PaNOSC

Photon and Neutron Open Science Cloud

H2020-INFRAEOSC-04-2018

Grant Agreement Number: 823852

Deliverable: D1.7 Mid-year summary 3

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Project Deliverable Information Sheet

Project Reference No.	823852
Project acronym:	PaNOSC
Project full name:	Photon and Neutron Open Science Cloud
H2020 Call:	INFRAEOSC-04-2018
Project Coordinator	Andy Götz (andy.gotz@esrf.fr)
Coordinating Organization:	ESRF
Project Website:	www.panosc.eu
Deliverable No:	D1.7
Deliverable Type:	Report
Dissemination Level	Public
Contractual Delivery Date:	31/05/2021
Actual Delivery Date:	04/06/2021
EC project Officer:	Simona Misiti

Document Control Sheet

Document	Title: Mid-year summary 3
	Version: 1
	Available at: https://github.com/panosc-eu/panosc
	Files: 1
Authorship	Written by: Jordi Bodera
	Contributors: Andy Götz, Rudolf Dimper, Tobias Richter, Fabio Dall'Antonia, Carsten Fortmann-Grote, Jean-François Perrin, Ornela De Giacomo, Thomas Holm Rod and Nicoletta Carboni
	Reviewed by: Andy Götz
	Approved: Rudolf Dimper

List of participants

Participant No.	Participant organisation name	Country
1	European Synchrotron Radiation Facility (ESRF)	France
2	Institut Laue-Langevin (ILL)	France
3	European XFEL (XFEL.EU)	Germany
4	The European Spallation Source (ESS)	Sweden
5	Extreme Light Infrastructure Delivery Consortium (ELI-DC)	Belgium
6	Central European Research Infrastructure Consortium (CERIC-ERIC)	Italy
7	EGI Foundation (EGI.eu)	The Netherlands

Table of Content

Project Deliverable Information Sheet 2
Document Control Sheet 2
List of participants 2
Table of Content 3
Introduction 4
Executive Summary 4
Summary of Executive Board meetings 4
Summary of Project Management Committee meetings 6
Progress of work packages 8
Work Package 1: Management 8
Work Package 2: Data Policy and Stewardship 9
Work Package 3: Data Catalog Services 10
Work Package 4: Data Analysis Services 11
Work Package 5: Virtual Neutron and x-raY Laboratory (VINYL) 11
Work Package 6: EOSC integration 13
Work Package 7: Sustainability 13
Work Package 8: Staff and User Training 14
Work Package 9: Outreach/Communication and Dissemination/Impact 15
Key Performance Indicators 17
Comparison between actual and forecasted project status 21
Risk Register 23
Next steps 26

Introduction

This document summarises the progress achieved in the project since the last management report. As such, this document and the one that will follow on month 42 will complement the four instances of annual workshop reports (months 12, 24, 36 and 48) to provide a regular update on the project, its management and comparing its current status with what was forecasted.

Executive Summary

Overall, the project is advancing according to plan. A virtual joint annual meeting attended by partners, ExPaNDS' project members, cluster representatives and observers took place in November 2020. A number of other meetings have been organised PaNOSC or been taken part in. The deliverables proposed in the grant agreement are being delivered mostly on time and are of high quality. The partners are engaged in the project, participate in regular meetings, recruitment has been successful. Our main repository documents our continuous activity: https://github.com/panosc-eu/panosc

Summary of Executive Board meetings

The fifth EB meeting took place on 9 November 2020 via zoom. All PaNOSC partners were represented by their EB representative. The meeting was chaired by T. Tschentscher (Eu.XFEL, chair of the PaNOSC EB for the period Nov 2019 - Nov 2020). The agenda of the meeting was as follows:

1. Review of progress and state of PaNOSC after 2 years
2. Use Cases and engaging with Users
3. Action plans per partner to address issues raised after 18M review
4. EOSC association
5. Nomination of new EB Chair
6. Topics to discuss with ExPaNDS
7. AOB

1. Review of progress and state of PaNOSC after 2 years

The Project Coordinator presented the status of PaNOSC WP by WP. He highlighted some specific points which impact or may impact the smooth progress of the WPs and in particular issues with the AAI integration, the fact that there was still only a small amount of open data, the need for more developers for WP4 on data analysis services, and at times a lack of engagement from partners to bring all actions forward with the highest possible degree of visibility.

He continued saying that overall, the project is advancing well but he thinks that the interaction with scientists at the partner sites still needs to be intensified.

EGI offered its help for the portal integration and ELI expressed the need for working with ExPaNDS on the long-term sustainability of the PaNOSC outputs.

2. Use Cases and engaging with Users

The Project Coordinator continued explaining that having relevant use cases will make PaNOSC, EOSC, and FAIR data management known to our user communities and considerably increase the value of the services developed in the project. The Executive Board recommended collecting as many scientific use cases as possible, also involving ExPaNDS, and then to review and prioritise them on relevance, impact, and the capacity to deliver.

3. Action plans per partner to address issues raised after 18M review

The Project Manager went through the points, which were discussed during a dedicated Project Management Meeting on 14 October 2020. This concerned human resource management at ILL and Eu.XFEL, a possible project extension, and the need to increase efforts for the development of the Portal and micro-services to connect to the infrastructure. With respect to requesting an extension of the project it was decided to review the situation at the end of year-3 and take a decision in function of the advancement of the project.

4. EOSC association

The EB discussed the importance of joining the EOSC-Association as members for having voting rights and for actively shaping the future direction of the EOSC implementation. Since it would be likely that universities would be numerous to join, the presence of the RIs to present their viewpoint would be even more important. The fact that the EIROforum has obtained a mandated voting right was welcomed.

5. Nomination of new EB Chair

Mark Johnson from the ILL was unanimously appointed as the new EB chair of PaNOSC for one year (Nov 2020 - Nov 2021).

6. Topics to discuss with ExPaNDS

The Project Coordinator introduced the item by stating that the work done by PaNOSC and ExPaNDS needs to continue beyond the project duration. He mentioned that centralising Open Data from the PaN RIs might be a way to ensure visibility and uptake by the PaN user community. Efforts should also be undertaken to investigate the possibility of a follow-up project. It was concluded that sustainability could be a topic to be discussed during a joint PaNOSC-ExPaNDS EB meeting in 2021.

7. AOB

No AOB items were discussed.

Summary of Project Management Committee meetings

The Annual Meeting[1] of the PaNOSC project took place online 9 to 11 November 2020 together with the European Open Science Cloud Photon and Neutron Data Service (ExPaNDS) project. A report of the Annual Meeting in 2020 can be found in the 2020 Annual Report deliverable D1.6. This section will cover the Project Management Committee (PMC) meetings of PaNOSC from then until May 2021.

Meeting 18/11/2020

The first PMC meeting after the annual meeting took place on 18 November during which the attendees exchanged their opinion about how the meeting developed and thanked the organisers.
WP leaders updated on progress made in their work packages, focusing on deliverables and milestones due soon.

The summary notes of this meeting[2] (as for all the others) are available in GitHub[3].

Meeting 02/12/2020

The meeting started with a summary of what has been delivered since the last meeting (D5.2, MS5.2, D1.6 and MS1.3) and on what was at the time still pending.

The new Open Research Europe publishing platform was discussed and the aim to review PaNOSC's KPIs and harmonise them with ExPaNDS was endorsed.

Meeting 16/12/2020

The meeting started with an agreement to inform the Project Officer about the delays affecting WP2 and WP6's deliverables and milestones. WP9 also informed of a new updated version of D9.1 - PaNOSC's Communication and Dissemination Plan. WP leader updates and progress on common KPIs with ExPaNDS and reviewing the risk register completed this meeting.

Meeting 06/01/2021

The first meeting of 2021 for PaNOSC focused on WP updates and news from the EOSC Association General Assembly (held 1712/2020) and the EOSC Future project.

Meeting 20/01/2021

During this meeting, the response of the Project Officer to our email (see meeting 16/12/2020) was discussed, in particular the fact that a new Project Officer will be appointed in April and that delayed deliverables/milestones should be a priority.

[1] https://indico.eli-beams.eu/event/376/

[2] https://github.com/panosc-eu/panosc/blob/master/Work%20Packages/WP1%20Management/Meetings/Project%20Management%20Committee/2020-11-18-PMC/2020-11-18-PMC_Summary.md

[3] https://github.com/panosc-eu/panosc/tree/master/Work%20Packages/WP1%20Management/Meetings/Project%20Management%20Committee

The funds available for EGI's contribution in PaNOSC and the new feature of PaNOSC's website to enable the submission of use cases were also discussed before a quick review of open issues and the WP progress.

Meeting 03/02/2021

The meeting started with the deliverable and milestones from WP2 and WP6 that were still delayed and the importance of submitting these as soon as possible.

The attendees also talked about the budget reserved for travel that is not being used since the start of the COVID-19 pandemic and work packages updates, including upcoming WP6-related AAI and WP8-related e-learning platform content creation trainings.

Meeting 17/02/2021

During the regular Project Management Committee meeting we discussed the outstanding milestones to achieve and the current issues preventing their achievement. This was followed by work package leader updates, an update regarding the EOSC Future proposal and the status of use case submission.

Meeting 03/03/2021

The meeting started again with an update of the progress towards achieving MS8.4 and MS6.2 and the good news that two services are now registered in the EOSC:

- PaN e-learning platform - https://marketplace.eosc-portal.eu/services/pan-learning-org-is-a-e-learning-platform-specifically-designed-to-provide-training-for-users-of-photon-and-neutron-sources
- PaN software catalogue - https://marketplace.eosc-portal.eu/services/panosc-software-catalogue

Conversations about repeating the common annual meeting in 2021 with ExPaNDS were reported and a dashboard showing the status of use cases was also shown helping to point out which partners need to increase their efforts to submit use cases. An update from WP leaders concluded the meeting.

Meeting 17/03/2020

The meeting started with a WP6 (EOSC Integration) presentation about the open questions that remain for service definitions related to PaN EOSC, some of which overlapped with WP7 (Sustainability).

Other topics covered were:

- It was agreed that PaNOSC will present the Software Catalog in an ExPaNDS' workshop, the 2021 annual meeting was discussed
- The planning of the first meeting to organise the 2021 Annual Meeting
- Use cases status and importance
- WP leader updates

Meeting 31/03/2020

During this meeting it was agreed to perform a WP per WP review, focusing on what PaNOSC is going to deliver through that WP, what is outstanding, challenges and remaining issues. An internal financial review for PaNOSC was also agreed, in order to understand what spending has taken place and where PaNOSC is over or under spending. It was also agreed that WPs will be producing dashboards to help assess the status and major items pending.

The meeting ended with a short update on the EOSC-Future (kick-off planned for 1st April 2021) and WP leader updates.

Meeting 14/04/2020

The meeting started by reporting that at the time it was still not known who the new Project Officer for PaNOSC would be. That was followed by a discussion on how PaNOSC is going to collaborate with EOSC-Future and a presentation on the Chatbot available in the PaNOSC website.

Before the usual WP leaders updates, reminders for the required deliverables and use cases took place.

Meeting 28/04/2020

During this meeting it was reported that a new Project Officer has been appointed which was followed by a short update on the EOSC-Future and the WP4 status internal review.

As part of the preparations for the WP4 review, the dashboard has been updated and challenges identified, in particular the fact that deploying the PaN Portal in production is not an easy task (each RI will need to do some work to interface it with their infrastructure) and it should be completed within 12 months.

The lengthy WP4 review did not leave space for the remaining items in the agenda and it was agreed to increase the frequency of meetings to weekly in order to accommodate the usual reporting and WP status reviews.

Meetings in May 2020

Meetings that took place during May are not covered in detail due to their proximity to the deadline for delivery of this document, however several meetings were planned to

- Report progress and deal with issues
- WP status internal reviews (WP4 and WP3)

Progress of work packages

Work Package 1: Management

Following the 2020 Annual Meeting, the deliverable D1.6 Annual Report was written and submitted.

The next deliverable for WP1 is this one (D1.7), however in the meantime WP1 has carried out the following actions:

- Coordination between partners and organisation of regular PMC meetings
- Coordination with ExPaNDS
- Representing PaNOSC at different meetings/conferences
- Communication with the PO
- Work with WP leaders to update the risk register
- Work with ExPaNDS on defining common Key Performance Indicators (KPIs)
- Work with WP leaders to obtain values for the agreed KPIs
- Start to plan and organise the next Annual Meeting
- Continue the project monitoring

  - Project Management Committee meetings
  - Maintenance of GitHub and documents stored
  - Starting a new Internal financial report
  - Set-up of WP dashboards
- Support for mailing lists

Overall, the Management work package is working well, however constant monitoring of the activities of other work packages and partners is required to ensure that PaNOSC remains on track and continues delivering as per the Grant Agreement.

Compared with a year ago, an increased collaboration with ExPaNDS is taking place and new monitoring initiatives are being taken (dashboards, effort on risk management, KPIs, etc.)

Work Package 2: Data Policy and Stewardship

The Work Package 2 on Data Policy and Stewardship had three main activities since the last annual meeting in November 2020. These are

1. the development of a Data Management Plan template generator,
2. the updating of data policies to be compatible with the PaNOSC Data Policy at all partners, and
3. the writing of the Data Policy Implementation Guidelines.

The Data Management Plan generator is part of "Task 2.5 Implement DMP template". The work involved reviewing the needs of all partners wrt DMPs. This was captured online in a google document[4]. The document is a complete summary of the needs of the partners and would merit to be converted into an internal publication for future reference.

The second main task was modifying the existing Data Policies in place or new ones being proposed to be compatible with the PaNOSC Data Policy. This involved consultations with the internal scientific committees to get their input and approval, then submitting the new data policy to top management for official approval. This task was carried out at each site and is well advanced at ESRF, ESS and Eu.XFEL. In the case of ILL the process has been identified and will be soon start. CERIC-ERIC and ELI have written new data policies which have been approved by their top management. The entire process has been documented in a detailed manner for MS 2.2. The milestone document is available from the PaNOSC github repository[5]. This document is of general interest to others who need to either adopt a data policy or modify an existing one.

The third task is the documenting of guidelines for implementing data policies based on the experience at the different sites. A draft of the document has been produced and the final version submitted. This document was delayed by 6 months. The main cause for the delay was the work overload of the WP leader. The quality of the deliverable was not affected however. The deliverable has been reviewed by an expert from ExPaNDS and found to be extremely useful for all PaN RIs who are implementing a data policy.

A number of activities which are related to WP2 have been happening in parallel. These include the submission

[4] https://docs.google.com/document/d/1LVJ0O95ZtqByk84TcyhdpMJYSLlFADNe8eIvNxUnJuU/edit?usp=sharing
[5] https://github.com/panosc-eu/panosc/blob/master/Submitted%20Milestones/07%20Adoption%20of%20PaNOSC%20DP%20Framework/PaNOSC-MS2.2%20-%20Adoption%20of%20PaNOSC%20DP%20framework.pdf

of the ESRF data repository for certification by CoreTrustSeal. The experience with this self-certification process is of practical interest to all partners. PaNOSC has been contacted by other projects like INSTRUCT for exchanging on best practices concerning data policies. PaNOSC together with ExPaNDS organised a Birds-of-a-Feather session at the online RDA 17 plenary even entitled "Sharing FAIR Data on COVID research at Photon and Neutron Facilities". The session was well attended and the quality of talks was very high and stimulating. Talks were mainly on the sharing of COVID-19 data from PaN RIs with speakers from European and the USA PaN RIs well represented as well as the EU COVID-19 data portal by the EOSC-Life cluster project.

Work Package 3: Data Catalog Services

In this reporting period partners have made good progress to establish harvesting endpoints to offer data to third party repositories via OAI-PMH. B2Find now propagates data automatically to OpenAIRE, which reduces the number of parties we need to register with. What effect that has on metadata mapping and potential data duplication (since some facilities are already harvested by OpenAIRE) is to be explored soon. Also establishing end points for the common search API is progressing well. A number of facilities have started offering data live from their data catalogues in compliance with the API from deliverable D3.1. The remainder of the partners have submitted representative metadata sets that can be served from the static API reference implementation. This was a useful preparation in the development of the federated search demonstrator, which has been submitted as deliverable D3.2. The federated search demonstrator provides a single interactive search interface that internally fans out queries to any number of configured search API endpoints from partner facilities and aggregates the results. The demonstrator implementation has a few shortcomings for practical use and did also highlight some potential issues with the search API. This is according to the development plan. It provides developers with early feedback cycles to direct efforts towards the most relevant areas and allows the project to iteratively improving solutions multiple times during its duration. Partners are discussing improvements to the baseline version of the common search API, exploring ideas to make the search aggregation more useful to the user by scoring and sorting results. In addition, having a running federated search service strengthened the integration with WP4. Collaborative work on the data portal is being intensified and a common task force with WP4 has been established.

As mentioned, more and more metadata have been collected from partners. As a result D3.2 contained a short and useful baseline list of search parameters that are common between facilities and serve the purpose of filtering query results in a meaningful way. The same has happened for roles of personnel connected to datasets. For specifying the experimental technique there is a discussion ongoing with WP3 of ExPaNDS that is expected to conclude with a practically useful result soon. This leaves the ontology work in good shape for where the project should be at this point. To be usable in the federated search the mappings between local and common names needs to be implemented and curated for each data catalogue instance.

The task to connect data sources to data catalogues held a best practices meeting on May 18th to 20th (https://indico.eli-laser.eu/event/3/). The topics included data ingestion, interactions with user office software, log books, file formats and ontologies. A summary of the workshop will be produced as a milestone document.

Work Package 4: Data Analysis Services

Since the last annual meeting, the decision was made to focus on the migration of the existing VISA (*Virtual Infrastructure for Scientific Analysis*) service from the ILL to serve as PaN portal production implementation. This plan was further concretised in terms of a road-map by the ILL developers, which lately included the decision to restrict built-in cloud service support to OpenStack compute infrastructures. With this plan, resources can be focused on providing the core VISA code-base as open-source, with the effort to adapt and add microservices being reduced to a feasible amount, and a much better perspective on maintenance and documentation for sustainable usage. It is expected that this adaptation of the work-plan will allow deployment of the VISA/PaN-portal at the partner sites beginning in summer. The PaNOSC partners have already started preparing their cloud infrastructure for the connection to the VISA service: either they already provide an OpenStack architecture in some form, or are ready to set such up soon.

A logging service has been developed by the CERIC partners to be integrated to the back-end architecture, and the search-API-related part of the web-UI could be advanced by ELI partners. This lately included a kick-off meeting on intensified collaboration with the search-API developers in WP3 for better alignment of search end-points and instrument-specific data models. At all levels of the portal back-end architecture, progress was made to support OpenID connect authentication, so to allow for flexible EOSC UmbrellaID login to the portal, granting individual access levels to data and compute resources depending on whether a facility user connects to own data (in case embargoed), or any community user connects to open data.

Contributing to the supported data analysis software stack, ESRF have further developed the h5web viewer software and provide its functionality as a plugin to JupyterLab. Moreover, adding to their set of tools they developed jupyterhub_moss, a Python package providing a JupyterHub Slurm spawner. Eu.XFEL have looked into existing ways and means to further enhance portability of Jupyter notebook repositories and their environments, on top of Binder ("postBuild"), as well as providing data retrieval methods from Binder/Jupyter ("start").

PaNOSC use cases have been addressed by a newly created task force, which resulted in the compilation and publication of 18 data analysis use cases so far, with contributions from most of the PaNOSC partners and a coverage of various experimental techniques (serial femtosecond crystallography, neutron scattering, X-ray optics and photon diagnostics just to name some) as well as simulation data.

Work Package 5: Virtual Neutron and x-raY Laboratory (VINYL)

We report here on the progress in WP5 since the submission of the latest deliverable (D5.2) and milestone (M5.2) which roughly coincided with the last PaNOSC annual meeting.

Project planning

Since the next deliverable and milestone are due in May 2022, we defined the following Internal Milestones (IMS) to monitor our progress:

- **IMS5.1**: Protocol for comparison of raw simulated to raw experimental data
- **IMS5.2**: Instrument database
- **IMS5.3**: Simulation database

- **IMS5.4**: Interfacing the instrument simulation database
- **IMS5.5**: Target database (NOMAD) interfaces

These IMSs reflect the goals set forth in Tasks 5.1, 5.3, and 5.4

Work on IMS5.1 is currently ongoing. A document outlining our recommendations for comparing raw experimental to raw simulated data is edited by the workpackage partners. In parallel, work on IMS5.3 and IMS5.4 has started with major activity from CERIC-ERIC, where a browsable collection of Oasys workspaces is created. These can serve users of our simulation services as a starting point to build their own simulation workflows in Oasys. Similarly, collections of Jupyter notebooks for SIMEX and McStasScript will be assembled. The workspaces and notebooks define beamline and instrument settings according to standard configurations in the respective RIs and enable users to simulate their experiments in these configurations.

We will now summarize the work that has been done with respect to the three major simulation platforms.

SIMEX

To improve the portability of SIMEX to different HPC systems (besides the DESY/Eu.XFEL system for which it was mainly developed), the build system and installation documentation were refactored. Furthermore, a new data API was developed to smoothen the user interface to the multitude of simulation data file formats. The so-far default radiation-matter interaction code XMDYN will be replaced by the Gromacs code which, in contrast to XMDYN, is freely available as an opensource code. A further change is the adoption of the extra-geom package to describe X-ray detector geometries in SIMEX, which makes the code compatible with the data analysis packages developed at Eu.XFEL and in PaNOSC WP4.

SIMEX was used for work described in a manuscript recently submitted for publication that analyses the impact of the hydration layer on X-ray scattering and reconstructability. Furthermore, two use cases have been submitted to the PaNOSC use cases database (use cases 9 and 14) which employ SIMEX to execute the proposed simulations.

McStas/McStasScript

During the last months, the code base of McStas has been reviewed extensively and review comments addressed and implemented. The interest in McStasScript has increased and several tutorials were given at international conferences and by invitation of collaborating laboratories. Within WP5, a use case that combines DFT target simulations with neutron scattering simulations driven by McStasScript was developed and submitted as use case 18 to the PaNOSC use cases database. Our partners at ILL use McStas to generate training and test data for a ML/AI based framework to predict neutron intensity distribution at the target position given the distribution at the beam origin. McStasScript is also used in online training material produced in WP8.

Oasys/X-ray optics simulations

A high level API based on the openpmd-api and reusing the metadata standards developed in WP5 (see deliverable D5.1) is now available for use in Oasys. The WISER code for coherent wavefront propagation was further developed and successfully applied to various simulation cases for beamlines at FERMI-ELETTRA.

Work Package 6: EOSC integration

Since the last annual meeting, the work on UmbrellalD, the PaN community AAI, has made important progress. All the community services that were performing users' authentication through UmbrellaID have been migrated to the new infrastructure powered by GÉANT: eduTEAMS. This achievement has authorised the deployment of new functionalities like the registration of user metadata that were approved during the last UmbrellaID management meeting. Services could now benefit from up to date and verified metadata like user's name and email address. For starting, this list of metadata has been defined as the minimum necessary for joining EOSC, but will be extended to meet new service needs once approved during the next UmbrellaID management meeting. New services, like the software catalogue, have joined UmbrellaID and were able to take benefit of such functionalities. Typically, such services do not have to manage local accounts anymore, they can simply rely on the UmbrellaID information.

To ensure a large adoption of UmbrellaID, a 1-day training (https://indico.psi.ch/event/10773/) has been organised by the WP6 team with the support of PSI and GÉANT experts. More than 40 IT specialists from the community were able to take advantage of this workshop and participated in the hands-on sessions.

The software catalogue has been refurbished and rolled out, new functionalities that were identified following the 2020 user community survey have been implemented and presented to the community. Apart from the integration of UmbrellaID, The most visible add-on is probably the integration of REST web services that allow querying the catalogue. This will simplify its integration with community services like data portals and other EOSC platforms. This service alongside the PaN training platform is now offered through the EOSC portal. The upgrade of the software catalogue and its general availability as a public service is also a deliverable (D6.4) of WP6. D6.4 was delayed by 6 months (due month 24) mainly due to the COVID pandemic. This work had to be handled by software developers who at the same time had to focus their activity on providing remote access solutions for their RI users. This period has been beneficial for activities such as the development of the remote analysis portal in WP4, but has slightly delayed less urgent work such as the software catalogue upgrade.

Following the first registration of PaNOSC services in the EOSC portal, we identified that we needed to provide more guidance to the community for delivering EOSC services. WP6 has started to draft service definitions (AAI, HelpDesk, monitoring, …). These definitions will be presented to the PMC and the ExPaNDs project for adoption. They should contribute to avoid delays for the registration of new services.

Regarding data transfer, out of the three use cases identified, two have adopted a solid solution (Globus for users' driven data transfer and Rclone for active across sites), the 3rd use case regarding data provisioning for EOSC services is still under active investigation.

Work Package 7: Sustainability

Since the last annual meeting, WP7 has progressed in all the tasks. In particular:

Most of the facilities completed the cost collection (task 7.2), however, due to some organisational changes, one of the facilities is missing, although the task is advanced and hopefully will be delivered on time to be included in the deliverable. Although quite some discussion took place to align the facilities and guarantee the uniformity of the costs reported, this was not fully achieved. When most of the facilities submitted their cost sheets, the costs were analysed, e.g. to identify the main cost drivers, and it became evident that this task was not trivial.

From this moment, the WP engaged in a discussion leading to a better definition of the cost lines. The comparison of approaches focused on the items included and the main reasons at the origin of cost differences that could not be simply explained by natural factors such as the volume of data, the number of users, number of instruments, or others. It emerged that the technical choices defined in the IT strategy of each partner site heavily influenced the costs and is one of the main cost drivers. The work then focused on defining better these aspects, by providing a meaningful representation of the cost analysis in the deliverable.

The task on business models for the PaN EOSC (task 7.3) also progressed, achieving a more straightforward definition of the value proposition and towards completing the canvas elements. This task relied on precious information obtained from the other WPs. Early in 2021, every WP leader was interviewed and asked about the outputs of their WPs, how much effort and how many resources would it require to maintain these developments after the end of the project and what model they could imagine for the governance of these outputs in the future. These interviews led to a better-defined value proposition and the identification of complementary costs to the ones collected by the partners in task 7.2. While most partners included in their cost sheets the additional costs for the future maintenance of the PaNOSC developments, none of them considered the cost of federated or centralised services. However, this information is crucial for the sustainability plan.

This work also led to a better understanding of possible models that could be applied to achieve the sustainability of the PaN EOSC. WP7 involves the ExPaNDS partnership in these discussions since the outputs of these two projects are shared, and their future actions are somewhat entangled.

The WP considers all the aspects that should be covered, ranging from governance to legal and financial aspects. The first discussions have taken place within WP7, intending to identify feasible strategies that should be validated with facility managers. There is a first interaction planned during the General Assembly of ERF-AISBL to occur at the beginning of the summer. The feedback collected during this meeting will contribute in refining the business models that will feed the sustainability plan. A broader interaction is planned at a later stage with different stakeholders to ensure that the scenario presented in the sustainability plan fits the needs of the PaN RIs, the users, the funders, and other relevant stakeholders.

Work Package 8: Staff and User Training

At the beginning of January, a new two year post-doc was recruited at ESS to work full time for the PaNOSC project (WP8 and 5).

Task 8.2 Integrating Jupyter Technology has come to a provisional conclusion with the completion of milestone 8.4 "Jupyter integrated with e-learning platform". Two solutions have been implemented, each with their advantages and disadvantages. One solution enables a simple launch of JupyterHub from pan-learning.org. The ESS IKON python workshop (https://indico.esss.lu.se/event/2568/) was hosted using this solution. Likewise, a course in McStasScript (WP5) for staff of the Jülich Centre for Neutron Science at Forschungszentrum Jülich. These courses demonstrated the functionality of PaN-learning JupyterHub. However due to some of its limitations, an alternative method of integrating Jupyter into PaN-learning using Illumidesk was set up. Illumidesk is currently operational and allows teachers to upload notebooks themselves (something which in Jupyterhub has to be carried out by system administrators). Unfortunately, it has issues of its own and solutions to these are currently being investigated. Ultimately only one out of Jupyterhub and Illumidesk will be maintained. The IllumiDesk solution was demonstrated at the Train-the-Trainers workshop (see below).

Task 8.3 Integrate ESFRIs in the e-learning virtual facility. The option to use Jupyter from pan-learning.org now makes it possible to run simulation codes considered in WP5 (McStas and SIMEX) from Jupyter. A GUI has been developed in Jupyter that allows users to run McStas from this GUI rather than through the Python API. A GUI will also be developed for SIMEX. These developments allow us to remove the web simulator, implemented in the Django framework, from pan-learning and hence improve its maintainability. Possible courses using these implementations are currently being explored.

Task 8.4 Staff training in e-learning platform. Following an internal user-story workshop, much of the focus has been preparing pan-learning for content creators. A two part Train-the-Trainers workshop (https://indico.esss.lu.se/event/2499/) was held that familiarised participants with moodle, the learning management system used for pan-learning.org, and the pedagogy of interactive and online learning. In addition a new course on how to easily record presentations on a laptop has been added to help teachers make video lectures on their own without expensive software. Staff from the Department of Science Education and the Niels Bohr Institute, University of Copenhagen, facilitated the training. As a result of the workshop, the appearance of the PaN-learning moodle has been improved to make it easier to look at, use and navigate. The upload limit has been increased to allow teachers to upload large files to courses such as videos. In-line LateX notation is now possible and external links can be set to immediately open in a new tab. Moreover, we have started to see the development of the first training material by facility staff, e.g. for the SwednESS summer school.

Task 8.5 Staff training in data stewardship. This task is still in its infancy, and currently an overview of existing courses from other research infrastructures is being created.

Training material will be developed with the aim of providing data stewardship training.

Other activities

Looking to the future, a risk analysis workshop took place jointly with ExPaNDS WP5 to assess possible risk to the success of the project. These were given a score based on their likelihood and how detrimental they would be. Avoidance and mitigation steps have been suggested and are being considered.

An important example is the creation of a github organization: https://github.com/pan-training. Its purpose is to keep track of technical and development issues and assign tasks to people within WP8 and ExPaNDS WP5. Most issues are focussed on fixing the limitations of the Jupyterhub and/or Illumidesk, integrating federated AAI, uploading and improving existing course material onto the moodle and creating new courses which use software developed as part of WP5.

In collaboration with WP6, pan-learning.org is now registered as an EOSC service. Moreover, federated AAI in the form of keycloak and UmbrellaID is currently being integrated.

Work Package 9: Outreach/Communication and Dissemination/Impact

Following the recommendations included in the project review report, and exchanges with the partners and ExPaNDS, in January 2021, *D9.1 - Communication and Dissemination Plan* has been reviewed and resubmitted. The document better highlights joint activities with our sister project, ExPaNDS, as well as the actions towards a wider engagement of the community of users of photon and neutron sources in onboarding them to EOSC, also in collaboration with the other Science Clusters.

By the end of 2020, PaNOSC launched a call for use cases, for users of photon and neutron facilities in PaNOSC

and ExPaNDS to submit factual examples of the use of the services being developed in the project(s). The goal is to demonstrate, to other actual and potential users, the current practices in data stewardship, data transfer, (remote) data analysis, data and experiments' simulation, and of collecting inputs to further improve and develop PaN EOSC services by addressing the specific needs of the research community. The first collection of use cases has been published on the project's website, as well as presented at user meetings, and we aim to present a wider portfolio at the 2021 project's annual meeting.

WP9 has been continuously interacting with all WP leaders to promptly populate all the online communication channels (website and social media accounts) with relevant updates on the project's achievements. This includes events, milestones, deliverables and publications (both on Zenodo and on OA peer-reviewed journals), also by distributing the information to partners, other EOSC cluster projects, PaN European initiatives (LENS and LEAPS) and networks (e.g., lightsources.org, neutronsources.org).

In line with the plan to increase the visibility and enhance the adoption of FAIR data practices, some of the services developed in the project have been presented through video interviews and demos – published on the PaNOSC YouTube channel and the website, and distributed via PaNOSC and the partner's social media channels, newsletters and mailing lists:

- Interview with Mads Bertelsen on performing McStas simulations with McStasScript (+ DEMO): https://youtu.be/2o-9MySCdWs
- Interview with Juncheng E on the photon experiment simulation environment SIMEX (with DEMO): https://youtu.be/Ei5DtrC-4BI
- Interview w Ibrahim Dawod on the use of SimEx and Gromacs for bioimaging theoretical simulations (with DEMO): https://youtu.be/SE4nwchbBMg
- Interview with Erik Knudsen on the McStas python interface McStasScript for X-ray telescope simulations (with DEMO): https://youtu.be/WO0Tw8qxS-4

In addition, interviews to PaN users on their views on Open Data and Open Science were released:

- Interview with PaN user, Dr. Elisa Bergami, on the benefits of Open Science for the Environmental Sciences: https://youtu.be/HTVmX1qfbS8
- Interview with PaN user Petr Čermák on the benefits of Open Data and Open Science: https://youtu.be/aHUaE-Eqv88
- Radio interview (in Italian) with Ornela De Giacomo and Alessandra Gianoncelli on CERIC, PaNOSC & the EOSC: https://youtu.be/4nwJAGxhYqo

On social media, as of 6 May 2021, both demos and videos reached over 2550 views.

To increase the knowledge about the benefits of the EOSC and of the services and technologies developed in the project, PaNOSC (and ExPaNDS), and the collected use cases, were presented at three PaN user meetings (DESY / European XFEL; ESRF; SOLEIL), both in plenary and poster sessions.

Collaboration with other ESFRI Science Cluster projects has also been ongoing:

- PaNOSC participated in bi-weekly "Science Cluster Coordination Meetings" during which the Science Clusters exchanged on their experiences in the projects and which allowed to align the activities in the newly initiated EOSC-Future project. The Science Clusters and the European e-infrastructures issued a common statement for feedback on the SRIA (https://zenodo.org/record/4044010). Furthermore, the Science Clusters also discussed their interaction within the EOSC Association and with the Commission.

- PaNOSC organized and took part in joint events, such as the PaN ESCAPE Data Management Workshop, where PaNOSC / ExPaNDS and ESCAPE shared their experience in automated smart data transfer and data management systems. Moreover, at the EOSC-hub / FREYA / SSHOC event: "Realising the EOSC" in November 2020, PaNOSC contributed to the plenary session "Thematic Discovery Marketplaces for the EOSC", aimed at showcasing the ecosystem of thematic marketplaces for the EOSC and their role as aggregators.
- More recently, at the RDA 17th Plenary Meeting, the five ESFRI cluster projects, the RDA community, and EOSC representatives came together again, to discuss the past, present and future of their collaboration during the journey of integrating thematic services into EOSC, at the House of Commons debate on commonalities and collaboration for thematic services, training and governance towards the EOSC.
- PaNOSC participated in the EOSC-Life AGM at the end of March during which many stimulating discussions took place. The coordinators of PaNOSC and EOSC-Life are in contact to explore how both projects can further profit from their mutual work and share outcomes.
- In the frame of the RDA 17 Plenary, PaNOSC and ExPaNDS also jointly organized a session on "Sharing FAIR Data on COVID research at PaN Facilities", and PaNOSC was also invited to contribute to the ARCHIVER project's RDA Virtual Plenary event "Service R&D for Archiving & Preservation for Research Environments".

WP9 contributed to the promotion of such events, in coordination with ExPaNDS.

Finally, KPIs in WP9 have been continuously monitored, with the goal of further fine-tuning communication and dissemination actions of the project.

Key Performance Indicators

During the last 6 months, PaNOSC has reviewed a set of Key Performance Indicators (KPIs), which has resulted in the following updated KPIs:

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on	2nd Value obtained	2nd value obtained on	3rd value obtained	3rd value obtained on
1	Percentage of PMs spent	0%	58%	31/08/2019	73%	31/05/2020	82%	31/03/2021
1	Percentage of target expenditure	0%	53%	31/08/2019	70%	31/05/2020	77%	31/03/2021
1	Deliverables submitted late (as per EC Portal)*[1]	0	3	30/11/2019	3	02/06/2020	7	29/04/2021
1	% deliverables submitted late (as per EC Portal)*	0%	30.00%	30/11/2019	18.75%	02/06/2020	33.33%	29/04/2021
1	Milestones submitted late (as per EC Portal)	0	1	30/11/2019	3	02/06/2020	4	29/04/2021
1	% Milestones submitted late (as per EC Portal)	0%	10.00%	30/11/2019	23.08%	02/06/2020	25%	29/04/2021
2	number of instruments on which Data Policy is implemented	16	21	30/11/2019	KPI not collected		85	02/03/2021

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on	2nd Value obtained	2nd value obtained on	3rd value obtained	3rd value obtained on
2	number of techniques / instruments for which metadata are defined	9	11	30/11/2019	KPI not collected		52	02/03/2021
2	Number of instruments available	97	97	30/11/2019	KPI not collected		149	02/03/2021
2	Percentage of techniques for which metadata is defined	33%	33%	30/11/2019	KPI not collected		39%	02/03/2021
2	number of metadata parameters defined	500	1,070	30/11/2019	KPI not collected		1862	02/03/2021
2	number of raw data DOIs	500	589	30/11/2019	KPI not collected		5550	17/03/2021
2	number of user defined DOIs	5	12	30/11/2019	KPI not collected		91	17/03/2021
2	number of downloads	5	10	30/11/2019	KPI not collected		69962	17/03/2021
2	number of citations of DOIs	5	10	30/11/2019	KPI not collected		675	17/03/2021
2	number of datasets	5	10	30/11/2019	KPI not collected		803555	17/03/2021
2	volume of data archived	5	10	30/11/2019	KPI not collected		27.709 PB	17/03/2021
2	number of datasets cited in publications	0	2	30/11/2019	KPI not collected		675	17/03/2021
3	number of PANOSC facilities represented in the NeXus advisory committee	2	3	30/11/2019	3	10/07/2020	3	02/02/2021
3	number of PANOSC facilities offering public metadata via OAI-PMH	0	1	30/11/2019	2	10/07/2020	2	02/02/2021
3	number of facilities offering a PANOSC catalogue search endpoint	0	0	30/11/2019	0	10/07/2020	1	02/02/2021
3	number of datasets with FAIR data leaving embargo period in the coming year	0	2	30/11/2019	328354	10/07/2020	KPI not collected	
4	Percentage of facilities offering JupyterHub or remote desktop for analysis services	50%	67%	30/11/2019	100.00%	20/05/2020	100%	29/04/2021
4	Number of unique users making use of JupyterHub or remote desktop for analysis services at partner facilities	0	156	30/11/2019	269	20/05/2020	1312	31/05/2021
4	number of techniques available through remote services	2	8	30/11/2019	49	20/05/2020	72	31/05/2021

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on	2nd Value obtained	2nd value obtained on	3rd value obtained	3rd value obtained on
5	Number of contributors to ViNYL	0	6	18/11/2019	KPI not collected		9	20/01/2021
5	number of users of ViNYL services at ESRF	0	0	18/11/2019	KPI not collected		0	20/01/2021
5	number of users of ViNYL services at ILL	0	0	18/11/2019	KPI not collected		0	20/01/2021
5	number of users of ViNYL services at XFEL	0	0	18/11/2019	KPI not collected		3	20/01/2021
5	number of users of ViNYL services at ESS	0	0	18/11/2019	KPI not collected		2	20/01/2021
5	number of users of ViNYL services at ELI	0	0	18/11/2019	KPI not collected		0	20/01/2021
5	number of users of ViNYL services at CERIC	0	0	18/11/2019	KPI not collected		2	20/01/2021
5	Number of modules included in ViNYL service	0	0	18/11/2019	KPI not collected		3	20/01/2021
5	Number of partner infrastructures that have used ViNYL service	0	0	18/11/2019	KPI not collected		3	20/01/2021
5	Number of DOIs for simulated data (by counting datasets with "ViNYL" labels on open-access repositories like Zenodo).	0	1	18/11/2019	KPI not collected		5	20/01/2021
5	Number of openPMD standard domain extensions merged into mainline openPMD repository	0	0	18/11/2019	KPI not collected		1	20/01/2021
6	Service Providers connected to eduTeams Umbrella	0	0	06/11/2019	2	28/05/2020	7	15/12/2020
6	Services connected and accessible through eduTeams Umbrella	0	0	06/11/2019	1	28/05/2020	9	15/12/2020
6	Users that have used at least one service	0	0	06/11/2019	0	28/05/2020	KPI not collected	
6	Number of partners that have set up data transfer	0	0	06/11/2019	1	28/05/2020	1	15/12/2020
7	KPIs finally not agreed							

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on	2nd Value obtained	2nd value obtained on	3rd value obtained	3rd value obtained on
8	KPIs changed from previous iteration							
8	Instrument techniques considered in the e-learning platform	6	6	30/11/2019	KPI not collected		KPI not collected	
8	Number of participants in training workshops from ExPaNDS	0	N/A		N/A		KPI not collected	
8	Number of participants in training workshops from PaNOSC	0	N/A		N/A		KPI not collected	
8	Number of external participants in training workshops	0	N/A		N/A		KPI not collected	
8	TOTAL Number of participants of training workshops	0	0	30/11/2019	KPI not collected		KPI not collected	
8	Number of entries in training catalogue	0	N/A		N/A		KPI not collected	
8	Number of visitors to training catalogue	0	N/A		N/A		KPI not collected	
8	Number of unique visitors to training catalogue	0	N/A		N/A		KPI not collected	
8	Number of redirects from the portal to the training catalogue	0	N/A		N/A		KPI not collected	
8	Number of ExPaNDS registered users in pan-learning.org	?	N/A		N/A		KPI not collected	
8	Number of PaNOSC registered users in pan-learning.org	?	N/A		N/A		KPI not collected	
8	Number of external registered users in pan-learning.org	?	N/A		N/A		KPI not collected	
8	TOTAL Number of registered users in pan-learning.org	?	N/A		N/A		KPI not collected	
8	TOTAL Number of completed courses in pan-learning.org	0 (but some on e-neutrons)	0 (but some on e-neutrons)	30/11/2019	25 (although not specifically tied to T8.7)	7/1/2020	KPI not collected	
8	TOTAL Number of (institutional) content providers for pan-learning.org	?	N/A		N/A		KPI not collected	
8	TOTAL Number of courses / workshops making use of the platform for teaching.	?	N/A		N/A		KPI not collected	

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on	2nd Value obtained	2nd value obtained on	3rd value obtained	3rd value obtained on
9	Number of social media posts mentioning PaNOSC*[2]	0	116	10/11/2019	234	31/05/2020	293	6/11/2020
9	Number of followers on Twitter*[3]	0	256	10/11/2019	407	10/06/2020	480	6/11/2020
9	Number of user meetings in which PaNOSC is presented*[4]	0	0	10/11/2019	2	10/06/2020	4	6/11/2020
9	Number of invited talks as PaNOSC representatives on topics of relevance for the project and its partners*[5]	0	11	10/11/2019	28	10/06/2020	38	6/11/2020
9	Number of visitors of the PaNOSC website*[6]	0	1186	10/11/2019	1947	10/06/2020	2403	6/11/2020
9	Number of PaNOSC-related publications' downloads on Zenodo *[7]	n/a	n/a		860	10/06/2020	1454	6/11/2020

*[1] Please note that this includes deliverables that have been resubmitted so they are up to date with project progress (like D1.2 or D9.1)
*[2] 375 as of 06/05/2021; *[3] 580 as of 06/05/2021; *[4] 7 as of 06/05/2021; *[5] 43 as of 06/05/2021; *[6] 3079 as of 06/05/2021 and *[7] 2785

The actual values of these KPIs will be updated regularly (the aim is to do so at least once a year) in order to show the progress made by the project.

Discussions with ExPaNDS have also taken place and the KPIs listed above includes many KPIs shared with ExPaNDS.

As more services become available and the number of users grows, it is planned to make some of these KPIs public in the later stages of the project.

Comparison between actual and forecasted project status

The following list shows all deliverables and milestones with their current status:

Milestone or Deliverable Id	Name	Due Date	Status
D1.1	Project initiation documentation	31-Jan-2019	Submitted
D1.2	Data Management Plan	31-May-2019	Submitted
D1.3	Mid-year summary 1	31-May-2019	Submitted
D1.4	Report of annual workshop 1	30-Nov-2019	Submitted
D1.5	Mid-year summary 2	31-May-2020	Submitted

D1.6	Report of annual workshop 2	30-Nov-2020	Submitted
D1.7	Mid-year summary 3	31-May-2021	Submitted
D1.8	Report of annual workshop 3	30-Nov-2021	
D1.9	Mid-year summary 4	31-May-2022	
D1.10	Report of annual workshop 4	30-Nov-2022	
D2.1	PaNOSC data policy	31-May-2020	Submitted
D2.2	DMP Template	30-Nov-2021	
D2.3	Guidelines on implementing Data Policy	30-Nov-2020	Submitted
D2.4	Integration of the policy	30-Nov-2021	
D3.1	API definition	31-May-2020	Submitted
D3.2	Demonstrator implementation	31-Mar-2021	Submitted
D3.3	Catalog service	31-Mar-2022	
D3.4	Implementation Report	31-Jul-2022	
D3.5	NeXus Metadata Schema	31-May-2022	
D4.1	Report data analysis capture	30-Nov-2019	Submitted
D4.2	Prototype remote desktop and Jupyter service	31-May-2020	Submitted
D4.3	Remote desktop and Jupyter analysis service deployed at EOSC	31-May-2022	
D4.4	Publicly accessible Demonstrator	30-Nov-2022	
D5.1	Prototype simulation data formats	30-Nov-2019	Submitted
D5.2	Documented simulation APIs	30-Nov-2020	Submitted
D5.3	Documented simulation tasks executable	31-May-2022	
D5.4	Software tested and released including interactive simulation and analysis workflow	30-Nov-2022	
D6.1	Data-hub	31-May-2020	Submitted
D6.2	Compute cloud	30-Nov-2021	
D6.3	AAI	30-Nov-2021	
D6.4	Software catalogue	30-Nov-2020	Pending
D6.5	Report on EOSC integration	30-Nov-2022	
D6.6	Integration of RIs in EOSC	30-Nov-2022	
D7.1	Photon and Neutron EOSC Stakeholder Feedback	31-May-2020	Submitted
D7.2	Photon and Neutron EOSC metrics and costs model	30-Nov-2021	
D7.3	Photon and Neutron EOSC Business model reference document	31-May-2022	
D7.4	Photon and Neutron EOSC Sustainability plan	30-Nov-2022	
D8.1	Report on lessons learned and future prospects for adopting best practises data stewardship at the PaNOSC facilities	31-Jul-2021	
D8.2	Report on lessons learned for adopting the e-learning platform at the PaNOSC facilities, task 8.4	31-May-2022	
D8.3	Teaching material for users of PaNOSC services, FAIR principles, and the PaNOSC facilities accessible in the e-learning platform at pan-learning.org, task 8.5-7	31-May-2022	
D8.4	Closing report including report from summer school, task 8.8	30-Nov-2022	
D9.1	PaNOSC's Communication and Dissemination Plan	30-Jun-2019	Submitted
D9.2	PaNOSC's Website	31-May-2019	Submitted
D9.3	PaNOSC's repository for internal communications	28-Feb-2019	Submitted
D9.4	Dissemination and Outreach activities	30-Nov-2022	
D10.1	POPD - Requirement No. 1	31-Dec-2018	Submitted
MS1.1	Project Initiation Stage completed	31-Jan-19	Achieved
MS1.2	First Annual Report	30-Nov-19	Achieved
MS1.3	Second Annual Report	30-Nov-20	Achieved
MS1.4	Third Annual Report	30-Nov-21	
MS1.5	Final Annual Report	30-Nov-22	
MS2.1	First version of PaNOSC DP Framework	30-Nov-19	Achieved
MS2.2	Adoption of PaNOSC DP framework	30-Nov-20	Achieved
MS2.3	Implementation of PaNOSC DP framework	30-Nov-21	
MS3.1	Survey of Catalogue APIS and Roadmap to EOSC Integration	30-Nov-19	Achieved

MS3.2	Anthology Feedback to API Tasks	30-Nov-19	Achieved
MS3.3	Catalogue Integration Best Practices Meeting	31-May-21	Achieved
MS4.1	Prototype data analysis services completed	31-May-20	Achieved
MS4.2	Data analysis services accessible through EOSC	31-May-22	
MS5.1	Simulation codes in PaNData Software Catalog	31-May-19	Achieved
MS5.2	Demonstration of simulation services	31-May-20	Achieved
MS5.3	VINYL Software release	31-May-22	
MS5.4	Validation of simulation services	30-Nov-22	
MS6.1	Implementation of AAI integration at the level of the Identity providers	30-Nov-21	
MS6.2	First release of PaNOSC services	31-May-20	Achieved
MS6.3	Second release of PaNOSC services, data and resources	30-Nov-21	
MS7.1	Stakeholder database ready	31-May-19	Achieved
MS7.2	First Sustainability Plan	30-Nov-22	
MS8.1	Joint WP4 & 8 plan	31-May-19	Achieved
MS8.2	Joint WP5 & 8 plan	31-May-19	Achieved
MS8.3	pan-learning.org up running	29-Feb-20	Achieved
MS8.4	Jupyter integrated with e-learning platform	31-May-21	Achieved
MS8.5	e-learning virtual facilities	30-Nov-21	
MS9.1	PaNOSC's Website Ready	31-May-19	Achieved

Risk Register

The table below lists the current risk register for PaNOSC and their assessment in order to give a summary of the risks of the project.

Risk Description	Likelihood	Impact	Risk Level	Consequence	Mitigation / follow-up Measures	Assessment
Participants become less engaged in project	Possible	Moderate	Medium	Failure to deliver as per grant agreement and to contribute to the EOSC	Bi-weekly meetings, Executive Board, face to face meetings, ensure targets of PaNOSC keep being aligned with those of partners	Accept risk after initial mitigation action and continue follow-up actions
Executive Committee deadlock prevents decision making	Unlikely	Minor	Low	Blockage could prevent PaNOSC from adapting to changing circumstances / environment / project challenges	The Executive Committee and its guidelines were created in a way to prevent blockages.	Accept risk after initial mitigation action and continue follow-up actions
Key staff members become unavailable (staff leaving, illness, etc.) ALSO: staff not recruited or delayed recruitment	Likely	Minor	Medium	Failure to deliver as per grant agreement (delays, deliverables/milestones not met) and to contribute to the EOSC	Bi-weekly meetings so partners and WP leaders can raise issues about staff	Accept risk and continue Follow-up actions
Local and/or legal conditions prevent a common data policy.	RISK OBSOLETE: Common DP framework published and accepted by partners. The implementation could be more difficult though.					
Data policy formally but not effectively accepted. The users may be forced to sign it but not fully comply with it (e.g. minimal metadata)	Possible	Moderate	Medium	Less open data available and/or less universally available metadata accross PaNOSC partners	Adequate involvement of the user communities of each facility especially for heterogeneous and distributed facilities.	Accept risk

Risk Description	Likelihood	Impact	Risk Level	Consequence	Mitigation / follow-up Measures	Assessment
Some of the facilities will not be able to meet the mandatory requirements for catalogue API	Unlikely	Moderate	Low	The Search API won't be as useful as originally planned (fewer results provided).	All parties have the source code of their metadata catalogues so they can adapt them. In case changes cannot be made allow exceptions to be granted.	Accept risk
Data Catalog not integrated with data sources	Likely	Moderate	High	API does not return any data and it becomes not very useful	Adequate effort to integrate data sources (e.g. experimental stations) with the catalog	Accept risk
Duplication of efforts/branching of projects/lack of sustainability	Possible	Moderate	Medium	Ineffecient use of research funds, decreased sustainability	Where we use open source and need to modify it to adopt to our needs, we aim to feed those changes back to the open source project. Keep contact with other projects and EOSC stakeholders to ensure no duplication of efforts / minimisation of duplication.	Accept risk but continue with follow-up actions
Lack of skilled staff	RISK REMOVED: Merged into “Key staff members become unavailable“					
Deliverables not met	RISK REMOVED: It is a consequence, not a risk					
Data formats not compatible	Possible	Moderate	Medium	Simulations not interoperable with data analysis	1) Merge openPMD extensions developed in WP5 into the openPMD main standard. 2) Implement Nexus or CXIDB as data format and metadata standard for simulated detector data.	Accept risk
APIs not compatible	Unlikely	Moderate	Low	It won’t be possible to integrate different software produced in separate WPs	ILL team is involved in WP3, WP4 and WP6 and key to the architecture of the project's software. WP5 members (Juncheng) regularly taking part in WP4 meetings.	Accept risk (very unlikely given good collaboration between WPs)
Compute resources not available	Unlikely	Minor	Low	Reduced usability and/or sustainability of software	Engage with EGI for HPC resources	Accept risk
Delay in staff recruitment	RISK REMOVED: Merged into “Key staff members become unavailable“					
EOSC delays	Likely	Minor	Medium	Some PaNOSC services not available through EOSC, but still available	Close collaboration with the EOSC and also ensure that services and tools developed within PaNOSC can work independently outside EOSC	Accept risk
Not clear definition of the EOSC structure	Likely	Minor	Medium	Change in the business model	Close collaboration with the EOSC involved partners and adaptation of business model and sustainability plan	Accept risk
Not clear definition of the EOSC stakeholders	RISK OBSOLETE: Direct collaboration with the EOSC-hub and others ensured this is no longer considered a risk					

Risk Description	Likelihood	Impact	Risk Level	Consequence	Mitigation / follow-up Measures	Assessment
Not common viewpoint	Possible	Moderate	Medium	threat to the sustainability of the PaN EOSC after the end of the project	Constant follow up to the stakeholders opinions through involvement in consultation and targeted communication	Accept risk and continue follow-up actions
Lack of engagement from partners in the costing exercise	Possible	Moderate	Medium	Errors induced in the following WP7 task, unreliable business models	Regular meetings and discussion with all partners on how to assess and report costs.	Accept risk & continue follow-up actions
Development of unsustainable business models	Possible	Moderate	Medium	No support for PaN EOSC services after the project	Constant interaction with stakeholders, clear comunication of requirements and added value	Accept Risk & continue follow-up actions
Ineffective sustainability plan	Possible	Major	High	The PaNOSC portal and other developments may stop being operational/quickly degrade after the project ends.	Early identification of requirements and proposal of actions to ensure the viability of PaNOSC outputs	Accept risk
e-neutron delays	RISK MATERIALISED: The risk materialised without any significant impact to the project					
WP4 deliverables to WP8 are not in due time	RISK MATERIALISED: WP4 was not able to help WP8 with integration of Jupyter in the e-learning platform and hence WP8 has found its own solution. That being said, recent work in WP4 may be adaptable for the e-learning platform. This risk has not had any negative impact as the two WPs are complementary and are both promoting Jupyter					
WP5 deliverables to WP8 are not in due time	Possible	Minor	Low	Training material cannot rely on simulations if not available. This will impact training material for techniques where associated simulations cannot be performed	Some people are active in both work packages, which currently is how we are keeping track of each other. It is first now that we can start to make more detailed plans of what specifically from WP5 should be used in WP8 and how it can be used. E.g. Oasys can for instance only be run via remote desktop and not from Jupyter	Accept risk and keep follow-up action
Difficulty to ensure outreach to key audiences	Possible	Minor	Low	Failure to properly inform and engage project's stakeholders	Update list of target groups on the basis of the stakeholders' database drafted / published in WP7 and ensure relevant communications are delivered to them when required	Accept risk & continue follow-up action
Delay compiling Scientific Use Cases	RISK MATERIALISED: The risk materialised in 2021 when the EB asked the WP leaders to engage with users and gather scientific use cases. The goal of 10 use cases per partner is proving difficult to reach and behind schedule. The reason is lack of communication between scientists and the project members and the need for a better description of what constitutes a use case. The project is working on solving this issue with the help of the EB.					
Risk not to successfully communicating results	Possible	Minor	Low	Failure to deliver as per grant agreement and to properly involve the user community in adopting FAIR open data practices	Regularly interact with WP1 and all other WPs to provide necessary support in informing relevant audiences once results are achieved	Accept risk

Risk Description	Likelihood	Impact	Risk Level	Consequence	Mitigation / follow-up Measures	Assessment
Difficulty managing communications due to the big variety of partners and clusters	Unlikely	Moderate	Low	Duplicated efforts; high travel costs for participation to many events on EOSC	Identification of PaNOSC representatives acting as spokespersons at relevant events; Identification of and participation in target events for specific audiences (e.g. user meetings for the user community); Identification of common actions together with other EOSC clusters	Accept risk & continue follow-up actions
ExPaNDS delays PaNOSC's WP3 due to extra work to manage everything	Possible	Moderate	Medium	WP3 would be late, with added complexity in its products (effort wasted to cover too many scenarios instead of focusing on the right items)	Tobias attending WP3 meetings in ExPaNDS and trying to steer them	Accept risk & continue follow-up actions

Next steps

The Third Annual Meeting is being planned, with dates in the autumn and the hosting organisation being PSI. This event will be a joint event with ExPaNDS for the second year running, further cementing the good collaboration between the two projects. Given the current COVID-19 pandemic it is very likely it will again be an online event.

PaNOSC aims to continue delivering for the remainder of the project as it has been doing so far, with special focus on:

- Ensuring that a PaN Portal is deployed and in operational at all PaNOSC's and at some of the ExPaNDS RIs
- Integrating the Search API within the PaN Portal
- Successfully implementing AAI with UmbrellaID for the PaN Portal, e-learning platform and Search API
- Creating training content in the e-learning platform linked to data and services available through the PaN Portal
- Implement the new data policy at as many RIs and instruments as possible
- Increase awareness of PaNOSC in the PaN Community by gathering more scientific use cases

However several challenges remain:

- Reach the production stage for the PaNOSC developments
- The interaction and integration with the European Open Science Cloud (EOSC)
- Engagement of certain stakeholders (e.g. users) with the project
- Collaboration with other Science Cluster projects

In order to address these challenges

- PaNOSC is reviewing the status of all WPs, starting with WP4 and WP3
- PaNOSC is reviewing its financial status

- PaNOSC is collaboration with ExPaNDS at a high level
- PaNOSC representatives are attending regularly EOSC meetings and conferences to provide feedback about the EOSC, help giving it shape and engage with other Science Cluster projects.
- PaNOSC is planning to attend user meetings and raise awareness of the EOSC and PaNOSC among the research institutes' user communities.