and the facility. Generally, there was little emphasis placed on long-term archiving, proper annotation of data with metadata or reuse of the data.

The situation at the facilities evolved significantly in the different work areas addressed within the ExPaNDS project, as we implemented the following actions:

- Providing templates and best practices for data policies and FAIR data management, aligning with the EOSC and empowering facilities to continuously self-assess their FAIRness in implementing these regulations;
- Supplying reference datasets and their corresponding cloud-ready analysis pipelines for a wide range of techniques, serving as exemplary models for future activities;
- Enhancing the findability of scientific Photon and Neutron data in the EOSC by contributing to widely-used metadata systems, enabling their content to be harvested through standard protocols;
- Developing ontologies for our experimental techniques and provide appropriate keywords for the Nexus file format;
- Making our training platform available in the EOSC marketplace;
- Engaging high level management from our facilities and national stakeholders in the project's outcomes by showcasing interviews with those individuals in published materials.



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

It is worth noting that during the COVID-19 pandemic the PaN community has underscored their importance to society by providing fast track access to their facilities responding to the vital need for determining the structure of the virus and contributing to the development of treatments.

While the project, due to the call's requirements, was limited to nationally funded laboratories, the user base of these facilities are international scientists. To prevent fragmentation of policies and mechanisms, ExPaNDS closely cooperated with the PaNOSC project, which encompasses ESFRI facilities and ERICS.

Lastly, it is important to highlight that in addition to aligning Photon and Neutron Science workflows with EOSC standards and making the results accessible through the marketplace, a primary objective of the project was to create a substantial positive influence on the facilities' daily operations. Ensuring the project's outcomes are sustainable by connecting with similar national initiatives and actively contributing to the future evolution of EOSC towards our highly significant scientific field is also a key focus. A large portion of this document discusses these endeavours in detail.



Table of Contents

Executive Summary	3
1. Objectives	7
2. Work Package 1: Management and Sustainability	11
2.1. Consortium coordination and project administration	11
2.1.1. Internal communication	11
2.1.2. Recommendations	12
2.1.3. Amendments	12
2.1.4. IPR issues	12
2.2. Financial management	12
2.3. Project quality assurance	13
2.3.1. Software quality	13
2.3.2. Risks	13
2.3.3. Key Performance Indicators	13
2.3.4. Deliverables and milestones	14
2.3.5. Gender equality	14
2.4. Internal communication	14
2.5. Coordination with PaNOSC and EOSC	15
2.5.1. Coordination with PaNOSC	15
2.5.2. Coordination with EOSC	17
2.6. Sustainability plans	19
2.7. Coordination of boards	20
2.7.1. Collaboration board	20
2.7.2. Technical coordination board	20
2.7.3. EOSC liaison group	21
2.8. DMPs for the project and ethics of handling data	21
2.9. Deliverables and milestones of the WP for the period	21
3. Work package 2: Enabling FAIR data for PaN national RIs	22
3.1. Alignment of policies	22
3.2. Data management planning and DMP	23
3.3. Mainstreaming of standards for data management	24
3.4. Persistent identifier infrastructure	26
3.5. Quality assurance and certification schemes for data repositories	27
3.6. Uptake of FAIR data practices	28
3.7. Deliverables and milestones of the WP for the period	29
4. Work package 3: EOSC data catalogue services for PaN national RIs	29
4.1. Develop an EU PaN ontology	30
4.2. Implement ontologies in metadata catalogues	31
4.3. Coordinate metadata catalogues and data life cycle	32
4.4. Integrate metadata catalogue services into EOSC	32
4.5. Training material	33
4.6. Deliverables and milestones of the WP for the period	33



5. Work package 4: EOSC data analysis services for PaN national RIs	34
5.1. Alignment of PaNOSC services	34
5.2. Testing and validation framework	35
5.3. Prototype data analysis services	35
5.4. Deploy data analysis services into EOSC	36
5.5. Deliverables and milestones of the WP for the period	37
6. Work package 5: Training activities through EOSC platforms	38
6.1. Organising at national RIs a series of staff and users training events	38
6.2. Evaluation of available e-platforms and provision of a training demonstrator	39
6.3. Collection, Preparation and Publishing of Training Material	40
6.4. Monitoring and Measuring of training effectiveness	41
6.4.1. Workshops organisation	41
6.4.2. PaN-training.eu	41
6.5. Deliverables and milestones of the WP for the period	42
7. Work package 6: Dissemination and outreach	42
7.1. ExPaNDS internal and external communications strategy and delivery	43
7.2. ExPaNDS' marketing and communication toolbox	47
7.3. Formulation and dissemination of metrics and refinement of KPIs	48
7.4. Deliverables and milestones of the WP for the period	48
Appendix: List of all ExPaNDS-related meetings	50



1. Objectives

The specific objectives for the project as described in section 1.1 of the DoA are listed in this section, with a summary description of the work carried out during the second period of ExPaNDS towards the achievement of each listed objective. Full details are given in the following chapters.

Objective 1	To deliver the EOSC to the wide variety of Photon and Neutron users
Description (DoA)	The EOSC offers an opportunity to improve the availability and effectiveness of data services for users to deliver their science. ExPaNDS aims to develop and deliver capability within the EOSC, which can be utilised across the user base.
ExPaNDS action	<p>The technical services provided by ExPaNDS in EOSC, as described in the architecture, were developed and advertised within the Photon and Neutron user community throughout the second period of the project.</p> <p>Use cases were developed for seven techniques, along with reference datasets and analysis workflows, allowing Photon and Neutron users to learn and train on the usage of our EOSC services.</p> <p>A training workshop dedicated to EOSC was provided by the project in April 2021, followed by several training workshops on our data services in 2022. ExPaNDS was presented at ten user meetings and scientific conferences⁷, raising the awareness of EOSC in the PaN community further.</p> <p>Finally, to communicate the importance of FAIR data handling and our involvement in the EOSC we recorded interviews with four key decision makers and senior management from our community.</p>
KPI	More than 16.500 scientists were reached by ExPaNDS dissemination. <i>Target: 14.000 scientists</i>

Objective 2	To enable FAIR scientific data at European national RIs
Description (DoA)	Data sharing within the EOSC by national RIs requires the enabling and adoption of FAIR data principles and practices within RIs, for managing and publishing data within the EOSC. A FAIR data framework of policies, practices and guidelines tailored to the Photon and Neutron domain will be developed in consultation with national RIs and their user communities which they in turn will adopt.
ExPaNDS action	ExPaNDS has developed a set of recommendations and guidance to further enable FAIR data practices and processes within the national PaN RIs, including frameworks for DMP and PID usage and recommended metadata. This has been extensively discussed and promoted within the ExPaNDS RIs at all levels to raise awareness and knowledge of the practical steps which can be undertaken for FAIR data. ExPaNDS then developed a tool to enable facilities to assess their state of FAIRness.

⁷ See the appendix for a complete list of ExPaNDS-related events in the second period.



KPI	All 10 ExPaNDS partner facilities plan to implement at least one FAIR-related action as a direct result of the project. 6 facilities have already completed concrete actions to adopt FAIR policies. <i>Target: 10/10 facilities with a plan to implement at least one FAIR-related action as a direct effect of ExPaNDS</i>
------------	---

Objective 3	To make national RI data accessible to user communities via the EOSC
Description (DoA)	National RI FAIR data will be made accessible and shareable across wide user communities to ensure RI's data catalogues conform to common standards and APIs and made available within the core EOSC Portal and data search services. This will make data more accessible to existing and potential new user communities, offering new opportunities for more effective data exploitation and reuse, and provide innovative data and catalogue services for the Photon and Neutron community and beyond.
ExPaNDS action	ExPaNDS has closely monitored and supported members in improving the metadata lifecycle management, from data cataloguing at experiment time to data sharing at publishing time. By leveraging existing search engines and collaborating with PaNOSC in the creation of a PaN-specific search portal, ExPaNDS has improved the findability of PaN national RIs data and metadata in EOSC. Before ExPaNDS, there was no commonly agreed terminology. This limited the ability for researchers to find data. This issue was addressed by developing a set of common controlled vocabularies for experimental techniques (PaNET), a NeXus ontology and a semantic mapping ontology (PaNmapping) which are now being adopted in many RIs. These vocabularies have been extensively discussed and promoted within the ExPaNDS RIs and included in the PaN training catalogue.
KPI	The level of maturity of our facilities data catalogues to enable FAIR services was evaluated in February 2023 as 70.6% ⁸ . <i>Target: 65%.</i>

Objective 4	To enrich the set of data services available in the EOSC with services from national RIs
Description (DoA)	National RI FAIR data processing and analysis services targeted at specific user communities will be made available and shared across wide user communities by ensuring RI's data services conform to common standards and APIs and will be made available within the EOSC Portal services. This will make services more accessible to existing and potential new user communities, allowing access to the tailored, scalable services to a wider community, offering new opportunities for more effective data analysis.
ExPaNDS action	After having selected them in the first period, based on their added-value and relevance to both PaN and EOSC, ExPaNDS made several data analysis

⁸ Raw data is available at <https://doi.org/10.5281/zenodo.7851972>



	<p>software infrastructure-agnostic so that users can run complete pipelines developed at different PaN facilities from their own lab.</p> <p>These data analysis software run either on Jupyter notebooks or virtual machines, both of which have been registered as EOSC services by several of our partners. This is all documented in our PaN training platform for users to discover and train on.</p> <p>All our partners involved with data analysis services in ExPaNDS are now providers in EOSC.</p>
KPI	<p>5 analysis pipelines are available in EOSC. <i>Target: 5 analysis pipelines.</i></p> <p>10 partners are registered as service providers in EOSC. <i>Target: 10/11 partners.</i></p>

Objective 5	To raise the level of awareness and competence in FAIR data practices within user communities
Description (DoA)	<p>FAIR data publication, data sharing and the potential value of sharing services and resources within the EOSC are not familiar to the wider user community served by national RIs. ExPaNDS seeks to raise the awareness of the potential benefits of FAIR data sharing within the EOSC by engaging with different stakeholder communities, particularly user groups. This will offer advocacy, guidance and training to raise the level of awareness and competency of FAIR data practices and use of shared services within the EOSC.</p>
ExPaNDS action	<p>Eight training workshops were organised during ExPaNDS' second period, to discuss and build community competence on the main outcomes of the project and on the use of workflows in the training platform. Half of them were hybrid and collocated with larger meetings to reach a good audience, the others being online.</p> <p>FAIR practices were advocated throughout the second period to different stakeholder groups, facilities senior management, instrument scientists, data and computational support, librarians, user scientists and the RDM community at large, through:</p> <ul style="list-style-type: none"> - our participation in workshops and conferences; - the structured discussions around the FAIR data policy and metadata frameworks; - the FAIR self-assessment.
KPI	<p>The total number of attendees to the ExPaNDS training workshops is 754. <i>Target: 500 participants in total to all ExPaNDS training workshops</i></p>

Objective 6	To engage with the EOSC programme to empower national RI user communities and to maintain sustainability
Description (DoA)	<p>In order to fully realise the potential gains from FAIR data sharing within the EOSC, the national RI community needs to engage with the broader EOSC programme to understand the potential synergies with activities across the EOSC, deliver their benefits back to the national RI communities, and to</p>



	<p>feedback experiences and requirements to influence the further development of the EOSC. ExPaNDS will participate in the emerging governance structures of the EOSC (e.g. Executive, Stakeholder forums, technical Working groups) to ensure it is represented within the evolution of the EOSC. Further, ExPaNDS will participate in the work undertaken within a wide variety of EOSC activities.</p>
ExPaNDS action	<p>ExPaNDS participated in the EOSC symposiums and the EGI conferences in 2021 and in 2022. We also took part in the survey to collect the key exploitable results of all Horizon 2020 EOSC-related projects in spring 2022. All eligible partners (i.e. not in Switzerland or the UK) are now in the EOSC-Association or in the process of applying. Our involvement in EOSC and in particular within the ESFRI science clusters, along with PaNOSC, allowed us to be part of the EOSC-Future consortium as well as in the OSCARS proposal to keep representing the Photon and Neutron community in EOSC in the future. We also have an “integrated use case” as the photon and neutron science in the FAIR-IMPACT project.</p>
KPI	<p>There are 4 members (ALBA, EGI, Elettra, MAX IV) and 2 observers (DESY, SOLEIL) in the EOSC Association. <i>Target: at least 5 partners in the EOSC Association.</i></p>



2. Work Package 1: Management and Sustainability

WP1 implemented the project's governance and management structure, undertaking overall project administration to ensure successful achievement of the project's objectives. It also addressed the general technical architecture and the sustainability of the provided services for EOSC.

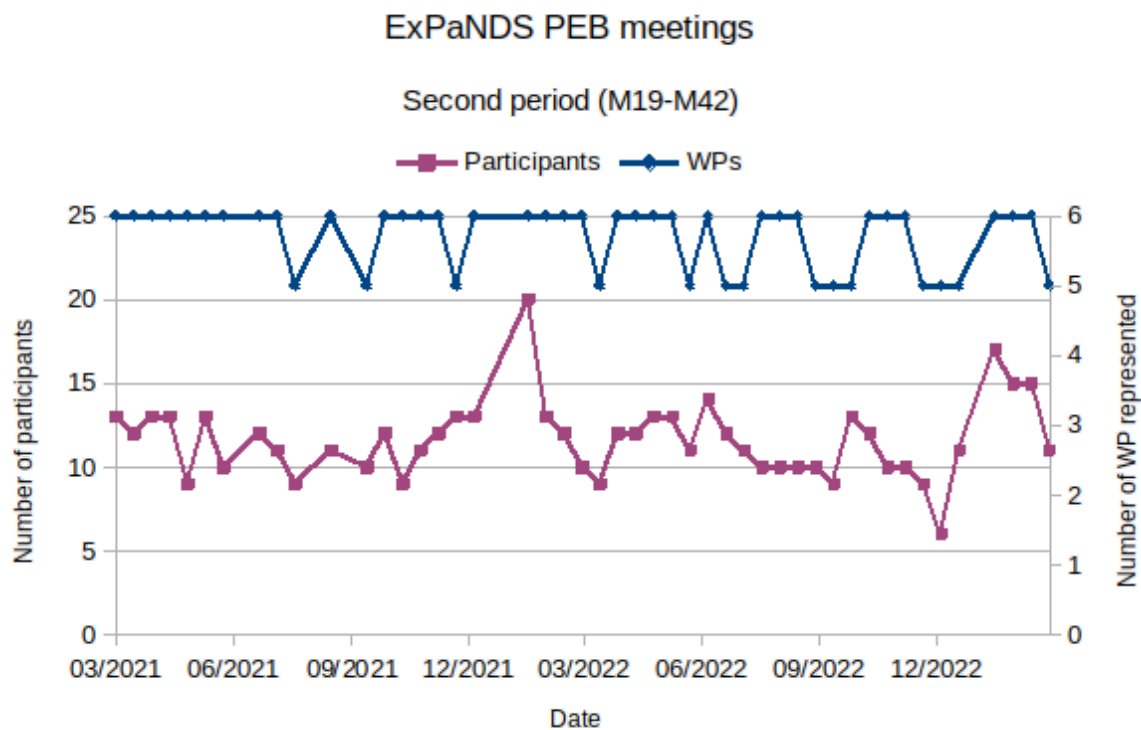
2.1. Consortium coordination and project administration

2.1.1. Internal communication

In the second period, WP1 (DESY) kept the meeting structure running, to ensure continued efficient communication and fulfilment of our objectives.

- Project executive board meetings

The project executive boards were held at a biweekly rhythm throughout the whole project. The participation in the PEB, as shown in the figure below, demonstrates the continued engagement of all WP leaders and co-leaders until the end of ExPaNDS. The peak in early 2021 is due to the second "extended PEB", which was organised by WP1 to prepare for the last year of the project's execution with all partners.



Statistics on PEB meetings during the second period: dates, number of participants and number of work packages represented.



- “Catch-up” meetings

To keep track of the WPs’ progress and alignment to their goals, catch-up meetings between WP1 and the other work packages have regularly taken place. Resulting “progress reports” were then made publicly available in GitHub⁹.

2.1.2. Recommendations

WP1 (DESY) kept track of the recommendations issued following the project’s mid-term review and regularly discussed their implementation with the concerned WPs, using the project’s issue tracking.

2.1.3. Amendments

After the mid-term review, a second amendment of the grant was prepared along with WP6 to rewrite the user engagement strategy, in line with the new paradigm brought by the COVID-19 pandemic. The updated strategy and associated amendment were approved by the European Commission in October 2021.

2.1.4. IPR issues

The project followed the principles of Open Science in relation to Intellectual Property Rights. All the project documents are published under a public licence (CC BY 4.0). Software developments are all open source, under different licences, and are publicly available in git repositories. The project’s data management plan was updated by WP1 (DESY) at the end of the project to reflect this.

WP1 also monitored GDPR-related issues in accordance with the POPD requirements issued during the first period. Notably, the PaN-training.eu platform’s conformance with GDPR was checked by HZDR’s data protection officer.

2.2. Financial management

Following the deviations in the use of resources during the first period and as reported in the first periodic report, WP1 (DESY) did a second interim report at M30 to check the targets set for each partner at mid-term were reached.

77% of the budget at M30 had been spent. All partners assured that the whole budget would be used by the end of the project. We therefore decided that no additional interim report at M36 was needed.

The second (D1.13) and third (D1.14) reports on cumulative expenditure incurred by the partners were delivered in December 2021 and December 2022, showing good progress in the spending rates of most partners.

⁹ <https://expands-eu.github.io/WP1-reporting/reports.html>



2.3. Project quality assurance

The project quality assurance processes defined in the first period continued to be observed by WP1 (PSI) in the second period.

2.3.1. Software quality

The testing and validation framework written in the frame of WP4 is applying EOSC Synergy's software quality rules (see WP4 report). WP1 (DESY) made several connections with the EOSC-Synergy project:

- arranging a presentation of Synergy's SQaaS in the ExPaNDs workshop on EOSC organised by WP1 in April 2021;
- arranging and chairing our participation in EOSC Synergy's session on EOSC at the EGI conference 2021, where we were invited to present our testing framework as a use case of Synergy's quality assurance developments.

2.3.2. Risks

Following a recommendation from mid-term review, the risk analysis was updated by WP1 (PSI) in January 2022, along with PaNOSC's project manager and the WP leaders.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Risk / Threat	Date of registration	Initial mitigation measure	Risk owner	Likelihood	Impact	Risk level	Consequences of occurrence	WP affected	Identified new vulnerability	Updated mitigation measure	Follow-up actions	Last update	STATUS	PaNOSC ref.	
16	14	Icat and SciCat cannot be extended	Nov 2018	The WP3 partners have developers in both teams, which will ensure that integration will be seamless and smooth.	WP3 leader	Unlikely	Moderate	Medium		WP3	Solved with publication of D3.3		Feb 2022	SOLVED		
22	20	Difficulties in deployment of metadata catalogue APIs	Nov 2018	Use input from WP3 and PaNOSC to make use of metadata catalogue test infrastructure.	WP3 leader	Possible	Moderate	Medium	Delay	WP3, WP4	Deployment of API identified in roadmap (D3.1)	Regular workshops in the frame of tasks 3.4 and 3.5	Feb 2022	ACTIVE	6, 7	
27		PaN search API not sustained after PaNOSC ends	Dec 2021	Define future governance and release management with PaNOSC	WP3 leader	Possible	Major	High		WP3		Will be followed-up as part of the sustainability task	Feb 2022	CLOSED		
28		VISA for PaN not sustained after PaNOSC ends	Dec 2021	Define future governance with ILL, identify what is extra development compared to ILL version	WP4 leader	Possible	Major	High	Extra load on partners who changed their production service to VISA, no community solution and lack of interoperability, duplication of effort	WP4	No alternative for VISA in the community	Will be followed-up as part of the sustainability task	Feb 2022	CLOSED		
29	U1	Duplication of efforts	Nov 2020		project manager	Possible	Moderate	Medium	Inefficient use of research funding	ALL WPs	Participation of all ExPaNDs WPs in the 5b project task forces, close collaboration with PaNOSC with regular meetings	Mapping document of collaboration and complementary areas being elaborated with PaNOSC Follow-up issue for training platform workflow feature	Jan 2022	ACTIVE	8	
31	U3	EOSC onboarding requirements incompatible with our services	Dec 2021	EGI setting up a working group on onboarding for ExPaNDs	WP1 leader	Likely	Minor	Medium	Need for new developments, delays, impossibility to onboard planned	WP3, WP4, WP5	Data policy and TRL level not sufficient Unique helpdesk vs. federated service needs a "Meet" between	Share experience between partners, with PaNOSC, and discuss with EOSC-Future / EGI	Jan 2022	ACTIVE	Materi alised	

Extract of the project's risk assessment sheet

2.3.3. Key Performance Indicators

The project's KPIs which were updated following the mid-term review were followed by each WP during the second period in a dedicated internal dashboard. WP1 (DESY) also kept the public overview of the main objectives and target KPIs in GitHub¹⁰.

¹⁰ <https://github.com/ExPaNDs-eu/ExPaNDs/wiki/ExPaNDs-objectives-dashboard-per-partner>



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

2.3.4. Deliverables and milestones

36 deliverables and 15 milestones have been submitted in the second reporting period. WP1 (PSI and DESY) suggested internal reviewers and checked the deliverables quality and consistency with other projects and WPs.

Difficulties with resources and the ongoing impacts of the COVID-19 pandemic until mid-2022 have led to several delays in the delivery of our milestones and deliverables. WP1 evaluated the impacts of these deviations and supported the WPs whenever possible.

2.3.5. Gender equality

Even though most of our facilities have equal opportunity officers, the PaN community has not yet achieved full gender balance and is still typically dominated by men. Being aware of this, we have strived, along with WP6, to provide female role models throughout our conferences, interviews and publications. We also set an example with a gender-balanced executive board.

It is the responsibility of each facility to monitor their gender equality, as described in their own internal policy. The number of female researchers on the grant is also monitored by each partner and reported on in the EC portal.

2.4. Internal communication

The second all-hands meeting took place online on 26th of October 2021, jointly organised with PaNOSC. A news piece¹¹ was released by WP6 on our website, providing a summary of the event, links to presentations and recordings.

The third all-hands meeting was the ExPaNDS closing event on 23rd and 24th of January 2023 at DESY. In addition to an internal meeting to prepare for the final report and to discuss the sustainability of our main outcomes, a public event was organised to present a user story showcasing the major results achieved by the project. This public event was an official satellite event of the DESY and European XFEL user meeting which enabled us to reach ~50 users who attended the event on site and ~30 users online.

All meetings ExPaNDS organised or was represented in are listed in appendix.

For internal communication, the existing ExPaNDS wiki was extended with new pages, providing for example a catalogue of ExPaNDS presentations, a one-stop-shop to all event-related resources, complementing expands.eu where the higher level of detail is more adapted to the audience targeted.

The project's issues were also monitored in ExPaNDS main repository issues tracker, after a progressive shift from SharePoint to make them more transparent to all project interesseees. In total, 42 issues were open and closed during the second period, with 150+ comments from ~25 participants.

¹¹ <https://expands.eu/2021/11/23/overview-of-the-2nd-panosc-and-expands-pan-eosc-symposium/>





ExPaNDS team at the closing event in Hamburg in January 2023

2.5. Coordination with PaNOSC and EOSC

2.5.1. Coordination with PaNOSC

In the second period of ExPaNDS, the interfaces and complementarities with PaNOSC were better defined and understood by the partners, making the need to draw a clear line between the two projects less relevant.

Externally, the participation in events and the communication of outcomes was completely aligned, showing a common face to the community, in particular to LEAPS and LENS. This materialised for example in the integration of ExPaNDS in the PaNOSC science cluster, both in the EOSC-Future project and in the OSCARS proposal, written in response to the HORIZON-INFRA-2023-EOSC-01-01 call. Internally, interfaces were kept up-to-date by WP1 (DESY) in the dedicated wiki page¹².

¹² <https://github.com/ExPaNDS-eu/ExPaNDS/wiki/Mapping-ExPaNDS-and-PaNOSC-activities>



To facilitate the evaluation of ExPaNDS activities in relation to PaNOSC, an update of the overview of the common activities is provided in the table below:

WP	Domain	Common activities
WP1	Risk analysis	<ul style="list-style-type: none"> - Align methodology for assessment and monitoring of risks - Review common risks together
WP1	Project DMP	<ul style="list-style-type: none"> - Share initial Data Management Plans and evolutions - Consult on handling of personal data
WP1	Sustainability	Converge on the sustainability policies applied to the technical WPs outputs
WP2	FAIR data policy*	Consult with the policy makers to consider the recommendations from the FAIR data policy framework
WP3	Common search API	Develop the search API to be used by all PaN facilities (meta)data catalogues
WP3	Metadata catalogues*	<ul style="list-style-type: none"> - Evolve to community data catalogues (e.g. ICAT and SciCat) - Make them externally reachable: OAI-PMH and PaN search API endpoints - Publish them as EOSC services
WP5	Training	Organise training workshops
WP6	Dissemination	Communicate, share events and news

Common activities with PaNOSC

For these activities, we aim to join forces as much as possible, or in the case of the activities marked with “*”, which intrinsically need to be replicated at each facility, we support each other by sharing experience and tools.

Another important aspect of the coordination with PaNOSC is identifying and managing the interfaces between both projects. The following table lists those interfaces, with the updated complementary focus of each project’s activity in that domain:

WP	Domain	PaNOSC focus	ExPaNDS focus
WP2	Instrument DMPs	Develop a PaN-ready Data Management Plan template	<ul style="list-style-type: none"> - Explore sources to collect the answers to the questions in the DMP template - Present a data model and an architecture for a possible active DMP implementation
WP2	Data policy framework	Update the PaNdata policy to make it FAIR aware	<ul style="list-style-type: none"> - Abstract the PaNOSC data policy framework to higher level principles to make it compatible with national policies - Consult the national RIs on the data policy framework



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

WP	Domain	PaNOSC focus	ExPaNDS focus
WP3	PaN Ontologies	Propose new definitions to NeXus	- Define an ontology for PaN techniques - Develop an ontology for NeXus terms, making them FAIRer
WP3	PaN data portal	Develop a single entry point portal to all open data in PaN (data.panosc.eu)	Provide endpoints to the data portal at national facilities
WP4	Data analysis workflows	Propose a PaN solution for remote data analysis: VISA	- Make a selection of data analysis software portable between facilities - Expose remote analysis services to the EOSC - Enhance VISA's portability and deploy it at national facilities
WP5	PaN-learning platform	- Update and add features to the neutron e-learning platform - Integrate it with PaN-training.eu	Develop a catalogue of training content: PaN-training.eu

Interfaces with PaNOSC

During the second period, a number of events were co-organised by both our projects:

- two online PaN EOSC symposiums (10/2021 and 10/2022);
- an alignment workshop in Copenhagen, focused on the expectations for the last year of PaNOSC (11/2021);
- an additional all-hands meeting at ELI beamlines near Prague, focused on cross work package tasks and strengthening the network between the two projects (06/2022);
- a common session at the EGI conference 2022 on the PaN community beyond ExPaNDS and PaNOSC (09/2022).

2.5.2. Coordination with EOSC

The general architecture of ExPaNDS with regards to EOSC was updated under the umbrella of WP1 (EGI)¹³. In addition, to help WP3 and WP4 achieve their respective objectives to onboard services in EOSC, EGI created a task force specifically dedicated to onboarding ExPaNDS services in the marketplace. PaNOSC partners also joined. One to one support was provided to partners to onboard their services (e.g SOLEIL, MAX IV).

A status overview of the onboarding of the ExPaNDS-related services to EOSC was constantly updated by the WPs and WP1 in the public objectives dashboard¹⁴.

In the frame of WP5, a workshop on EOSC was organised by WP1 (EGI and DESY) in April 2021¹⁵. It addressed how ExPaNDS delivers its data services to PaNOSC and EOSC, with a

¹³ <https://doi.org/10.5281/zenodo.6958045>

¹⁴ <https://github.com/ExPaNDS-eu/ExPaNDS/wiki/ExPaNDS-objectives-dashboard-per-partner>

¹⁵ <https://indico.desy.de/event/29297/>



first day on data catalogues: getting our open datasets harvested by the community and a second day on data analysis: getting our software in the PaN software catalogue, getting pipelines to run, and ensuring their quality over time.

ExPaNDS participated in the EOSC symposium 2021 and 2022 (see the complete list of talks in appendix):

Persistent identification as the factor of FAIR

- The access model is not defined explicitly;
- In terms of resolving, in current model all kinds of the existing data representing the given Object are stored separately;
- The data represented on the diagram are existence-agnostic about each other;
- The Client obtains provenance from the trusted service that is actually self-signed.

Diagram illustrating the relationships between data objects:

```

    graph TD
      Collection -- aggregates --> DO((DO))
      DO -- is represented by --> Repository
      DO -- is described by --> Metadata
      DO -- is referenced by --> PID[Persistent ID]
  
```

ExPaNDS logo and funding information are visible at the bottom of the slide.

“Openly reproducible PIDs as a factor of FAIRness in data sharing practices” presentation at the EOSC Symposium 2021 (speaker: Andrey Vukolov, ELETTRA Sincrotrone Trieste)¹⁶

PaN Training Catalogue and E-Learning Platform combined in one Overall Portal

ExPaNDS PaN-training.eu and PaNOSC E-learning.PaN-training.eu

Central place to access all ExPaNDS / PaNOSC training materials and our e-learning platform, define training workflows to guide the scientist through the materials, harvest events and other e-learning platform courses from external providers or PaN RIs:

- Website,
- GitHub,
- Video (youtube),
- Training datasets, (from Zenodo)
- Workshop minutes,
- ...

The central e-learning platform for blended learning courses with PaN content and additional features:

- Jupyter notebook integration,
- Simulations (neutrons, VINYL),
- Virtual lab and
- Wiki.

Logos for LEAPS, LENS, lightsources.org, moodle, jupyter, and MediaWiki are shown.

EOSC SYMPOSIUM 2022 logo is in the top right corner.

“The Training Catalogue for Photon and Neutron Data Services” presentation at the EOSC Symposium 2022 (speaker: Oliver Knodel, HZDR)¹⁷

¹⁶ [Breakout 1: Rules of participation & PID policy, Day 4 #EOSCSymposium2021](#)

¹⁷ [Day 2: Training & Skills for EOSC: Lightning talks](#)



Thanks to ExPaNDS, by the end of the project:

- 6/11 partners are now in the EOSC-Association;
- 10/11 partners are EOSC providers.

2.6. Sustainability plans

The main focus of WP1 (DESY and PSI) for the last year of the project was to build a sustainability plan with our WPs for each of their outcomes. This resulted in the publication of 9 “sustainability sheets”, following the same template to present in a condensed manner:

- Description of the solution, its target audience, its benefits;
- Accessibility, including links to relevant documentation and feedback mechanism;
- Competitors, readiness level and EOSC integration status;
- Plans and conditions for sustainability and exploitability.

Outcome	TRL	Sustainability sheet DOI
FAIR implementation framework	Prototype	10.5281/zenodo.7804712
NeXus ontology	In production	10.5281/zenodo.7702023
PaNET ontology	In production	10.5281/zenodo.7684470
Ontology API service	Prototype	10.5281/zenodo.7744338
Search API for data catalogues*	Prototype	10.5281/zenodo.7743913
OAI-PMH endpoint for data catalogues	In production	10.5281/zenodo.7789134
VISA*	Pilot	10.5281/zenodo.7788839
Testing and validation framework	Pilot	10.5281/zenodo.7867091
PaN training platform	In production	10.5281/zenodo.7732603

*Sustainability sheets for ExPaNDS outcomes - PaNOSC outcomes with significant ExPaNDS contribution are also included and marked with **

Following the presentation of ExPaNDS and PaNOSC’s outcomes to the LEAPS plenary in October 2021, LEAPS and LENS chairs were invited to the 2nd PaN EOSC symposium to state where each of their members were in the road to adopt both our project’s outcomes. Their response to the survey showed a very good uptake rate.



The survey was updated a few times until the final picture we got at the PaNOSC final event:

FACILITY	FAIR data policy	DMPs	DOIs	Nexus HDF5	Search API	Open Data Portal	AAI	Jupyter Lab	VISA	VINYL/OASYS/McStas	Pan-learning/training
ALBA	P	P	WIP	WIP	WIP	WIP	P	Y	WIP	N	U
DESY	WIP	WIP	WIP	Y	WIP	P	WIP	Y	U	Y	WIP
CERIC-ERIC	Y	WIP	Y	WIP	Y	Y	Y	Y	Y	Y	Y
DIAMOND											
ELETTRA	Y	WIP	Y	Y	Y	Y	Y	Y	Y	Y	Y
ESRF	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
ELI-ERIC	Y	Y	P	Y	Y	Y	WIP	Y	Y	Y	Y
ESS	Y	Y	Y	Y	Y	Y	Y	WIP	WIP	Y	Y
EuXFEL	Y	WIP	Y	WIP	Y	Y	WIP	Y	WIP	Y	Y
FELIX	Y	P	WIP	U	U	WIP	U	U	N	N	U
HZB	Y	P	WIP	Y	P	Y	P	U	U	U	U
HZDR	Y	WIP	Y	N	U	Y	Y	Y	P	WIP	Y
ILL	Y	WIP	Y	Y	WIP	Y	Y	Y	Y	Y	WIP
MAX-IV	WIP	U	Y	Y	Y	Y	Y	Y	U	U	U
PSI	Y	WIP	Y	WIP	Y	Y	WIP	WIP	N	N	N
PTB	Y	WIP	Y	WIP	N	Y	N	N	N	N	N
SOLARIS											
SOLEIL	Y	WIP	WIP	Y	WIP	WIP	Y	WIP	WIP	U	Y
SESAME	Y	U	P	Y	P	WIP	P	P	N	Y	N

Yes, already adopted (Y)

Planned to be adopted (P)

Not Planning to be adopted (N)

Under evaluation (U)

In progress of being adopted (WIP)

Adoption (in October 2022) of PaNOSC+ExPaNDS outcomes¹⁸

2.7. Coordination of boards

2.7.1. Collaboration board

Three collaboration board meetings were organised by WP1 (DESY) in the second period.

- 23/04/2021: Rehearsal for the midterm review
- 17/11/2021: General update and plans for last year
- 24/01/2023: Sustainability

2.7.2. Technical coordination board

As the technical direction of the project had already been established, the TCB met only once in the second period, as a hybrid event co-located within the ExPaNDS closing event. In this TCB meeting, the board members were consulted on ways in which the technical output from ExPaNDS might be further adopted. The results of that meeting were collated and used as additional input when developing the different sustainability sheets.

¹⁸ <https://indico.esrf.fr/event/66/contributions/264/>



During this reporting period, the ExPaNDS technical coordinator (DESY) undertook several new roles that enable us to promote ExPaNDS outcomes and build connections within the wider research community:

1. PaN facilities community representative on the EOSC-Future Technical Coordination Board

Within that capacity, we are able to provide input, steering EOSC-Future in the direction of ExPaNDS goals while advocating ExPaNDS deliverables and building cross-cluster connections within the larger research community.

2. Member of the Technical Advisory Board of the DAPHNE4NFDI project

This allows us to advocate the adoption of ExPaNDS outcomes within the DAPHNE project, furthering the sustainability of our outcomes.

3. Leader of Working Group 3 on data within LEAPS

Through this position, we are building frameworks to disseminate ExPaNDS outputs and sustain the collaborations established during ExPaNDS (and PaNOSC). Continuing these connections will allow LEAPS member facilities to continue collaborating on the wide range of topics on FAIR data. Such collaborations will include neutron facilities whenever possible.

2.7.3. EOSC liaison group

EGL participated in the INFRA-EOSC-05b task force on Service Onboarding, discussing with the other projects the onboarding activities and issues encountered by PaN and reporting back to our facilities. It was also used as input for the ExPaNDS internal task force on onboarding EGL set up in early 2022 (as already reported above).

Our link to the EOSC also materialised with our participation in EOSC-Future's science project for PaN. At the facility level, 6 of our partners joined the EOSC-Association during the second period of ExPaNDS.

2.8. DMPs for the project and ethics of handling data

The project's DMP was updated to represent the final status of ExPaNDS "data" at the end of the project¹⁹.

2.9. Deliverables and milestones of the WP for the period

#	Deliverable	Partner	Due date	Actual date
1.09	Periodic progress report	DESY	28/02/2023	03/05/2023
1.13	Financial report 2	DESY	31/12/2021	29/12/2021
1.14	Financial report 3	DESY	31/12/2022	30/12/2022

¹⁹ <https://doi.org/10.5281/zenodo.3672926>



#	Milestone	Partner	Due date	Actual date
7	Second all-hands meeting organised	DESY	30/11/2021	26/10/2021
8	Third all-hands meeting organised	DESY	31/12/2022	24/01/2023
9	Successful completion of final review	DESY	28/02/2023	-

3. Work package 2: Enabling FAIR data for PaN national RIs

WP2 extended and deepened the adoption and use of FAIR data principles within the Photon and Neutron community to allow publication and access of national RI data and services within the EOSC.

The comprehensive set of deliverables and recommendations compiled in WP2 form an actionable FAIR Implementation Framework for PaN RIs. This toolkit equips our facilities with the necessary resources to further their implementation of FAIR data processes and ensure that their operations align with the EOSC Rule of Participation, which mandates that EOSC resources adhere to the FAIR Principles.

3.1. Alignment of policies

After the publication of the draft data policy framework in the first period²⁰, formal consultations with the facility staff in charge of data policies were carried out between February and May 2021 with each ExPaNDS partner facility. They led to the publication of the final data policy framework²¹, which now proposes 21 elements for RIs to consider when formulating data policy.

The work on the FAIRification of data policies was featured by the FAIRsFAIR project as the first of its implementation stories²².

²⁰ <https://doi.org/10.5281/zenodo.4014811>

²¹ <https://doi.org/10.5281/zenodo.5205825>

²² <https://doi.org/10.5281/zenodo.5040078>



Implementation Story

Data policy framework for photon and neutron research infrastructures

Theme 1: Defining a FAIR-enabling Policy Environment

The ExPaNDS project is the European Open Science Cloud (EOSC) Photon and Neutron Data Service: a collaboration, funded by the Horizon 2020 programme, between ten national Photon and Neutron Research Infrastructures as well as EGI, an e-infrastructure that provides advanced computing services for research. These research infrastructures are exemplars of “facilities science”, providing specialised instruments and techniques for use by researchers from a wide range of fields. One strand of work in the project is on “Enabling FAIR data for photon and neutron national research infrastructures”, and includes development of a data policy framework to allow the facilities to adopt a coherent approach to FAIRness of the data that they generate.

The activity examined here concerns revisions to the ExPaNDS data policy framework presented in the project’s Final data policy framework for Photon and Neutron RIs. The earlier draft policy framework (ExPaNDS deliverable D2.1: Draft extended data policy framework for photon and neutron RIs) had been published in September 2020, and drew on the FAIRsFAIR Policy enhancement recommendations. Before that, there had been joint work with the PaNOSC project, which had produced its own data policy deliverable in May 2020. However it became apparent that ExPaNDS partners were keen to explore the various themes of the data policy framework in more depth, especially with a view to providing greater flexibility in approach, so a programme of consultations with each of the ten ExPaNDS partner facilities was undertaken.

04 April 2022

Zenodo

ExPaNDS



Announcement of the ExPaNDS implementation story on [fairsfair.eu](https://www.fairsfair.eu)

3.2. Data management planning and DMP

The work on DMPs, led by HZB, was to develop a framework for practical DMPs and was undertaken in close collaboration with PaNOSC with teams contributing and building upon each other’s deliverables. Members of the work package interacted with the RDA Active Data Management Plans activity.

The WP reviewed some selected DMP workflows and roles in combination with project management lifecycles from the literature, particularly considering OAIS and PMBOK as contributing insights into the workflows and lifecycles in the PaN RI context.

Joint work with PaNOSC led to the development of a DMP template for PaN RIs²³. This template consists of over 100 questions which should be addressed to capture a complete picture of the data and computational context of an experiment; this is clearly too complex for most users.

Consequently, in ExPaNDS T2.2, the experimental context was analysed so that this could be radically simplified by considering the experimental lifecycle and supporting systems so that information could be extracted from different sources to answer most of the questions, thus leaving a manageable set of questions for a particular experiment. This is described in D2.4²⁴.

D2.8²⁵ takes these concepts further, describing how the DMPs can be implemented and automated within a data model supporting active DMPs, so they can be generated as simply as possible and the information within them used to steer the experiment and publish FAIR

²³ <https://doi.org/10.5281/zenodo.5639428>

²⁴ <https://doi.org/10.5281/zenodo.5636096>

²⁵ <https://doi.org/10.5281/zenodo.7223438>



This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 857641.

data. This includes a consideration of current DMP tools, such as DMPOnline, Data Stewardship Wizard and Research Data Management Organiser, and their suitability for implementing these recommendations.

3.3. Mainstreaming of standards for data management

Following the initial analysis in the first period²⁶, which concluded with a set of 31 recommended minimum metadata fields which facilities should publish as part of FAIR data, recommendations on how this metadata might be implemented were then addressed. The metadata recommendations were realised in the PaN search-API, developed within WP3 conjunction with PaNOSC, and the formal mapping recorded in D2.7²⁷.

²⁶ <https://doi.org/10.5281/zenodo.4312825>

²⁷ <https://doi.org/10.5281/zenodo.6821676>





Metadata framework in the PaN data continuum²⁸

²⁸ <https://doi.org/10.5281/zenodo.7680072>



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

This work also developed mappings between the core metadata and the metadata formats required by general data discovery services used within the EOSC, particularly B2Find and OpenAIRE, as well as the Dublin Core and DataCite metadata schemas were described, so that facilities can map their catalogues onto these broad discovery services.

Further, the use of NeXus to provide detailed experimental parameter data was discussed. Additional extensions such linking to FAIR research software information, the use of PROV-O for recording provenance information, and PREMIS for preservation information were considered briefly, with recommendations for further development.

This task was carried out by ALBA, with contributions from all WP2 partners, in particular UKRI.

3.4. Persistent identifier infrastructure

This task led by UKRI with close involvement of HZB and Elettra, aimed at providing PaN facilities with recommendations on adopting, integrating and further exploiting PIDs within the facilities process and as part of FAIR data publishing.

The work of the task initially targeted the use of PIDs for identifying and publishing data, as this is a fundamental use case, which some of the participating RIs have been in the process of implementing. A project workshop was designed to share best practice on the deployment and use of data PIDs and took place in January 2021.

The task also considered the growing family of recommendations for the use of PIDs to identify different resources. It aimed to provide guidance on the state of the art to navigate this complex landscape, building on previous work undertaken in the FREYA project, and to recommend where the PaN community might usefully benefit.

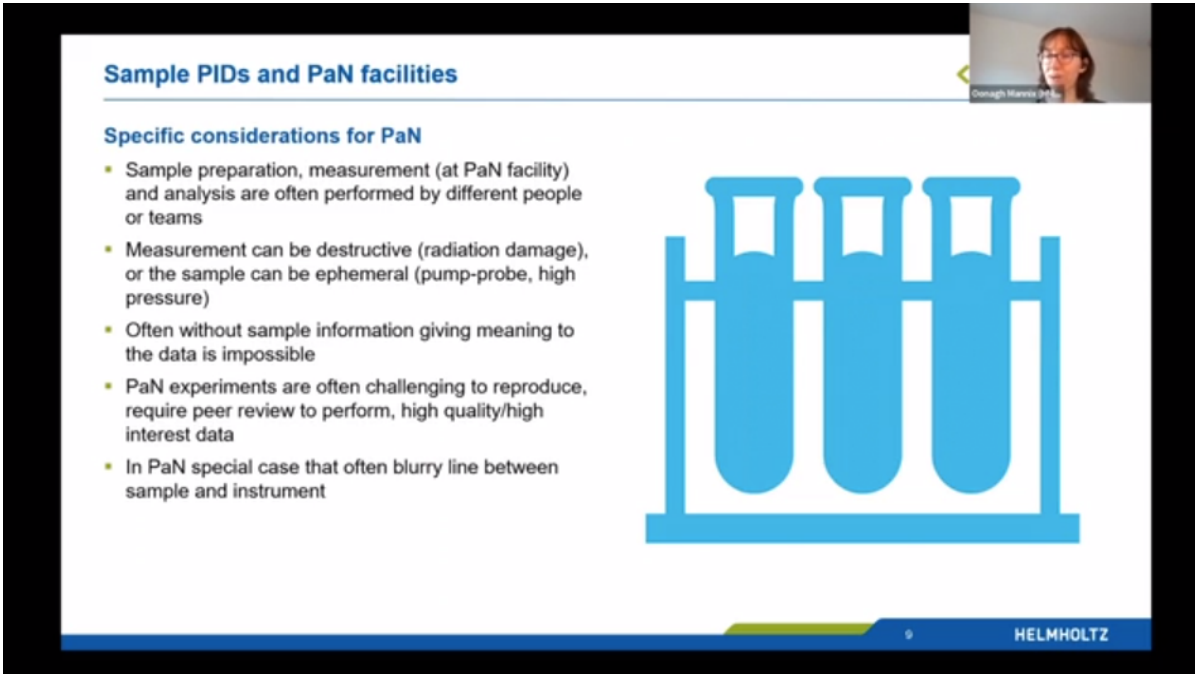
In particular, Organisation PIDs, Instrument PIDs and Person PIDs (e.g. ORCIDs) were seen as mature proposals which could be adopted immediately or with relatively small effort. Finally, how PID information from PaN RIs might be integrated into the PID-Graph as developed in FREYA and in Zenodo was explored, although its implementation is left to future work.

This work was presented in a second workshop in October 2021, which provided an update on the state-of-the-art and recent developments in the use of persistent identifiers that are popular elsewhere but are not yet widely adopted by research facilities.

The recommendations on PIDs are presented in D2.5²⁹.

²⁹ <https://doi.org/10.5281/zenodo.5905351>





Sample PIDs and PaN facilities

Specific considerations for PaN

- Sample preparation, measurement (at PaN facility) and analysis are often performed by different people or teams
- Measurement can be destructive (radiation damage), or the sample can be ephemeral (pump-probe, high pressure)
- Often without sample information giving meaning to the data is impossible
- PaN experiments are often challenging to reproduce, require peer review to perform, high quality/high interest data
- In PaN special case that often blurry line between sample and instrument

HELMHOLTZ

Workshop on PIDs for facilities held on the 22nd of October 2021 (speaker: Oonagh Mannix, HZB and HMC)³⁰

3.5. Quality assurance and certification schemes for data repositories

The area of FAIR assessment has been rapidly changing during the course of the project, so UKRI first undertook a survey of current best practice. It considered in detail the RDA FAIR Data Maturity Model, and from the FAIRsFAIR project, the ACME FAIR and the F-UJI tool. Further, FAIRsFAIR undertook an alignment with CoreTrustSeal for certifying data repositories, resulting in a CoreTrustSeal+FAIRenabling framework that provided additional valuable input.

However, these tools and methods produced by other projects, such as FAIR EVA from EOSC-Synergy and FAIR Enough, focus on assessing the FAIRness of individual data sets, and data repositories holding such data sets. For ExPaNDS partners it is more valuable to focus on whether the experimental process and its computational environment results in the generation of FAIR data. Our assessment method should be tailored to that aim, so these approaches were not considered directly suitable for application in ExPaNDS.

Consequently, UKRI developed a method to assess the FAIRness of facilities experimental processes, based on adapting the best practices from the published methods. A questionnaire was developed as a self-assessment tool, with a mixture of multiple choice and qualitative, textual questions to cover all the FAIR principles.

This was then used within a self-assessment exercise with all RI partners undertaking a detailed response. This allowed each facility to have a realistic assessment of their progress towards FAIR, without grading facilities as being more or less advanced.

³⁰ <https://vimeo.com/676015892>



Feedback from the facilities then formed recommendations towards modifying the method for future reruns. The method and results of the self-assessment are detailed in D2.6³¹.

3.6. Uptake of FAIR data practices

WP2 partners led by UKRI organised and participated in a large number of presentations, workshops, and other activities promoting the uptake at FAIR data practices, aimed at different groups of stakeholders, with a focus on facilities senior management, instrument scientists and other facilities staff, and user communities.

This included preparing presentation and training material in conjunction with WP5 and WP6.

UKRI worked closely with FAIRsFAIR, notably via the Synchronisation Task Force which FAIRsFAIR organised in conjunction with other EOSC-5b projects. WP2 (UKRI) shared the work of the project via this task force, for example contributing to workshops and providing adoption stories.

D2.9³² discusses the advocacy activities undertaken aimed at different stakeholder groups, together with an assessment of impact.



PaNOSC & ExPaNDS panel session at ESOF2022 - Open Data for healthier societies: a virtuous cycle? (Speaker: Isabelle Boscaro-Clarke, Diamond)³³

³¹ <https://doi.org/10.5281/zenodo.7246802>

³² <https://doi.org/10.5281/zenodo.7572045>

³³ <https://youtu.be/O2dN-3hfYus>



3.7. Deliverables and milestones of the WP for the period

#	Deliverable	Partner	Due date	Actual date
2.3	Final data policy framework for Photon and Neutron RIs	UKRI	31/08/2021	20/08/2021
2.4	DMPs for Photon and Neutron RIs	UKRI	30/11/2021	30/11/2021
2.5	Advanced infrastructure for PIDs in Photon and Neutron RIs	UKRI	28/02/2022	03/03/2022
2.6	Self-evaluation Photon and Neutron RIs for FAIR data certification	UKRI	30/11/2022	19/12/2022
2.7	Final Recommendations for FAIR Photon and Neutron Data Management	UKRI	30/05/2022	11/07/2022
2.8	Active DMPs for Photon and Neutron RIs	UKRI	30/11/2022	21/12/2022
2.9	Report on promotion of FAIR data within Photon and Neutron RIs	UKRI	31/01/2023	03/02/2023

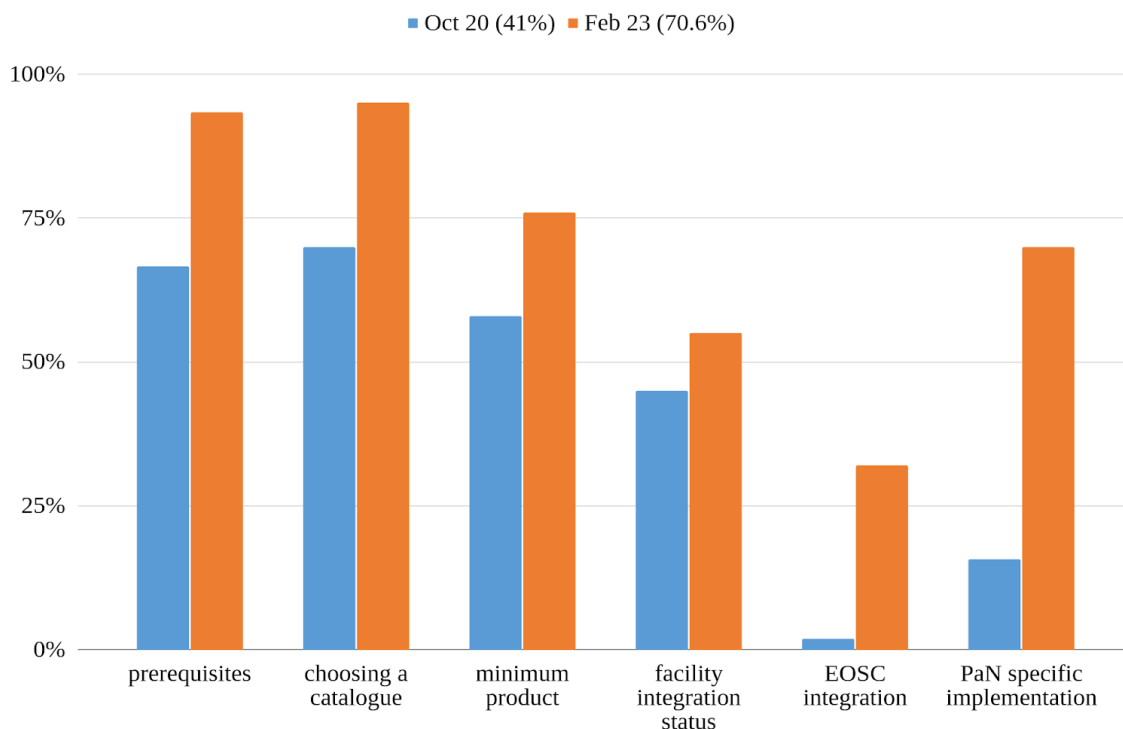
#	Milestone	Partner	Due date	Actual date
11	Production of final FAIR data framework	UKRI	30/11/2022	30/12/2022

4. Work package 3: EOSC data catalogue services for PaN national RIs

WP3 improved the importance and recognition of the value and need of cataloguing data with standardised metadata in facilities. It drove and supported wider adoption and usage of community data catalogues, SciCat and Icat, and of PaN and EOSC data commons, making PaN data more findable and interoperable.

The roadmap defined in the first period enabled monitoring the status of our partner facilities towards federated metadata catalogues. The following figures provide an overview of our progress throughout the whole metadata lifecycle at our facilities.





Facilities status towards federated metadata catalogues evolution between beginning (Oct 20) and end (Feb 23) of ExPaNDS³⁴

4.1. Develop an EU PaN ontology

Diamond has led the design and development of a set of ontologies to support FAIR data implementation in PaN data catalogues. Three ontologies were developed to facilitate consistent semantics for terms within the PaN domain, which involved providing global persistent identifiers, community-agreed labels and synonyms, and human-readable definitions, annotations and references:

- PaNET: Photon and Neutron Experimental Techniques Ontology. This simple ontology provides a taxonomy of PaN techniques, with new techniques being defined as subclasses of multiple, more elementary, technique classes.
- NeXusOntology: An ontology of NeXus format definitions (the dominant metadata model with the PaN domain). NeXusOntology is a representation of the formal NeXus definitions and is created automatically from NeXus definition files.
- PaNmapping: A semantic mapping ontology for the Photon and Neutron Science domain. This ontology aims to map between the overlapping entities in the most common PaN metadata schemata NeXus format and CSMD/ICAT data schema, as well as integration with the DCAT v2 Ontology and DublinCore.

PaNET is accessible in BioPortal³⁵. It is currently used:

- in the federated PaN search API via the ontology service, developed as part of T3.3, which (in turn) is used by the PaNOSC data portal;
- in the PaN training platform via the ontology service;

³⁴ The raw data is available at <https://doi.org/10.5281/zenodo.7851972>

³⁵ <https://bioportal.bioontology.org/ontologies/PANET>



- at Diamond to categorise publications.

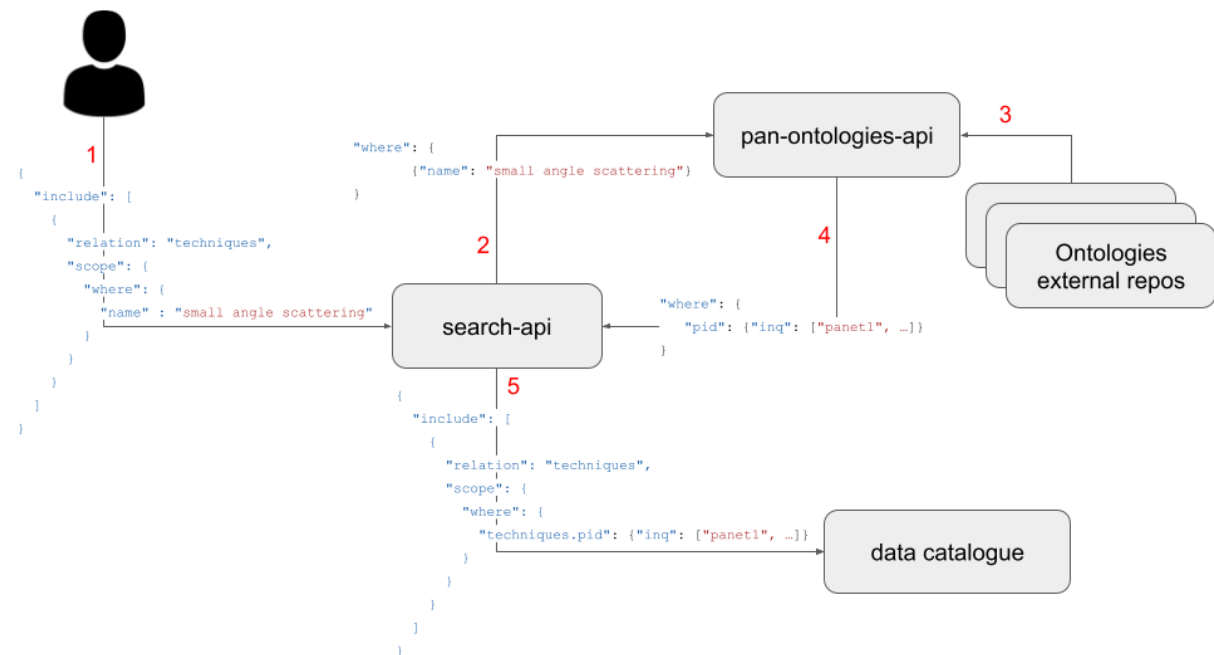
The NeXus ontology has been completely handed over to the NeXus international advisory committee³⁶, thanks to our partners' engagement.

This work is detailed in D3.2³⁷.

4.2. Implement ontologies in metadata catalogues

PSI led the implementation of the pan-ontologies-API, a RESTful service that links the PaNET ontology to the search-API. At a high level, the service works by pulling and caching information from external ontology repositories, and in particular, translating the query from the user input (e.g. in data.panosc.eu search portal³⁸) to a new query sent to the search-API after having added the logic from the PaNET ontology.

Our goal with PaNET is to expand the user's query to include the ontology graph from PaNET. In particular, queries will include all the descendants of the user defined technique.



pan-ontologies-API data flow for the PaNET ontology

To foster further its federation, PSI provided such a service in a federated fashion³⁹, enabling all services of all work packages of the ExPaNDS and PaNOSC members to use it.

PSI extensively shared the pan-ontologies-API with dedicated meetings and training materials.


³⁶ <https://www.nexusformat.org/NIAC.html>

³⁷ <https://doi.org/10.5281/zenodo.4806026>

³⁸ <https://data.panosc.eu/>

³⁹ <https://pan-ontologies.psi.ch/>





PAUL SCHERRER INSTITUT
PSI

Photon and Neutron Techniques Ontology Service

PaNET Ontology Service

The PaN technique ontology service is a software that allows clients to expand a photon and neutron technique (PaNET) term into an equivalent search query that includes all PaNET sub-term: those that are more specific research techniques. It can also display the PaNET ontology.

Organisation: [Paul Scherrer Institute](#)

☆☆☆☆☆ (0.0 / 5) 0 reviews Add to comparison Add to favourites

Access the service

🔓 FULLY OPEN ACCESS

PaNET Ontology Service in the EOSC marketplace⁴⁰

To further deploy the PaNET ontology in community catalogues, HZB also provided mechanisms to populate the ICAT technique table with the terms from PaNET.

4.3. Coordinate metadata catalogues and data life cycle

HZB, in collaboration with ExPaNDS and PaNOSC partners, led the dissemination of best practices and standards around data life cycle in PaN, through periodic meetings and webinars. In particular, the webinars on data catalogues, EOSC integration, PaNET and metadata ingestion from experiments were well received and proved to be extremely useful, as testified by the improvement of WP3 KPIs.

PSI also developed a self-contained, stand-alone metadata catalogue release that facilities could download to test and try-out locally⁴¹. PSI and ESS contributed to the SciCat catalogue and Diamond and UKRI to ICAT.

Effort was also put into “informal training”, i.e. direct collaboration between ExPaNDS and PaNOSC members, on an on-demand basis, strengthening the relationships between facilities and proving their cooperation towards the common goal of FAIRness.

The level of maturity of facility data catalogues to enable FAIR services was monitored to assess overall and individual progress in the roadmap set in the first period⁴². Approximately every six months, PSI circulated these questions to all partners and collected and analysed the results.

4.4. Integrate metadata catalogue services into EOSC

MAX IV led the dissemination of guidelines for the integration of the PaNOSC data portal into the EOSC and the use of Umbrella-ID, the PaN AAI infrastructure. The main activities performed in this task were:

- PSI, EGI and HZB co-organised the training workshop on EOSC with WP1, notably providing presentations and material explaining the functionality of OAI-PMH and how to implement this protocol in SciCat and ICAT;
- Deploying the SciCat and ICAT OAI-PMH components, developed in the first period of the project, in the facility data catalogues (Elettra, HZDR, ISIS, MAX IV);
- Deploying the PaN federated search API endpoint in the facility data catalogue (Elettra, HZB, ISIS, MAX IV and PSI);
- Providing a production endpoint for the PaNOSC data portal (PSI and MAX IV).

⁴⁰ https://marketplace.eosc-portal.eu/services/eosc.psi.panet_ontology_service

⁴¹ <https://doi.org/10.5281/zenodo.5205909>

⁴² <https://doi.org/10.5281/zenodo.4146819>



Details on the demonstration of ICAT and SciCat, along with APIs, compatible with ExPaNDS federated EOSC services can be found in D3.3⁴³.

Thanks to WP3, by the end of the project, our facilities have taken a big step forward in making their open data available in EOSC:

ExPaNDS Partner Facility	PaN catalogue search endpoint	Data Catalogues with the OAI-PMH protocol
ALBA	N/A	N/A
DESY	N/A	N/A
Diamond	N/A	N/A
Elettra	In place	In place
HZB	In place	In place
HZDR	N/A	In place
ISIS	In place	In place
MAX IV	In place	In place
PSI	In place	In place
SOLEIL	In place	In place

Status of ExPaNDS facilities metadata catalogues in EOSC

4.5. Training material

SOLEIL led the creation and cataloguing of the portfolio training material in WP3, by providing a description and references of the available material at the time of D3.4⁴⁴, specifying the target group for each one of them. This work has been facilitated by the use of the PaN-training catalogue, developed as part of WP5.

WP3 constantly updated and added training material and workflows to the PaN-training portal and co-organised with WP1, WP5 and WP6 a total of 4 training workshops (PSI, SOLEIL, DESY, EGI, HZB, HZDR, UKRI and MAX IV).

4.6. Deliverables and milestones of the WP for the period

#	Deliverable	Partner	Due date	Actual date
3.2	Release V1.0 ExPaNDS ontology available as an EOSC online service	PSI	31/05/2021	04/06/2021

⁴³ <https://doi.org/10.5281/zenodo.6363591>

⁴⁴ <https://doi.org/10.5281/zenodo.6390324>



3.3	Demonstrate ICAT and SciCat released with APIs compatible to ExPaNDs federated EOSC services	PSI	28/02/2022	16/03/2022
3.4	Portfolio material to support training for target groups at all Research Infrastructures (RIs)	PSI	31/03/2022	11/04/2022

#	Milestone	Partner	Due date	Actual date
12	ExPaNDs Metadata Catalogue Release	PSI	31/08/2021	24/08/2021
13	Metadata catalogue as EOSC service	PSI	28/02/2022	28/02/2022

5. Work package 4: EOSC data analysis services for PaN national RIs

WP4 provided Photon and Neutron users with the ability to find and run analysis workflows with EOSC aligned data services. We focused on five challenging data analysis pipelines implementation as remote data analysis services taking into account the needs of representative scientific communities while supporting the diversity of the institutions' existing computing infrastructures. Each of these five use cases has been developed by a different facility.

The coordination of efforts by ExPaNDs helped its partner-facilities to provide PaN scientists in Europe with standardised data analysis pipelines available through the EOSC portal and to conceive an overall modular architecture around generic APIs. They also opened the road to enlarge the audience of facilities using the VISA portal⁴⁵.

5.1. Alignment of PaNOSC services

During the second period of the project, the alignment with PaNOSC focused on the deployment of VISA at the national RIs. WP4 enhanced VISA's portability, allowing it to be deployed at national PaN facilities and integrated with their diverse cloud technologies, notably using Kubernetes (DESY, ALBA)⁴⁶.

Meetings on VISA were biweekly held, gathering all ExPaNDs and PaNOSC contributors to discuss the alignment with the requirements and needs of the partner facilities, distribute the development effort and share facility experiences.

This task led to the elaboration of a memorandum of understanding (MoU) between PaNOSC and ExPaNDs facilities, steered by SOLEIL, to continue the collaboration under a formal contract after the projects end.

⁴⁵ <https://doi.org/10.5281/zenodo.7788840>

⁴⁶ ILL is now contemplating this Kubernetes deployment for their own use, demonstrating the benefit of having VISA as an open source community project already.



WP4 also participated in the workshop organised with PaNOSC⁴⁷ to discuss the integration of the new UmbrellaID AAI into the infrastructures of our facilities.

5.2. Testing and validation framework

The goal of this work was to establish a framework that assures that the selected data analysis services can be validated against reference data sets.

MAX IV developed a framework for testing and deploying software that is made available for users across research facilities with heterogeneous compute infrastructures. It focuses on validating Jupyter notebook services as a concrete example.

This work was based on the best practices introduced by the EOSC Synergy project⁴⁸. The testing and validation framework for Jupyter notebooks is set up and in use at MAX IV and can be easily configured to work at other ExPaNDS partners' infrastructures. The code repositories are publicly available, including the associated docker images, the code used to produce them and a setup deployable on HPC infrastructures⁴⁹.

This task is described in D4.3⁵⁰ and was the object of our contribution to the workshop organised by EOSC-Synergy at the EGI conference 2021 "Bridging from theory to practice with EOSC-Synergy" where we were invited to present our quality approach as a use case of the use of Synergy's outputs.

5.3. Prototype data analysis services

As part of the overall work plan, prototype analysis services have been set up and tested in different configurations (through Jupyter notebooks or ready to use remote desktops) to provide users with software, tools and workflows to analyse data associated with five out of eleven selected use cases described in deliverable D4.2 on reference datasets⁵¹.

State-of-the-art mechanisms to ease the exchange of data analysis solutions between our national RIs for running different data analysis pipelines representing particular experimental techniques were documented, with a special section dedicated to describe dependencies on core services shared with our partner project PaNOSC.

We demonstrated that, for the selected use cases, we were able to build analysis pipelines based on state-of-the-art cloud technologies (OpenStack, Kubernetes, Proxmox, etc.) which we deployed and ran at multiple computing infrastructures across our partners.

Completion of this task is described in D4.4⁵² and was carried out jointly between DESY and SOLEIL, with contributions from HZDR, EGI, PSI, MAX IV, Diamond and ALBA which all worked on their home infrastructure to align with the project's roadmap.

⁴⁷ <https://indico.psi.ch/event/12701/>

⁴⁸ <http://dx.doi.org/10.20350/digitalCSIC/12543> and <https://doi.org/10.20350/digitalCSIC/12533>

⁴⁹ <https://gitlab.com/MAXIV-SCISW/JUPYTERHUB>

⁵⁰ <https://doi.org/10.5281/zenodo.5718671>

⁵¹ <https://doi.org/10.5281/zenodo.4558708>

⁵² <https://doi.org/10.5281/zenodo.6305000>



5.4. Deploy data analysis services into EOSC

Completion of this task, which comprises the core effort of WP4 and draws on all previous tasks and deliverables, is described in D4.5⁵³. The main activities which led to the deployment of the five typical analysis workflows into the EOSC were:

- make the five use cases facility-agnostic, by either creating portable Jupyter notebooks or bundling software into containers for execution at any facility, providing the right environment
 - PSI, HZDR, Diamond, ISIS and DESY;
- test the workflows on different infrastructures
 - PSI, DESY, Diamond, HZDR, EGI;
- register as an EOSC provider
 - ALBA, Diamond, EGI, Elettra, HZDR, MAX IV, PSI, SOLEIL and ISIS;
- publish a data analysis service in the EOSC marketplace
 - DESY, Diamond, EGI, PSI and ISIS;
- document the process to execute the pipelines for the users
 - PSI, HZDR, Diamond, ISIS, DESY, EGI and SOLEIL;
- publish as a workflow in the PaN-training.eu catalogue
 - PSI, HZDR, ISIS, DESY and SOLEIL.

Technique	Partner	PaN Training	EOSC Portal access
4D Full field tomography	PSI	Yes	PSI, DESY, Diamond
Terahertz Spectroscopy	HZDR	Yes	Via PaN Training Portal
Electron scanning diffraction imaging	Diamond	No	Diamond Remote Desktop Service
Small-angle neutron scattering	UKRI	Yes	ISIS Data Analysis as a Service, DAaaS
Serial crystallography	DESY	Yes	DESY VISA Portal Service

Five ExPaNDS use cases deployed in EOSC

⁵³ <https://doi.org/10.5281/zenodo.7417726>



1.5. Run sorting raw data, choose one option

The main part of the sorting process is to correct the timing of each signal pulse. For each pulse the timing jitter is determined and assigned.

```
In [5]: %matplotlib inline
        """Run single position by providing the mirror position in mm, written in the filename"""
        #sortbin.sorting(position='51.1')

        """Run all position"""
        sortbin.sorting()
```

Data folder: 038_0p7THz_LSCO_2mmZnTe_31K_13.06.2018_0535_02
 Sorted file: 2018/2018-June_datasorting/2018-06-13/038_0p7THz_LSCO_2mmZnTe_31K_13.06.2018_0535_02_sorted.dat
 Start sorting
 Time elapsed for preparation: 0.04 s
 "Check "jitter038_0p7THz_LSCO_2mmZnTe_31K_SD_101523Hz_51.10mm13.06.2018_0535_02.png" if fit is good.
 Time elapsed after multiprocessing sorting: 388.37 s
 Time elapsed after writing sorted file: 407.08 s
 Sorting done

1.6. Run binning data

The image shows two plots. The left plot, titled 'Determine jitter time', shows 'signal / a.u.' on the y-axis (0 to 2000) and 'time / pixel' on the x-axis (0 to 2000). It displays a sharp peak at approximately 1000 pixels, with a blue line for 'raw data' and an orange line for 'butter fit'. The right plot, titled 'Sorted data', shows 'signal / a.u.' on the y-axis (-0.20 to 0.20) and 'pixel' on the x-axis (31000 to 35000). It displays a dense, rectangular distribution of blue data points.

Analysis step of the TELBE spectroscopy notebook on the HZDR jupyter notebook service

5.5. Deliverables and milestones of the WP for the period

#	Deliverable	Partner	Due date	Actual date
4.3	Testing and Validation framework	DESY	31/08/2021	30/11/2021
4.4	Analysis Services	DESY	28/02/2022	28/03/2022
4.5	Deployment of Analysis Services for EOSC within the EOSC-hub	DESY	31/12/2022	30/12/2022

#	Milestone	Partner	Due date	Actual date
14	Analysis services prototypes	DESY	28/02/2022	28/03/2022
15	Analysis services for EOSC	DESY	31/12/2022	30/12/2022



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

6. Work package 5: Training activities through EOSC platforms

WP5 aimed to educate and train staff and users from PaN Research Infrastructures to exploit the services developed during the project, by organising a series of events and delivering training materials on a dedicated platform.

6.1. Organising at national RIs a series of staff and users training events

WP5 organised a series of training events to promote ExPaNDS outcomes. Events were organised in close collaboration with the technical work packages (WP2, WP3 and WP4), which provided material and defined the key needs, and with WP6, which oversaw organisational and promotional aspects. These training activities focused on the following topics: FAIR principles (already provided in the first period), EOSC services, metadata catalogues, the PaNET ontology and the VISA data analysis infrastructure. PaNOSC partners were also involved in the co-organisation or participated actively in these workshops:

- The second ExPaNDS workshop took place in April 2021 (EOSC Workshop, 6th and 7th April 2021).
- A hands-on session on the PaN ontologies API (Online, 17th February 2022) and a Workshop on Metadata Catalogues (Online, 4th April 2022) were organised and allowed us to create a dedicated learning path which gathers three workshops and related materials, such as videos, guidelines and repositories.
- To continue the promotion of main ExPaNDS WP3 results (catalogues, APIs, ontology...), an open session was organised (hybrid meeting, Saint Aubin, 3rd November 2022) as satellite of an internal ExPaNDS/PaNOSC meeting.
- We organised three workshops on the demonstration, set up and distribution of the VISA portal:
 - on the 14th of June 2022, Dolní Břežany (hybrid)
 - on the 16th of September 2022 (Online) and
 - on the 22th to the 28th of February 2023, SOLEIL (hybrid).
- A cross-WP workshop titled “Follow a user story! A journey through the PaN data services” took place in Hamburg on the 24th January 2023 (hybrid) and gathered 5 key presentations: FAIR journey, EOSC demo, PaN-training platform demo, VISA portal and Search API/Harvesting presentations. The workshop was organised as a satellite event of the Users’ annual meeting at DESY/EuXFEL and was part of the ExPaNDS closing event.

In addition, local events were held with users and staff at the various facilities, e.g., HZDR, and SOLEIL. Dedicated meetings between the WP5 team and ExPaNDS facilities’ staff have also been organised to work on the use of workflows in PaN-training.eu, resulting in a significant uptake on this feature, especially within WP3 and WP4.

The two annual reports published in the second period (D5.2⁵⁴ and 5.3⁵⁵) allowed close monitoring of the status of this task, including the training plan set up following a mid-term recommendation and continuously updated since.

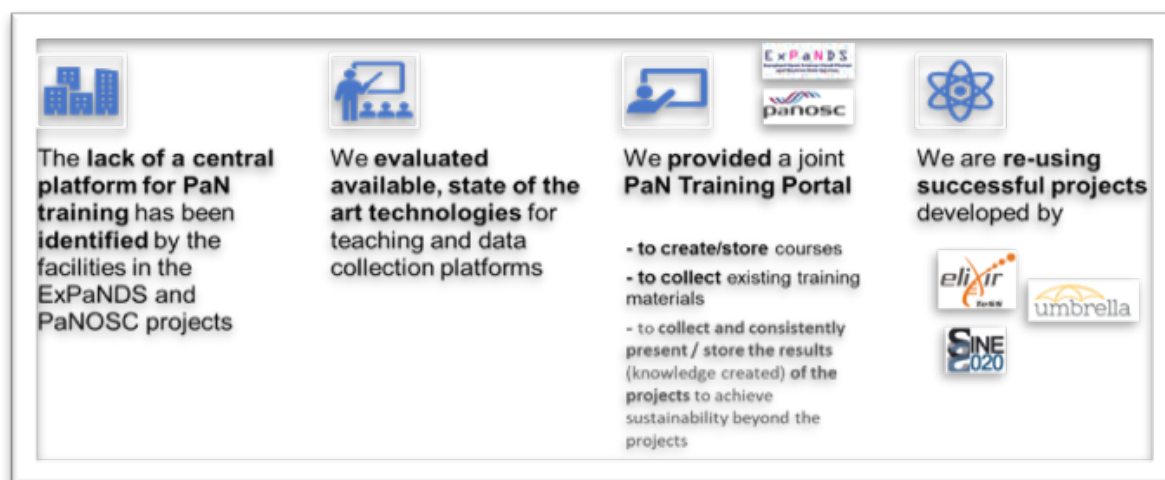
⁵⁴ <https://doi.org/10.5281/zenodo.5468567>

⁵⁵ <https://doi.org/10.5281/zenodo.7022922>



6.2. Evaluation of available e-platforms and provision of a training demonstrator

The evaluation of the available e-platforms and provision of a training demonstrator was summarised in D5.4⁵⁶ and the description of the platform and its content is available in D5.5⁵⁷.



Summary of ExPaNDS T5.2 goals and results

The PaN training portal was officially released in August 2022. This first version of the portal provided a catalogue for:

- the registration of project and third-party training events;
- training materials such as tutorials, videos, repositories;
- training workflows to show learning steps and dependencies between materials.

Thanks to the alignment of WP5 with PaNOSC, the portal offers a tightly coupled e-learning system (Moodle) with PaN-specific learning courses and Jupyter notebook support.

The PaN-Training catalogue has been registered as a service in the EOSC⁵⁸, enabling the training content to a wider user base beyond the PaN community (HZDR).

The training catalogue is based on TeSS⁵⁹, an established solution developed by the ELIXIR community. With the catalogue, ExPaNDS partners are strengthening the role of EOSC services for the PaN scientific field, with a focus on providing a sustainable training, learning and documentation platform for the PaN community, including initiatives such as LEAPS, LENS, and Laserlab Europe.

Several features were added during the second period of the project, including:

- integrating the PaNET ontology in the platform's metadata schema;
- enhancing the content's visibility in Google;
- offering DOIs for materials, using Zenodo's API;
- enabling login with UmbrellaID.

⁵⁶ <https://doi.org/10.5281/zenodo.5171766>

⁵⁷ <https://doi.org/10.5281/zenodo.7023247>

⁵⁸ <https://marketplace.eosc-portal.eu/services/pan-training-catalogue>

⁵⁹ <https://tess.elixir-europe.org/>



The training catalogue was a joint development between HZDR and SOLEIL.

6.3. Collection, Preparation and Publishing of Training Material

ExPaNDS developed training content covering the main results of the project. All training materials relevant to the PaN community have been uploaded to the catalogue:

- WP2 produced materials related to FAIR data, including workshops and guidelines. 13 media products were developed in the framework of this WP and are available via the platform: “Promotion of FAIR principles, data policy framework, data management”;
- In addition to all materials from WP3 (13), a workflow showcasing a “Learning path: Metadata catalogue services including PaN ontologies” has been created in the platform covering the data services;
- Workflows detailing the use cases described in deliverable D4.4 (see WP4) have been created. ExPaNDS WP5 in conjunction with WP4 worked on the development of these training workflows to better display the dependencies and connections between the different services and data pipelines involved in all the reference analysis services in EOSC, in the PaNOSC data portal or in RI’s dedicated services.

The full list of materials (workshops, guidelines, GitHubs, Wikis, posters ...) developed in the framework of the ExPaNDS project can be found in the project Training Activity plan in the Annex of D5.6⁶⁰.

As of February 2023, there are 190 materials, 324 events and 11 workflows in the training catalogue. Resources are either created manually or integrated in the platform via scrapers. The majority of these materials have been tagged as videos (82), followed by slides (24), git/Git/github/Git projects (15), documents (14), tools (9) and jupyter notebooks (6). These products are complemented by 48 Moodle courses from the e-learning platform. Most of the materials are provided with a DOI.

The materials in our catalogue are described by a metadata schema following the RDA metadata recommendations⁶¹ with additional elements from the IEEE LOM⁶² and the ENVRI⁶³ training catalogues as described in our deliverable D5.4⁵⁶. The classification of the materials is the responsibility of the submitting user or a curator of our community. In this context, we integrated the PaNET ontology developed by WP3 as an optional field *scientific topic* in our metadata schema. Specifically, the user can find materials with a particular scientific topic and any of its children.

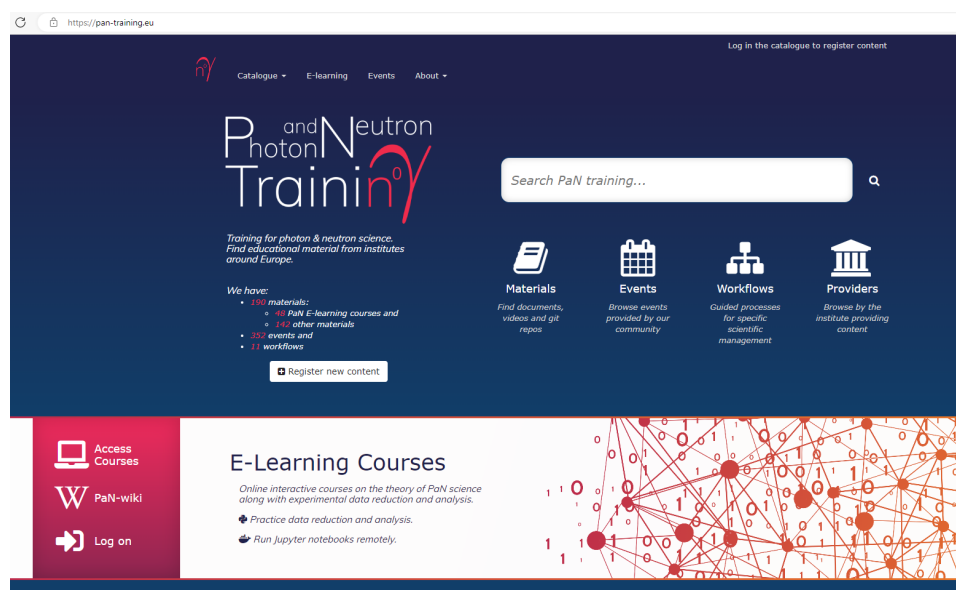
⁶⁰ <https://doi.org/10.5281/zenodo.7642688>

⁶¹ <https://doi.org/10.15497/RDA00073>

⁶² IEEE Standard (IEEE 2002) for Learning Object Model (LOM)

⁶³ <https://doi.org/10.5281/zenodo.3903340>





PaN training platform at pan-training.eu

The collection, preparation and publication of training material was a joint effort between HZDR and SOLEIL.

6.4. Monitoring and Measuring of training effectiveness

6.4.1. Workshops organisation

Due to the COVID-19 pandemic, ExPaNDS had to rethink the way workshops were held. This is the reason why workshops were done remotely or in a hybrid mode, allowing attendees to stay in their facilities for safety reasons (during lockdown periods) and were also useful to allow participants that would otherwise not be able to join. Furthermore, these workshops were recorded, making them re-usable, and easily findable on the PaN-training platform.

Surveys were distributed to the participants after each workshop and the number of attendees were collected for our KPIs.

6.4.2. PaN-training.eu

The first prototype of the training catalogue was released and presented (live demonstration) during the mid-term review meeting in June 2021. In retrospect, it was the right decision to open the catalogue to the public as early as possible, because this enabled us to motivate our scientist to evaluate the platform and register first content early in the project. With the planned release date more than one year later, the number of materials would be significantly lower.

The PaN-training platform was presented at various events, such as the EGI Conference 2022, the EOOSC Symposium 2022, the LEAPS plenary assembly 2022 and the 18th ESUO general assembly⁶⁴, to get feedback from PaN users and the wider community. In addition,

⁶⁴ See complete list of events with ExPaNDS participation in the appendix.



possible joint developments were discussed with ELIXIR and representatives of the EOSC. As a direct result of the EOSC Symposium 2022, we moved forward with the final registration of PaN-training as a service within the EOSC marketplace.

GDPR-compliant analytics were set up in the training catalogue to measure usage of the platform and its different components.

The outcomes of this task are presented in D5.6⁶⁰.

6.5. Deliverables and milestones of the WP for the period

#	Deliverable	Partner	Due date	Actual date
D5.2	Annual report on training activities, workshops held, materials published and investigations performed	SOLEIL	31/08/2021	01/09/2021
D5.3	Annual report on training activities, workshops held, materials published and investigations performed	SOLEIL	31/08/2022	16/09/2022
D5.4	Demonstrator for using of e-learning platforms	SOLEIL	30/09/2022	30/09/2022
D5.5	Dedicated websites and e-platforms with the teaching materials	SOLEIL	31/08/2022	16/09/2022
D5.6	Report on lessons learned and future prospects for adopting best practices	SOLEIL	31/01/2023	20/02/2023

#	Milestone	Partner	Due date	Actual date
17	Training websites and e-platforms online available	SOLEIL	31/10/2022	27/09/2022

7. Work package 6: Dissemination and outreach

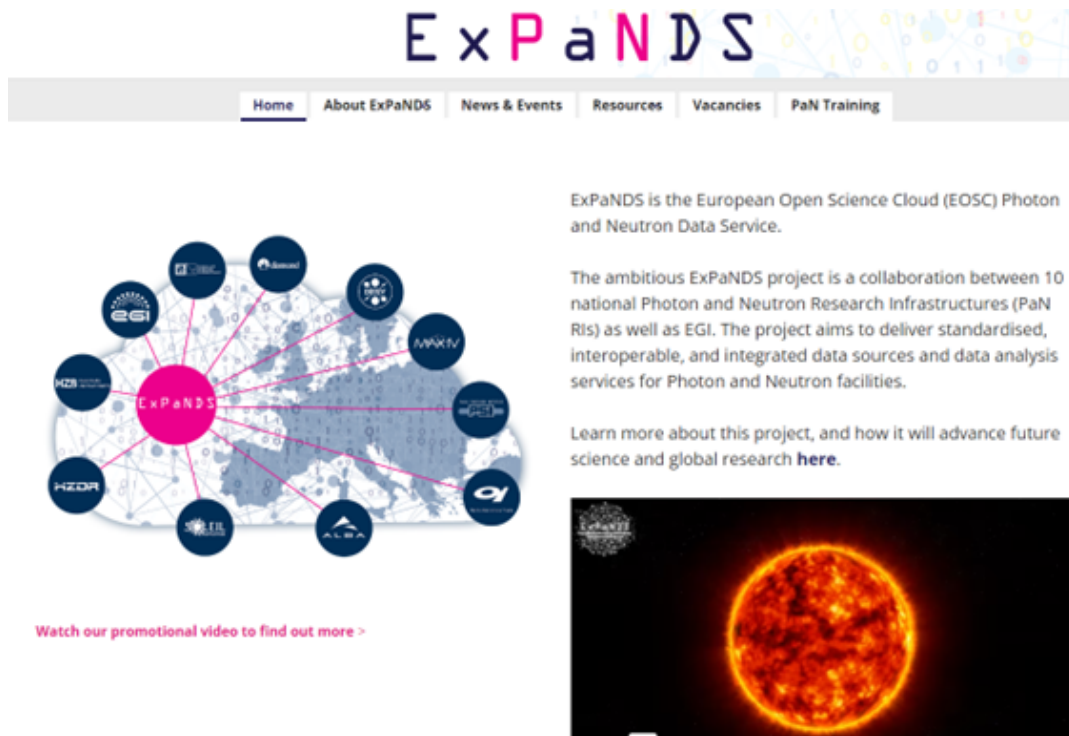
WP6 set up the communication and dissemination plan of the project and supported the internal and external communication within and between the work packages and also for webinars and conferences.

Awareness of the EOSC services was raised and all stakeholders at the Photon and Neutron RIs were engaged.



7.1. ExPaNDs internal and external communications strategy and delivery

- Website (D6.02) - www.expands.eu



Screenshot of the expands.eu website

Our website received a number of views and was visited by over 10,000 visitors during the 42 months of our grant.

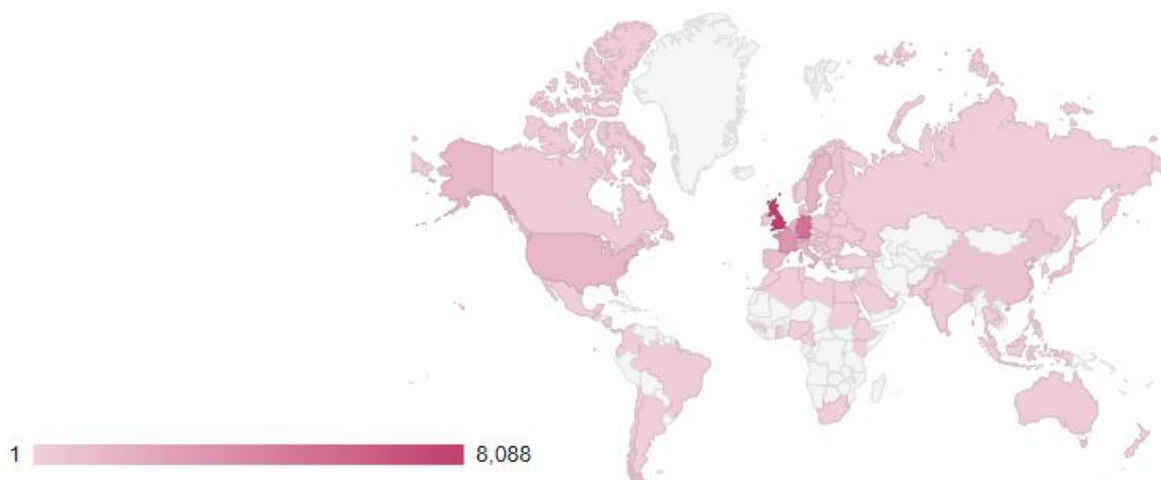


Image and statistics sourced: ExPaNDs website Wordpress - My stats - all time at 13.03.2023

Total views:	27,149
Total visitors:	10,520
Average monthly views:	250



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

As you can see from the below top ten countries chart, most of visitors were from across Europe:

1. United Kingdom	8,088 views
2. Germany	5,182 views
3. France	3,146 views
4. Sweden	1,343 views
5. Italy	1,332 views
6. United States	1,215 views
7. Switzerland	950 views
8. Spain	913 views
9. Netherlands	687 views
10. Finland	651 views

- **Twitter account:** @ExPaNDs_EU

We've conducted a total of 360 tweets until the end of February 2023 and averaged 8 tweets a month, which has resulted in over 222,000 impressions with over 9,600 profile visits throughout the duration of the grant.



Image and statistics sourced: ExPaNDs Twitter analytics until end of Feb 2023

- Creating an 8-minutes ExPaNDs **promotional video** on FAIR data for PaN science, with beautiful aerial images of our member facilities, starring senior scientists and project members⁶⁵.

⁶⁵ <https://player.vimeo.com/video/637958452>





ExPaNDs Promotional Video

1 year ago | Plus

Diamond Light Source + Suivre

More from Diamond Light Sou

Lire automatiquement la vidéo s



ExPaNDs promotional video

- A series of 7 **user interviews**⁶⁶ was produced to present the benefits of ExPaNDs from different perspectives and key stakeholders, covering:
 - *Heritage Science* with Dr Eleanor Schofield
 - *Tomography* with Dr Roberto Volpe
 - *Data for the Future* with Professor Dr Volker Guelzow
 - *Pandemic Preparedness* featuring Dr Annette von Delft and Dr Ed Griffen
 - *Industry* with Dr Kamel Madi
 - *Connecting data* with Professor Kristina Edstrom
 - *AI Drug Discovery* featuring Dr Alpha Lee and Dr Ben Perry

1. Heritage Science



Video link: <https://vimeo.com/672837958/210d727e76>

3. Sustainability



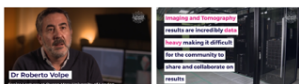
Video link: <https://vimeo.com/672819190/23a91bd4c0>

5. Industry



Video link: <https://vimeo.com/672818617/f9ba076823>

2. Tomography



Video link: <https://vimeo.com/672819456/ba18a8a3ee>

4. Pandemic Preparedness



Video link: <https://vimeo.com/672818926/a7e7d07c0d>

6. Connecting Data



Video link: <https://vimeo.com/672818339/8f56566aae>

7. AI Drug Discovery



Video link: <https://vimeo.com/672818081/f774b3637d>



User interviews

- Creating **senior management engagement interviews**⁶⁷ with facility member senior managers and directors.
 - Professor Dr Helmut Dosch from DESY
 - Dr Caterina Biscari from ALBA
 - Professor Roger Eccleston from ISIS Neutron and Muon Source
 - Professor Andrew Harrison formerly from Diamond Light Source

⁶⁶ <https://expands.eu/videos/one-minute-user-videos/>

⁶⁷ <https://expands.eu/senior-engagement-interviews/>



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



Senior management engagement interviews

- Showing the different **use cases** identified among PaN facilities by ExPaNDs during the grant period⁶⁸:
 - Reflectometry Case Study
 - Tomography Case Study
 - ExPaNDs COVID Case Study
 - Full field Tomography Case Study
 - Kinetic SAXs Case Study
 - Small Angle Scattering Case Study
 - Serial Crystallography Case Study
 - Terahertz Spectroscopy Case Study
- Topic-based **webinars** including updates from WP either live or recorded⁶⁹:
 - *Life Sciences* with Professor Dr John Helliwell
 - *Cultural Heritage* with Professor Dr Anjali Goswami
 - *Tomography & Imaging* with Dr Mark Basham
 - *Industry as data users* with Dr Kamel Madi *et al.*
 - *General update* with Dr Adrian Mancuso
 - *WP6 Update* to global Lightsources.org community
- Published article in EU Researcher magazine⁷⁰

Finally, we have also submitted (Feb 2023) an open access publication article in the Synchrotron Radiation News Magazine (SRN), which is due to be published in April 2023⁷¹, outlining the key achievements of our grant.

⁶⁸ <https://expands.eu/resources/use-cases/>

⁶⁹ <https://expands.eu/2023/01/26/expands-topic-based-webinars/>

⁷⁰ https://issuu.com/euresearcher/docs/expands_eur31_h_res

⁷¹ <https://doi.org/10.1080/08940886.2023.2186664>



Supporting open science

The ability to share information is crucial to scientific progress, yet it is not easy to manage and share the vast amounts of data generated at Photon and Neutron Research Infrastructures (PaN RiS). We spoke to **Professor Dr Patrick Fuhmann** and **Dr Sophie Servan** about the work of the ExPaNDS project in creating a framework for effective data management and supporting open science.

The PaN RiS found across Europe play an important role in scientific research, helping scientists from a wide variety of different disciplines gain deeper insights into major questions in their fields. Facilities like DESY, HZDR and HZB in Germany, ALBA in Spain, MAX IV in Sweden, PSI in Switzerland, Elettra in Italy, SOLEIL in France, and Diamond Light Source, UKRI / STFC in the UK are home to sophisticated equipment which are used to generate photon and neutron beams. "What we do at DESY is accelerate electrons. We have mechanisms to produce electrons from high-energy, intense rays of very brilliant photon beams which are extremely short," explains Professor Patrick Fuhmann, Group Leader of "Research and Innovation" at DESY, an accelerator centre based in Hamburg. The wider societal benefits of these facilities lies in what is done using these bright beams, Professor Fuhmann says experiments are conducted at the very end of research what is called a beamline. "The beamlines can be used for experiments on batteries, quantum technology and medical applications for instance. Essentially, they are relevant for any topic where you need to look at processes which are occurring on very small scales," he outlines.

ExPaNDS project
These types of experiments often generate enormous amounts of data, mainly in the form of images, which can become progressively more difficult to manage and share with other researchers as the volume increases. This issue lies at the heart of the ExPaNDS project, a European Union's Horizon 2020 research and innovation programme grant which brings together 10 European PaN RiS, as well as the European Grid Infrastructure (EGI) federation. "Our goals in the project are two-fold. First of all, we want to create a policy framework for the facilities, so that they have a certain way of processing the data on the policy level. This is about agreements between a facility

and their customers - the scientists who produce the data," says Professor Fuhmann. A second major goal in the project relates more to the technical level. "We discuss possible formats for the data with the different facilities, and we encourage them to use the same formats. This relates also

to the meta-data of the data, such as the frequency, the energy level and the time," continues Professor Fuhmann. "We want to encourage the facilities and the beamlines to record this data in a similar fashion, so that it can be more easily understood later on." The data and meta-data captured from samples at these Photon and Neutron (PaN) facilities are mainly recorded in a standardised format called Nexus. This is part of the


wider goal of ensuring that data is shared according to the FAIR (Findable, Accessible, Interoperable, Reusable) principles. "Meta-data fields have been defined and agreed, and this enables interoperability," says Dr Sophie Servan, leader of the Management and Sustainability work package within ExPaNDS.

On the technical level, the standardised file format ensures that different applications can open files, while Professor Fuhmann says it's also important that there is a common vocabulary used in association with the data. "We are trying to convince everyone to use the same key words, which is important to help Artificial Intelligence (AI) applications scan the data effectively," he stresses. "The data relevant to a researcher is stored

somewhere, but it may be difficult to find it. The meta-data catalogue holds only the most important key words of data. This catalogue holds all the key words of all the data generated by these PaN experiments in Europe, and researchers can search it."

A researcher may be interested in the lung tissue of a rat for example. They are able to search the catalogue for relevant experiments, using clear search terms, and they then receive a list of data, as well as details about where the experiment was conducted. "A researcher can download the image in that Nexus format, so they are better able to understand what it represents," explains Professor Fuhmann. There is also the possibility to add data from experiments conducted in the past, although Professor Fuhmann says the best solution may be to simply do the experiment again, with the new generation of synchrotrons set to give researchers even deeper insights into the nature of matter. "The brilliance of the beams at the current facilities is increasing so fast that in most cases it is cheaper to simply repeat the experiment rather than convert the old data," he stresses. "In many areas of science an experiment can be conducted again and again, and with improved facilities, it can be done better."

Accelerating scientific progress
The wider aim here is to help accelerate scientific progress by encouraging data sharing through the European Open Science Cloud (EOSC). Making data more widely available helps your scientific progress, as has been shown by increasing levels of access to the internet. "The more you make research available to everyone, the faster science accelerates," says Professor Fuhmann. Nevertheless, it is not always easy to encourage scientists to share their findings and adopt a less proprietary attitude to their research by putting the FAIR principles at the heart of policy frameworks, the project aims to make data-sharing the default. "The users of the synchrotrons may not even know what's happening, but they can benefit from this commitment to the FAIR principles," outlines Dr Servan. "We cannot always change the attitude of the scientists, but we can change a little bit the attitude of our management, and we can make the routine work of incorporating the data into the catalogue automatic. A scientist doesn't have to think too much about how they can meet the FAIR principles, they can focus their energy on their research." As soon as an experiment starts and a detector is generating data, the most important meta-data is extracted and then automatically put into the catalogue. The general policy is that all of the data generated in experiments should be included in the catalogue, which can then help guide research in the future. "Some scientific researchers use published data to set the initial parameters of models, for example in spectroscopy. The more access they can get to accurate, well-described data from past experiments, the better they can fine-tune their model," says Dr Servan. PaN facilities are also an important research tool across a wide range of other areas, from climate change, to energy, to quantum computing. While DESY itself is not directly involved in the development of quantum computers, Professor Fuhmann says research at the facility can help lay the foundations for further progress. "For example, we're not involved in the creation of an atom, but we help to make sure that the material is flexible, sufficiently strong, and doesn't age too fast," he outlines. "We can investigate the properties of materials at DESY." The situation with quantum computers is more complex, as researchers have to consider quantum effects and how far certain materials are able to reflect them. However, the aim is still to evaluate whether a given



48 EU Research www.euresearcher.com 49

Article in EU researcher magazine on ExPaNDS

The tomography case study developed in ExPaNDS gathered interest and was referenced in a number of different publications:

- International Labmate Online⁷²
- Institute of Materials, Minerals & Mining⁷³
- Environmental Magazine⁷⁴
- Bioenergy Insight⁷⁵
- Knowledia⁷⁶

7.2. ExPaNDS' marketing and communication toolbox

In ExPaNDS second period, we expanded the marketing toolbox with updated presentations for conferences and a reusable bottle with ExPaNDS branding.

We also continued rolling out the roadmap and vision, in particular:

- delivering presentations to relevant large international conferences⁷⁷,

⁷²

<https://www.labmate-online.com/article/microscopy-and-microtechniques/4/diamond-light-source-ltd/g-etting-synchrotron-and-neutron-big-data-the-fair-way/3261>

⁷³ <https://www.iom3.org/resource/a-close-up-of-biochar-porosity.html>

⁷⁴ <https://environmentalmag.org/news/scientists-unravel-the-porosity-of-biochars-in-a-world-first/>

⁷⁵ <https://www.bioenergy-news.com/news/qmul-research-could-lead-to-biochar-breakthrough/>

⁷⁶

<https://news.knowledia.com/us/en/articles/queen-mary-scientists-first-to-unravel-porosity-of-biochars-at-diamond-916190c551161163339f8cc2f5a3991d6c2c6493>

⁷⁷ <https://doi.org/10.5281/zenodo.7464347>



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

- organising a librarian symposium with WP2, and reporting discussion points and key insights gained during the event hosted on the 30th September 2021⁷⁸.

Our attendance of AAAS in 2023, after the grant closure, yielded a number of fruitful conversations with the most notable being with a Program Scientist and Program officer with NASA's Transform to Open Science (TOPS) programme, inviting ExPaNDS to deliver a 15 minute talk about our grant and findings.

7.3. Formulation and dissemination of metrics and refinement of KPIs

As an update of the introduction to a common taxonomy produced in the first period, WP6 created a benchmarking of terminologies used within SRAO, EuroSciVoc and OECD subject vocabulary against the proposed PaN discipline taxonomy⁷⁹.

ExPaNDS surveyed and interviewed 12 internal scientists to assess the topics of FAIR data, data catalogue services, data analysis and data management from which we set out objectives of what knowledge we wanted to gain. This then allowed us to design a larger and more detailed questionnaire. This was piloted with our 52 ExPaNDS ambassadors, who had previously agreed to support our ExPaNDS activities. We received a healthy response and their feedback allowed us to shape the final survey.

This was subsequently launched to the large Diamond user database, containing over 14,000 scientific contacts. We collected and analysed their responses and our findings were presented at the ExPaNDS closing event.

In addition, we have created a generic blueprint questionnaire that any collaborating facility or those further afield could utilise, should they wish to run a similar exercise.

7.4. Deliverables and milestones of the WP for the period

#	Deliverable	Partner	Due date	Actual date
6.05	ExPaNDS representation material for relevant large international conferences	Diamond	31/08/2021	04/03/2022
6.06	ExPaNDS representation material for relevant large international conferences	Diamond	30/04/2022	23/06/2022
6.07	ExPaNDS representation material for relevant large international conferences	Diamond	31/12/2022	30/12/2022
6.11	ExPaNDS presentation for conferences	Diamond	30/06/2021	09/08/2021
6.12	ExPaNDS presentation for conferences	Diamond	28/02/2022	04/03/2022
6.13	ExPaNDS presentation for conferences	Diamond	30/06/2022	15/07/2022

⁷⁸ <https://doi.org/10.5281/zenodo.5973069>

⁷⁹ <https://doi.org/10.5281/zenodo.5639003>



6.2	Assistance to the development of the taxonomy strategy for the relevant data catalogues	Diamond	31/08/2021	01/09/2021
6.21	Promotional video	Diamond	31/05/2021	02/06/2021
6.22	Librarian symposium report	Diamond	31/10/2021	04/02/2022
6.23	User interviews	Diamond	31/12/2021	04/02/2022
6.24	Senior level engagement	Diamond	30/06/2022	01/08/2022
6.25	PaN communities use cases	Diamond	31/12/2022	30/12/2022

#	Milestone	Partner	Due date	Actual date
22	AAAS & ESOF attendance	Diamond	31/08/2021	28/02/2023
23	AAAS & ESOF attendance	Diamond	30/04/2022	28/02/2023
24	AAAS & ESOF attendance	Diamond	31/12/2022	28/02/2023
26	Ambassador engagement from all facilities user community	Diamond	31/07/2022	01/09/2022
29	Taxonomy Strategy	Diamond	31/08/2021	01/09/2021



Appendix: List of all ExPaNDS-related meetings

All events and workshops (co-)organised by ExPaNDS and all presentations given by ExPaNDS members during the **second period** of the project are listed in the following table:

Date	Topic	Title
2021/03/23	PaN training	Train the Trainers - PaNOSC/ExPaNDS workshop
2021/04/06	EOSC	ExPaNDS workshop on EOSC and Delivering data services to EOSC
2021/04/22	FAIR	Sharing FAIR data on COVID research at PaN RIs at RDA17 ⁸⁰
2021/05/18	Data catalogues	PaNOSC WP3 Catalogue Integration Best Practices Meeting
2021/06/08	Mid-term review	ExPaNDS mid-term review (GA 857641)
2021/06/10	FAIR	Introduction to FAIR and research data management to support FAIR data ⁸¹
2021/06/21	FAIR	Reusing the FAIR experiment
2021/06/18	Data catalogues	ExPaNDS WP3 status update
2021/06/18	PIDs	Openly reproducible PIDs as a factor of FAIRness in data sharing practices
2021/06/18	Engagement	ExPaNDS and the neutron and synchrotron light source science community
2021/09/30	FAIR	ExPaNDS Librarian Symposium Report and presentations ^{82, 83}
2021/10/19	Data analysis	EGI conference: Bridging from theory to practice with EOSC-Synergy
2021/10/22	FAIR	Persistent Identifiers (PIDs) for Facilities Research Workshop - Presentations
2021/10/26	PaN EOSC symposium	Booklet of presentations from the PaN EOSC Symposium 2021
2021/11/15	FAIR	Assessing the FAIRness of a prototypical PaN instrument at BESSY II Talk
2021/11/16	Data catalogues	Data Hub Use Cases for ExPaNDS and PaNOSC

⁸⁰ [IG Research Data Needs of the Photon and Neutron Science Community Sharing FAIR Data on COVID Resear - YouTube](#)

⁸¹ [UKCH Data Management Workshop - Abigail McBirnie - YouTube](#)

⁸² <https://vimeo.com/675857001>

⁸³ <https://vimeo.com/675860373>



Date	Topic	Title
2021/11/29	Management	PaNOSC alignment workshop
2021/12/07	Data catalogues	ExPaNDS session for gap and issue assessment
2022/01/12	FAIR	Invigorating the FAIR principles for surface science
2022/01/20	FAIR	ExPaNDS: Bringing FAIR photon and neutron data to the European Open Science Cloud
2022/01/24	Data catalogues	DAPHNE4NFDI: Science driven data management solutions for the user community
2022/02/03	FAIR	The role of Helmholtz Partners in FAIR related Photon and Neutron EU Projects - see pages 138-158
2022/02/07	FAIR	Towards FAIR research data management at CECAM workshop
2022/02/17	PaNET ontology	PaNET ontology workshop
2022/03/02	FAIR	Metadata framework validation presentations ⁸⁴
2022/04/04	Data catalogues	Workshop on metadata catalogues ⁸⁵
2022/04/05	EOSC	Science des Photons et des Neutrons et EOSC : PaNOSC et ExPaNDS
2022/04/08	EOSC	EOSC, un atout pour la recherche
2022/05/03	UmbrellaID	UmbrellaID Workshop: Keycloak ^{86, 87}
2022/05/09	PaN training	PaN training platform presentation at the LEAPS and LENS network meeting
2022/05/11	Open Science	SOLEIL et Science Ouverte
2022/06/14	All-hands	PaNOSC & ExPaNDS Prague F2F meeting
2022/07/06	FAIR	FAIR self-assessment workshop #1
2022/07/18	Engagement	Public Awareness of Research Infrastructures (PARI 2022)
2022/07/13	Open data	Open Data for healthier societies: a virtuous cycle? ⁸⁸
2022/08/26	FAIR	ExPaNDS - Towards FAIR and open photon and neutron data at IXS 2022

⁸⁴ <https://vimeo.com/685211186>

⁸⁵ <https://vimeo.com/695852712>

⁸⁶ <https://umbrellaid.org/zoom/2022-05-03-UmbrellaID-Workshop-Session-1.mp4>

⁸⁷ <https://umbrellaid.org/zoom/2022-05-03-UmbrellaID-Workshop-Session-2.mp4>

⁸⁸ <https://youtu.be/O2dN-3hfYus>



Date	Topic	Title
2022/08/29	FAIR	What the ExPaNDS project provides as tools promoting FAIR open data sources and data analysis services for Photon and Neutron facilities at 18th ESUO GA
2022/09/08	Engagement	ALBA user meeting and poster
2022/09/16	Data analysis	VISA workshop ⁸⁹
2022/09/18	Open data	Challenges towards Open Data in science
2022/09/22	EOSC	PaN community beyond ExPaNDS and PaNOSC
2022/09/26	FAIR	FAIR self-assessment workshop #2
2022/10/10	FAIR	FAIR data in the Photon and Neutron community at IBERGRID
2022/10/17	Open Science	Open science and data at SOLEIL at GANIL
2022/10/18	PaN EOSC symposium	3rd PaN EOSC symposium: sustainable data from PaN facilities
2022/10/19	Engagement	International Conference on Research Infrastructures (ICRI 2022)
2022/11/03	Data catalogues	Joint ExPaNDS and PaNOSC WP3 meeting ⁹⁰
2022/11/04	Data management	HZDR data management day
2022/11/15	PaN training	The Training Catalogue for Photon and Neutron Data Services ⁹¹
2022/12/05	PaN training	Presentation to the DAPHNE4NFDI project consortium
2023/01/20	Data analysis	Participation to the round table participation on data analysis management
2023/01/23	All-hands	ExPaNDS closing event
2023/02/20	Engagement	Topic-based webinar: Life Sciences
2023/02/21	Engagement	Topic-based webinar: Cultural Heritage
2023/02/22	Engagement	Topic-based webinar: Tomography & Imaging
2023/02/23	Data analysis	Technical meeting on VISA and hands-on session
2023/02/23	Engagement	Topic-based webinar: Industry as data users

⁸⁹ <https://vimeo.com/753854451>

⁹⁰ <https://vimeo.com/775112529>

⁹¹ <https://youtu.be/KeqMNC3fwQc?list=PLbISfqJh3TsvLr2xF7xaLiqFVYW1nvQz4>



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

Date	Topic	Title
2023/02/24	Engagement	Topic-based webinar: General update
2023/02/28	Engagement	Topic-based webinar: Grant update to the global Lightsources.org community

