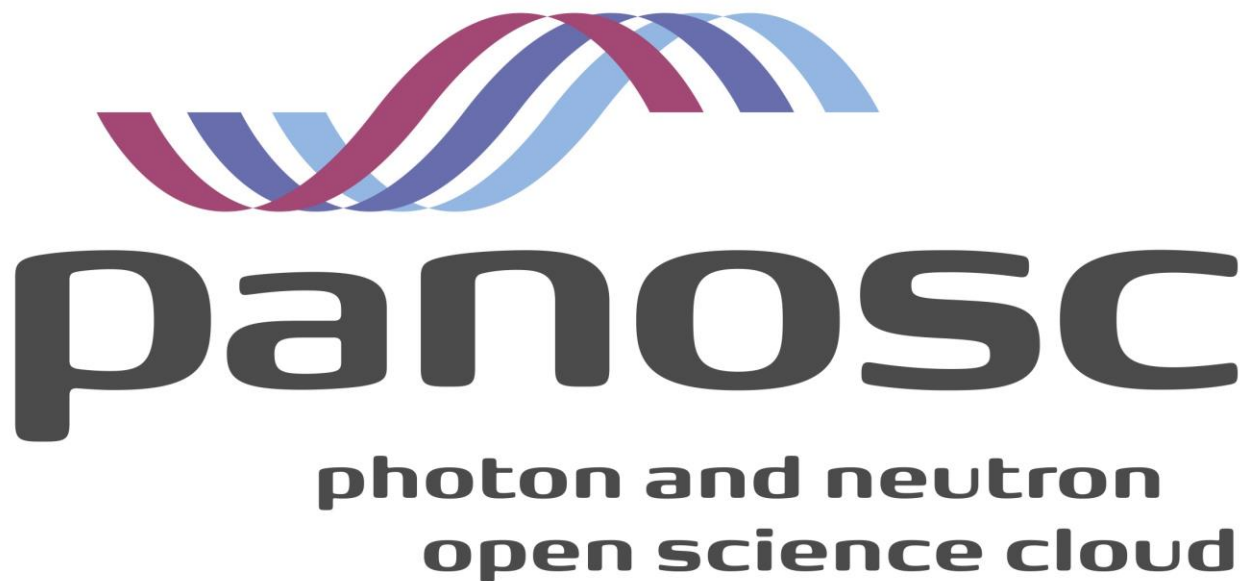


PaNOSC

Photon and Neutron Open Science Cloud

H2020-INFRAEOSC-04-2018

Grant Agreement Number: 823852



Deliverable: D1.5 Mid-year summary 2

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:	D1.5
Deliverable Type:	Report
Dissemination Level	Public
Contractual Delivery Date:	31/05/2020
Actual Delivery Date:	
EC project Officer:	René Martins

Document Control Sheet

Document	Title: Mid-year summary 2
	Version: 1
	Available at: https://github.com/panosc-eu/panosc/tree/master/Submitted%20Deliverables
	Files:
Authorship	Written by: Jordi Boderà
	Contributors: Andy Götz, Tobias Richter, Hans Fanghor, Carten Fortmann-Grote, Juncheng E, Jean-François Perrin, Ornella De Giacomo, Thomas Holm Rod, Nicoletta Carboni
	Reviewed by: Andy Götz
	Approved: Andy Götz

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	5
Executive Summary	5
Summary of Executive Board meetings	5
Third Executive Board meeting - 5/11/2019	5
Fourth Executive Board meeting - 20/1/2020	6
Summary of Project Management Committee meetings	9
Meeting 20/11/2019	9
Meeting 04/12/2019	9
Meeting 18/12/2019	9
Meeting 15/01/2020	10
Meeting 29/01/2020	10
Meeting 19/02/2020	10
Meeting 04/03/2020	10
Meeting 18/03/2020	11
Meeting 01/04/2020	11
Meeting 08/04/2020	11
Meeting 15/04/2020	11
Meeting 22/04/2020	11
Meeting 29/04/2020	12
Meetings in May 2020	12
Progress of work packages	12
Work Package 1: Management	12
Work Package 2: Data Policy and Stewardship	12
Work Package 3: Data Catalog Services	14
Work Package 4: Data Analysis Services	15
Work Package 5: Virtual Neutron and x-ray Laboratory (VINYL)	15
Work Package 6: EOSC integration	16
Work Package 7: Sustainability	16
Work Package 8: Staff and User Training	17
Work Package 9: Outreach/Communication and Dissemination/Impact	19

Key Performance Indicators	21
Comparison between actual and forecasted project status	24
Next steps	26

Introduction

This document summarises the progress achieved in the project since the last management report. As such, this document and the ones that follow on months 30 and 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 joint annual meeting attended by partners observers and ExPaNDS' project representatives took place and deliverables proposed in the grant agreement are being delivered in time, with the partners engaged in the project, regular meetings, recruitment and continuous activity in our main repository:

<https://github.com/panosc-eu/panosc>

Summary of Executive Board meetings

Third Executive Board meeting - 5/11/2019

The 3rd PaNOSC Executive Board meeting was held on the 5 November 2019 during the 1st Annual Meeting held in Trieste. The meeting was chaired by Rudolf Dimper (ESRF) as chair of the EB. The agenda was:

1. Approval of the agenda
2. 1st year financial report and review meeting in Brussels
3. Commitment of partners
4. Deliverable planning
5. Risk management
6. Participation and/or collaboration with future INFRAEOSC-03 and 07 projects
7. Election of the next EB chair
8. Venue and date of the next annual meeting (please bring along your agendas)
9. AOB

The EB discussed the issue of lack of WP leaders and involvement from some partners in certain work packages. They proposed that the partners who are in charge of WPs where WP leaders have changed need to propose a formal replacement plan. The WP leaders were asked to assist in following up work packages by providing intermediate milestones to the project manager (Jordi Boderà). The PMC was charged with providing a plan on how to engage with users once the pilot services are available.

The EB took note that the WP leaders needed to engage more with the risk register together with the project manager to do a regular update.

The EB took note of the mid-term review meeting which was foreseen to take place in Brussels. EuXFEL proposed they could possibly provide a room for the meeting.

The EB discussed the request from the EOSC EB for PaNOSC to provide a position paper outlining the view of PaNOSC w.r.t. to the EOSC and what PaNOSC needs are and what it provides. The EB agreed to review the draft and send their comments to the coordinator before the EOSC week in Budapest.

The upcoming EOSC calls INFRAEOSC-03 and INFRAEOSC-07 were discussed by the board members. It was decided PaNOSC needs more information before it can make a decision whether to participate in these calls or not.

The EB accepted the offer from Allan Weeks (ELI-DC) to host the second Annual Meeting at ELI-Beamlines in Prague. The EB thanked ELI-Beamlines for this offer.

The EB unanimously agreed to elect Thomas Tschentcher (EuXFEL) as chair of the EB for the 2nd year of the project.

Fourth Executive Board meeting - 20/1/2020

The fourth PaNOSC Executive Board meeting was held on the 20 January 2020 via remote video conference. The meeting was chaired by Thomas Tschentcher (EuXFEL) as chair of the EB. The agenda was:

1. Approval of the agenda (ALL)
2. Report from Project Coordinator (A.Götz)
3. Participation and/or collaboration with future INFRAEOSC03 and 07 projects (ALL)
4. Agenda for Review meeting in Brussels (A.Götz)
5. Collaboration with ExPaNDS (ALL)
6. Deliverable planning for 2020 (A.Götz)
7. Update on risk management (A.Götz)
8. Venue and date of the next annual meeting (ALL)
9. AOB

The following summary of the progress was presented:

1. WP1 (ESRF) –
 - reviewed Risk Analysis metrics
 - updated Risk Analysis + KPI tables
 - meeting with ExPaNDS management on a regular basis
2. WP2 (ESRF, ESS, ILL, ELI, CERIC) –
 - produced new draft Data Policy Framework (DPF) + continuing reviewing of DPF
 - analyzing DPF with FAIRsFAIR criteria for FAIR
 - ESRF Data repository certification training with FAIRsFAIR (in 2 weeks)

3. WP3 (ESS, ILL, ESRF, ELI, CERIC) –
 - progress on rolling out OAI-PMH as the first step for EOSC harvesting
 - registering with re3data as well as on the technical side for offering the end point
 - done a few iterations on the federated search API, with feedback from PANOSC partners and ExPaNDS
 - planned a joint WP3 meeting with ExPaNDS to agree the baseline of the API (11+12 Feb) + aligned plans for startup of ExPaNDS
 - discussed a joint meeting on ontologies - still working out if our respective goals are compatible
 - agreed to deliver mock/stub of search API for WP4 prototype (due March, I believe)
 - had session with WP6 on authentication
 - no progress on the integration of data sources task yet
4. WP4 (XFEL, ILL, ESRF, ELI, CERIC) –
 - development of Data Analysis portal has started @ ILL
 - ELI new hire just started, ESRF new hires start next month
 - no contact with ExPaNDS
5. WP5 (XFEL, ESS, CERIC) –
 - proposed API for calling simulation codes
 - continued developing simulation codes (SIMEX, McStas, OASYS)
6. WP6 (ILL, CERIC, EGI, GÉANT, STFC, DESY) –
 - AAI: in addition to Christos we have now Slavek and Mario on board from GEANT, we have organised hands-on 2H00 fortnightly telcos in order to make more significant progress.
 - AAI: We have connected Confluence to the UmbrellaID/EDUTeams platform, It works unfortunately it will be really usable only after the 2nd phase of the pilot when we will have more attributes (with confluence plugin the mapping of account is done with email addresses and unfortunately the current IdP does not provide such attributes)
 - Data transfer for archiving: Test transfer between UKRI and ILL have been achieved, we are now working on automatising the transfer with FTS3. We have organised 1H00 weekly meeting between ILL and UKRI (Ian)
 - Data transfer for analysis services: Tomorrow we should have a demo of OneData transfer between Ceric and EGI.
 - EGI contacting Globus to setup a contract for communities like PaNOSC
 - DESY providing compute resources to ELI-ALPS
7. WP7 (CERIC) –
 - completed the database of stakeholder and selected the survey tool
 - contacted EOSC Sustainability WG chair (R.Lück)
8. WP8 (ESS) –
 - working on Task 8.1 migration of e-learning platform to ESS (on track)
 - organizing face-to-face meeting with ExPaNDS for Spring 2020 in CPH and

workshop later in the year.

9. WP9 (CERIC, ESRF, ILL) –

- produced 3 videos
- communicated on all fronts
- maintaining website up-to-date
- working with ExPaNDS and other cluster communications groups

The participation of all partners in all work packages was discussed. Some partners have not been able to hire staff for all work packages and have therefore not been actively participating in some work packages. This concerns only a few partners and work packages. These were discussed during the EB and will be followed up by the partners.

The list of meetings where PaNOSC was present was presented. PaNOSC is being solicited to collaborate with the other INFRAEOSC-04 clusters in the areas where there is overlap, especially in WP4 and WP8. The lack of available resources (there are no dedicated resources foreseen in the proposal) to participate more actively is slowing down the collaboration. However, the coordinator is convinced there is a strong interest in developing common tools and sharing know-how and solutions among the clusters. This is one of the goals of the EOSC.

A discussion took place about the involvement of the PaNOSC partners in INFRAEOSC-03 and INFRAEOSC-07. It was agreed that Rudolf Dimper (ESRF), will participate in the INFRAEOSC-03 meetings and represent PaNOSC and ExPaNDS. Tiziana Ferrari (EGI) briefly presented the EGI concept for INFRAEOSC-07 and requested that PaNOSC provides a letter of support to the proposal. The goal of the EGI proposal is to provide access to EOSC resources.

An agenda for the PaNOSC mid-term Review Meeting was discussed. It was proposed to have more demonstrators than presentations to highlight what has been achieved. The meeting was tabled for end of June / beginning July in Brussels. Finally the meeting will be held remotely via video conference due to travel restrictions imposed by the COVID-19 pandemic.

The collaboration with the PaNOSC sister project ExPaNDS was presented. The PaNOSC Project Coordinator and Project Manager have met with their ExPaNDS counterparts and discussed how to collaborate. It has been agreed to hold monthly management meetings. A list of common tasks and outcomes has been prepared and presented to the ExPaNDS collaboration. This will be the basis of common activities.

Summary of Project Management Committee meetings

The 1st Annual Meeting¹ of the Photon and Neutron Open Science Cloud (PaNOSC) took place on 4th and 5th November 2019 in Trieste, Italy. This section will cover the Project Management Committee (PMC) meetings of PaNOSC from then until May 2020.

Meeting 20/11/2019

The first PMC meeting after the annual meeting in Trieste took place 20th November in which the deliverables due end of November 2019 were discussed. Work Package (WP) leaders also made balance of the annual meeting and updated on progress in their respective areas.

The summary notes of this meeting² (like of all the others) are available in GitHub³.

Meeting 04/12/2019

After submitting the deliverables due end of November 2019, it was clear that the internal process for managing the concurrent delivery of several deliverables wasn't ideal and the process, formats and timescales for submitting deliverables were discussed in depth, with an agreement reached to share advanced versions of them at least two weeks before the due date in order to have time for an internal review and validation.

It was also discussed that for month 18 (May 2020) six deliverables will be due and that in order to ensure that all of them can be reviewed and validated it is important to advance some of them if possible.

Finally, information obtained from our Project Officer (PO) about the future Review Meeting was shared.

Meeting 18/12/2019

The meeting started with a request for all WP leaders to ensure Milestone supporting documents are stored in GitHub.

Some WP leaders also agreed to advance some of their deliverables due end of May and the call for proposals INFRAEOSC-03 was discussed among the PaNOSC partners.

¹ <https://indico.esrf.fr/indico/event/36/>

² https://github.com/panosc-eu/panosc/blob/master/Work%20Packages/WP1%20Management/Meetings/Project%20Management%20Committee/1st%20Year/2019-11-20-PMC/2019-11-20-PMC_Summary.md

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

Updates from WP leaders included a call to do more regarding KPIs and the management of risks.

Meeting 15/01/2020

INFRAEOSC-03 was the subject that opened the meeting which was followed by a discussion about the Review meeting, expected to take place in Brussels around the month 18 of project execution. In particular it was agreed that doing demos to show the work done so far would be a good approach.

Another hot topic of the meeting was the collaboration with ExPaNDS, for which some measures were announced (participation on each other's mailing list, joint meetings, etc.).

Meeting 29/01/2020

As usual, a PMC meeting took place again two weeks after the previous one. For this instance the opening subject was feedback from an Executive Board meeting where initial discussions about a joint annual meeting with ExPaNDS took place.

More exchanges followed regarding the dates and draft agenda for the Review Meeting in Brussels and also the organisation of the next annual meeting with ExPaNDS.

Meeting 19/02/2020

During the PMC meeting in February discussions continued for the Review meeting in Brussels (2nd half of June preferred), the next Annual Meeting (9th to 11th November proposed) in Prague organised by ELI and the INFRAEOSC-03 call for proposals

It was also agreed to create a PaNOSC space on Zenodo to upload deliverables there and generate DOIs for them.

Meeting 04/03/2020

During this meeting it was announced that the ESRF will seek further details to PaNOSC's PO about the Review Meeting in Brussels and it was discovered that the milestone 8.3 *Pan-learning.org up and running* would be late by approximately one month (new expected delivery date for end of March).

Jamie Hall was nominated as the PaNOSC representative for ExPaNDS' Technical Coordination Board.

Further news were shared about the INFRAEOSC-03 call and Research Institutes from either PaNOSC or ExPaNDS that could be interested in participating.

Meeting 18/03/2020

The second PMC meeting of March started by confirming that no one had any issue/ conflict of interest with the proposed reviewer by our PO. It was also confirmed that due to the ongoing COVID-19 pandemic the review meeting will take place remotely.

It was also announced that a replacement for a key person for WP8 (who is on long-term sick) was found and that deliverable D7.1 PaNOSC Stakeholder feedbacks was progressing satisfactorily.

Meeting 01/04/2020

The first point of the agenda confirmed that pan-learning.org, our e-learning platform is now available and that milestone MS8.3 was accomplished.

Regarding the Review Meeting, dates were discussed and two possible dates agreed up (which were later proposed to our Project Officer) and it was also agreed to meet more frequently in order to prepare well this event.

Meeting 08/04/2020

This meeting focused entirely in the upcoming Review Meeting.

The date for the Review Meeting had been confirmed with the PO (16th June) and then the attendees agreed to develop and propose a schedule as well.

The list of possible attendees to this remote Review Meeting was also reviewed and refined (WP leaders, ensuring all partners are represented, former WP leaders when adds value, a representative from our Executive Board and from ExPaNDS and supporting staff for demos), with the idea to propose again to our PO a desired number of attendees.

The meeting ended discussing the idea of having a significant number of software demonstrations as the central point for many WP during the Review Meeting.

Meeting 15/04/2020

This meeting focused on progress made by different WPs, as there were many deliverables due at the end of May 2020.

Meeting 22/04/2020

The Review Meeting preparations were followed during this meeting, refining the list of proposed attendees, the schedule and the demonstrations to be done.

A table with all the key tasks/events and dates for the Mid-Term Report and Review meeting was shared.

Meeting 29/04/2020

This was another meeting focusing on WPs recent work and issues. Most WPs reported their deliverables to be well advanced and some required final reviews / collaboration from others. The EOSC Rules of Participation document was also discussed.

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. Four meetings were planned :

- 6th and 20th May focusing on the preparations for the Review Meeting scheduled for 16th June
- 13th and 27th May focusing on WP progress and deliverables due at the end of May

Progress of work packages

Work Package 1: Management

Following the 1st Annual Meeting, the deliverable D1.4 Annual Report was written and submitted on time.

The next deliverable for WP1 is this one, however in the meantime WP1 has supported 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
- Preparing for the Review Meeting and Mid-term Report
- Plan and organise the next Annual Meeting
- Setting-up and monitoring internal processes
 - Internal milestones to better follow project progress
 - Maintenance of GitHub and documents stored
 - Internal financial report
 - Follow-up of issues
- Support of mailing lists
- Support the risk management process
- Support the introduction of Key Performance Indicators (KPIs)

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 under control and carries on delivering as per the Grant Agreement.

Work Package 2: Data Policy and Stewardship

WP2 worked the last 6 months on producing the PaNOSC Data Policy framework in collaboration

with the ExPaNDS WP2 members. The first draft, based on the PaNdata Data Policy framework, was produced during the breakout meeting at Trieste during the 1st Annual Meeting. Input was gathered from the partner sites, respectively ESRF, ESS, CERIC-ERIC, ILL, EuXFEL, and HZB. The first draft included feedback from the partner sites based on their experience adopting and applying the PaNdata framework. The next phase consisted in inviting all members of the ExPaNDS WP2 to contribute to modifying the draft to fully take into account the FAIR principles. A series of meetings (roughly ten) were required to review in detail every point in the framework and produce the final version which was also deliverable D2.1. The 41 criteria of the RDA FAIR Data Maturity Model (FDMM) were used to verify the implementation of the FAIR principles. The new PaNOSC data policy covers 90% of the FAIR criteria. The remaining criteria are dependent on the implementation at each site. The main improvements to the framework with respect to the PaNdata framework are:

- respects all 4 FAIR principles and 90% of FDMM criteria
- distinguishes between metadata, raw, processed, and auxiliary data
- extension of best practices to software used and electronic logbooks
- extends and clarifies many terms in the glossary
- addition of implementation notes
- termination of custodianship clause
- support for machines and machine learning
- recommendation to use the community standard metadata standard (Nexus/HDF5)

The next step is for each partner to adopt the policy and update their existing policy (if they already have one). This activity will be led by CERIC-ERIC.

ExPaNDS WP2 participated actively in the preparation of the PaNOSC data policy framework with the aim of having a common data policy framework for the whole community. ExPaNDS will continue the work of PaNOSC to produce a common data policy for the PaN community. PaNOSC WP2 will contribute actively to this activity. The goal is to have one data policy framework for the photon and neutron community to replace the PaNdata policy.

One of the partners, ESRF, has been selected as part of the FAIRsFAIR⁴ programme to certify their data catalogue by the Core Trust Seal certification. Two members of the ESRF attended a workshop⁵ organized by FAIRsFAIR in Den Haag on the certification process in February 2020. They will continue with the process until the end of 2020.

The WP2 members of PaNOSC and ExPaNDS participated actively in the EOSC FAIR activities by attending and providing feedback on the FAIRsFAIR deliverables.

⁴ <https://www.fairsfair.eu/>

⁵ <https://www.fairsfair.eu/events/fairsfair-1st-certification-support-workshop-data-repositories>

Work Package 3: Data Catalog Services

Participants of the work package met at the general assembly, as well as a meeting at ESS in Lund in February. This meeting was held jointly with the ExPaNDS project. The cataloguing work package in ExPaNDS, also labeled WP3, has by design close ties with PaNOSC and this meeting was there to align efforts and to converge on the search API deliverable. The main topics of the meeting were:

- Review different catalogue systems in use by partner institutions
- Search API and harvest API discussions, OAI-PMH status
- API data model, search and query formulation, units conversion and handling
- Integration with analysis service work packages
- Finalisation of API, preparation for deliverable

Presentations from this workshop are available on the event page:

<https://indico.esss.lu.se/event/1373/>

Notes taken during the meeting are also available at

<https://confluence.panosc.eu/pages/viewpage.action?pageId=18186323>

Numerous participants from ExPaNDS took part in a joint meeting the first time; they were introduced to the other participants and to the work that had been carried out up to the meeting. Their presence ensured that activities between PaNOSC and ExPaNDS are aligned as much as reasonably possible. Alun Ashton gave an introduction to WP3 in ExPaNDS, which must cover most of the topics WP3 covers in PaNOSC, with a special emphasis on developing an ontology in Task 3.2.

Differences and similarities between the search and harvesting APIs were discussed. Facilities need to agree what level of granularity is necessary in order to publish data. A coarser granularity model is easier to define and implement, but may not provide the details that users expect when searching. On the other hand the granularity should not be site specific and still give common properties the same terms, in order to fully open the data to non-domain (or site) experts. The WP will start with a coarser definition for metadata and work towards standardising more and more search terms during the course of the project. In a similar way to what is planned for the harvesting metadata schema.

Also of interest is the issue of units handling: how measured quantities stored and/or searched for using various units can be handled. There are some existing software packages aiming to solve similar problems and these need to be investigated. A baseline version of the API will be delivered as D3.1 at the end of May.

With the baseline of the search API in sight, the integration aspect with WP4 (Data Analysis Services) came into closer view. It lies in the fact that search results are a natural input to data analysis services: the user searches for some data, then would like to proceed with analysis/visualisation of the data he or she found. On top of ensuring the functionality above, it is a possible requirement from WP4 that search results be previewed/visualised on the user interface before proper data analysis is attempted. For this to be supported, pre-generated image files need to be attached to the dataset and/or datafiles.

In this period work started on the catalogue integration task, led by ELI. The scope was discussed in the Lund meeting and video conference meetings are planned between ELI and CERIC on this

topic.

Work Package 4: Data Analysis Services

WP4 continued to work on planning and creating an architecture for the remote data analysis services. In addition to the survey (submitted as D4.1) that was conducted with all the partner facilities to learn about their requirements, an additional extended survey was performed later to gather more information, which included the ExPaNDS partner facilities.

The feedback from the surveys, as well as a number of meetings, was used to create and refine early drafts of the portal architecture, which is now being developed using an agile and iterative approach. This agile approach allowed us to release a very early prototype of the portal, which has been deployed and tested by most PaNOSC facilities. This approach will be continued, with updates being tested by all partners as they come out so that feedback can be quickly taken into account, ensuring that the portal is suitable for all partners. One of the challenges is to integrate the different variants of computing infrastructure that is used by the different facilities.

Additionally, facilities have made progress on deploying local instances of JupyterHub and remote desktop services to enable preliminary remote analysis while the PaNOSC portal is being developed. This allows users to get accustomed to the concepts of Jupyter, Jupyter Notebooks, and JupyterHub, so that the move to the portal is easier and requires less training. By providing these services now, we help improve the state of remote data analysis right from the start and start to engage future users.

Further investigation done for the portal design has lead to the use of other technologies which enhance the FAIRness of the data and analysis, such as Singularity containers for the packaging and archiving of computational environments, and Spack for creating a mirror of all of the source code required to fully recreate an environment with no need for external sources.

PaNOSC partners have worked on multiple open source projects, including Jupyter itself, as well as tools used by many of the partner facilities such as h5py (used by NeXuS), h5glance, hdf5plugin, pyFAI, and a number of other projects which aim to enhance remote analysis, such as improvements to remote desktop tools and ways to remotely inspect HDF5 files.

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

After reaching the first milestone (M5.1: Addition of simulation codes to the pandata software catalogue) and D5.1 (openPMD domain extensions), the work in WP5 is now focused on the harmonization of simulation APIs. The simulation API constitutes the backbone of the next Milestone (M5.2: Demonstration of simulation services) and Deliverable D5.2: Release of documented simulation APIs.

From April 20th to 30th, members of WP5 have joined their forces during a development sprint. After reviewing the design principles and software requirements for the API, the partners split up into three groups to work on

- 1) the abstract base classes of the API (pyvinyl library),
- 2) a prototype source-to-experiment simulation for high-power laser generated neutron diffraction including DFT sample simulation,

- 3) demonstration of x-ray diffraction simulation with SIMEX inside the graphical simulation framework Oasys, and
- 4) exploration of a NoSQL database solutions for simulation data.

The outcomes of this sprint were documented in several demo Jupyter notebooks showing how SimEx takes the output data of OASYS complying to the openPMD wavefront domain extension defined in D5.1 and generates the diffraction patterns using GAPD SimEx diffraction calculator with the wavefront data. Other notebooks demonstrate how to use ASE to simulate a powder diffraction sample and the pyvinyl API to perform the neutron scattering simulation including the structural relaxation, neutron-matter interaction and signal generation. All notebooks can be found at <https://github.com/PaNOSC-ViNYL/workshop2020/tree/master/demo>

Work Package 6: EOSC integration

Work on the deployment of UmbrellaID on the eduTEAMS infrastructure has continued and the infrastructure upgraded. New services have been set up like the confluence platform or the ILL central SSO service based on the Keycloak software. They have both taken advantage of the availability of the OIDC (OpenID Connect) protocol available on eduTEAMS to simplify the integration setup. The PaNOSC training platform plans to integrate UmbrellaID as Identity provider in the next few weeks.

Regarding the data transfer activities, we have mainly focused our effort on the integration of the EGI DataHub solution with RI archive and Jupyter Services. A pilot is running and will be demoed during the review meeting; unfortunately, the COVID-19 crisis across Europe has slowed down the deployment effort in this period as the IT team of the RIs had to play a key role in enabling teleworking and remote experiment when possible. We now have to reproduce this set up for more RI and run large acceptance tests.

The archive transfer activity has also made some significant progress as the technical solution has been set up and tested between UKRI-STFC and ILL datacenters. Like the EGI dataHub pilot this activity has also suffered from the current crisis concerning the performance tests as the network bandwidth was artificially limited in order not to disturb the remote activities of the RIs. Nevertheless these activities are aligned with the PanOSC DoW program.

Work Package 7: Sustainability

During this period, the partnership further refined the list of the main stakeholders, proposing also the contact points. Many of the activities of this task were described in deliverable 7.1, delivered on time. The personal contacts of the stakeholders were not included in the deliverable for privacy reasons, being a public deliverable. In addition to the identification, the partnership has taken any opportunity to actively interact with its main stakeholders, and also follow the activities of other initiatives that included consultations with stakeholders, trying to avoid duplication of efforts. Following and influencing (whenever possible) the trends and outputs of the EOSC WGs and other initiatives (FAIRsFAIR, RDA, EOSC-Hub) became an important activity for the sustainability of the

PaNOSC. WP7 (and PaNOSC) partners conveyed its feedback in different ways. One of these channels were the EOSC consultation events, where we took part in the sessions related to sustainability (e.g. The EOSC Coordination Day, Budapest 28-29/11/2019; EOSC Consultation Day, 18/05/2020). Finally, WP7 established a connection with the Sustainability WG through personal exchanges with the Co-Chair, and coordinated the collection of feedback from the partners to different documents, as requested by the EOSC Secretariat (e.g Sustainability WG, tin man documents, RoP, etc.).

The second main activity carried out was the definition of the metrics and costs for the Photon and Neutron EOSC community. This involved an exercise where the goal, scope and methodology were discussed and agreed upon. The list of the services, activities and resources to be included in the data workflows is being refined by the partners as the cost collection in the partners' sites proceeds.

Work Package 8: Staff and User Training

During the last six months, focus has been on:

1. Getting pan-learning.org into operation
2. Getting content in pan-learning.org and aligning with ExPaNDS

1. Pan-learning.org into operation

Pan-learning.org is now in operation (since end of March, 2020) running on a server hosted by ESS (Milestone 8.3). pan-learning.org is the result of migrating e-neutrons.org to ESS. In the process, security and stability has been improved and the layout of the web site has been changed to better mimic its role for both photon and neutron scattering and its support from PaNOSC. Continuous integration based on Travis has been implemented. Alexandre Stefanov, ESS and Peter K. Willendrup DTU/ESS have done most of the work related to migrating the platform to ESS.

Pan-learning.org can also be accessed from the domain name elephant.school, where elephant is an acronym for elearning for photon and neutron training. Whether pan-learning.org, elephant.school or a third domain name should be the persistent one still needs to be decided.

The new pan-learning.org site has already successfully been used for a course on neutron scattering at University of Copenhagen with 25 students enrolled, however, without using the simulation engine (i.e. McStas).

ESS specific courses and instrument simulations are available in pan-learning.org, mostly due to another EU project SINE2020, but it nonetheless encompasses the internal milestone IMS8.8.

A solution for integration of Jupyter into e-learning.org (IMS8.2) has been identified. The idea is to use <https://github.com/jupyterhub/ltiauthenticator>, which enables a Moodle and Jupyter instance to communicate. More information can be found at <https://moodle.org/mod/forum/discuss.php?d=362641>. However, in the meantime other

integration solutions have been identified, which also needs to be evaluated. For instance as seen in the NSF funded WHOLETEALE project (<https://www.wholetale.org>) or as discussed in the Moodle forum (<https://moodle.org/mod/forum/discuss.php?d=362641>).

Kareem Galal at ESS has started to integrate federated AAI with assistance from WP6 (ILL and GÉANT). Likewise, WP6 (ILL) is exploring the opportunities for getting cloud computing resources for pan-learning.org from DESY or EGI.

2. Getting content in pan-learning.org and aligning with ExPaNDS

Unfortunately for this work package, a person crucial for the planning and execution of training activities has fallen ill and is now on sick leave. This, together with COVID-19, has delayed the planning of a workshop in Task 8.2 on how to use pan-learning.org. Two persons who are or previously have been involved in the development of the e-learning platform have agreed to take over the organization of the workshop. One of those persons is a staff member at the Niels Bohr Institute, University of Copenhagen (UCPH), and the other is a faculty member from the Department of Science Education also at UCPH. The specific terms are still being negotiated. The workshop is now planned to be split into two workshops, with the first one planned to be held Q4 this year and a follow-up in Q1, 2021. For the participants, we will aim for 50% being “data oriented” members of PaNOSC and the other half being more science focused, i.e. people with experience and interest in teaching other scientists. Currently, we are collaborating with ExPaNDS WP5 and the coordinators of the two projects on identifying the right people to participate in such a workshop.

In an effort to coordinate between ExPaNDS and move the work package forward, we have assembled a task force consisting of the work package leaders for WP5 in ExPaNDS and for WP8 in PaNOSC, namely Thibaud Cayla (SOLEIL) and Thomas H. Rod (ESS) in addition to Peter K. Willendrup (DTU/ESS) and Florian Gliksohn (ELI) from PaNOSC and Uwe Konrad (HZDR) and Nazare Guimard (SOLEIL) from ExPaNDS.

A Gantt chart with both ExPaNDS and PaNOSC activities have been generated and will be used as the starting point for coordinating activities between the work packages. The Gantt chart is located at [Google Docs](#). At a meeting on May 18, 2020, with the PaNOSC and ExPaNDS coordinators it was agreed that the two projects should collaborate on getting content in the platform, and that we should start to plan more specifically for which courses should be developed so that we can get the appropriate assistance from work packages and other staff to develop those courses.

Work Package 9: Outreach/Communication and Dissemination/Impact

A report summarizing the outreach activities carried out during the 1st year of implementation of the project has been drafted and published on PaNOSC's GitHub repository: <https://bit.ly/2T12rVr>

As foreseen in PaNOSC communications' strategy, the work in WP9 has supported the promotion of the work carried out in all WPs, as follows:

- **WP1** – Regular exchanges with the project coordinator, also through participation in meetings of the Project Management Committee.
- **WP2** – Promotion of the latest achievement in the WP: release of the PaNOSC data policy framework.
- **WP4** – A summary of deliverable D4.1 was published on the PaNOSC website and further promoted via its social media channels to the project's community and partners. A video interview with WP4 leader is being produced, with a summary of the goals and expectations foreseen in the WP, and a presentation of the OSCOVIDA platform (<http://oscovida.github.io/>), which showcases the technology being used and developed in PaNOSC for the analysis of PaN data science in the context of the COVID19 situation, and which can be used by citizens and policy makers worldwide to better understand the unprecedented pandemic: <http://bit.ly/2PS9eyl>
- **WP5** – A summary of Deliverable D5.1 was published on the PaNOSC website and further promoted to PaNOSC stakeholders through the various channels available: <http://bit.ly/2Egldz3>
- **WP6** – Dissemination activity (publication and distribution of an article and a video) about the achievement of the 1st phase of implementation of UmbrellaID integration with EduTEAMS, which implies the migration of services from existing UmbrellaID infrastructure to the new Umbrella AAI powered by GÉANT. The video also gives an overview of the following phases, which will allow the new Umbrella AAI to be available to the PaN user community.
- **WP7** – WP9 will take into account the work carried out in WP7 and to be published in D7.1 – PaN EOSC stakeholder feedback, to increase the impact of the communications' activities foreseen in the project.
- **WP8** – WP9 has been interacting with WP8 to design a new logo for pan-learning.org training website, and optimize the descriptions in the platform, to then start a wider promotion across stakeholders.

OTHER ACTIONS IN WP9

- Support in the promotion of all events organized in other WPs.
- Update of templates and other visual materials.
- PaNOSC Zenodo Platform set up: <https://zenodo.org/communities/panosc/>
- Video publication and promotion:

- Interview with PaNOSC coordinator, Andy Götz:
<https://www.youtube.com/watch?v=Vn0Gb6nHRyw>
- Interview with Mathew Bowler on structural biology and the EOSC:
<https://www.youtube.com/watch?v=IYVXoBzeXNo&t=1s>
- Interview with Jon Taylor on EOSC: <https://www.youtube.com/watch?v=bZ7Q-71PrVo>
- Presentation of PaNOSC at public events (January 2019 - EOSC day at CNRS; January-February 2020 - DESY, European XFEL and ESRF user meetings; April 2020 - FAIRsFAIR 2nd Synchronisation Workshop; May 2020 - European RIs for a smarter future digital conference, ORSO Workshop)

COLLABORATION WITH ExPaNDS

- PaNOSC and ExPaNDS have regularly interacted to coordinate and harmonize the communications actions, also in preparation of the next projects' annual meeting scheduled in November 2020. Both projects mutually support each other to communicate about their goals and disseminate results also via their social media channels.
- ExPaNDS' section on the PaNOSC website has been updated with more information about the joint work carried out in the frame of the two projects.
- Two videos mentioning PaNOSC have been released by ExPaNDS. The 1st one presents the general aims and scopes of the projects, and the 2nd one highlights the benefits of open FAIR data for the users of free electron laser, synchrotron light and neutron sources. Both videos are available here: <https://bit.ly/3dDg0Cn>

COLLABORATION WITH OTHER PROJECTS and CLUSTERS:

- Regular meetings with EOSC cluster projects and FAIRsFAIR have taken place. Major common tasks in the projects have been identified, together with the main contact persons involved. The aim is to link them and promote harmonization of the work, avoiding duplication of efforts in the various common tasks. Due to the COVID-19 pandemic, a joint event previously scheduled at ESOF 2020 had to be cancelled, but interaction is ongoing. All projects have the need to highlight use cases and a 1st draft of a common template has been shared to better communicate about the potentials of the EOSC and the possible uses that can be made of the services developed in the projects.
- PaNOSC, in coordination with the FILL2030 H2020 project, has been carrying out the preparatory work for an action to advocate the editors of scientific journals to actively support the citation of data DOIs in their published articles. A letter will be sent once a video promoting the use of data DOIs across the PaN user community, will be released. The video, sponsored by PaNOSC, is an update of the one previously published by ILL, and has received the support of all photon and neutron facilities in PaNOSC and ExPaNDS projects, as well as LEAPS and LENS initiatives.

Key Performance Indicators

PaNOSC has agreed to define a set of Key Performance Indicators (KPIs), with most of the indicators and initial values decided around the time of the first annual meeting (November 2019).

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on
1	PaNOSC percentage of PMs spent	N/A	58%	31/08/2019
1	PaNOSC percentage of target expenditure	N/A	53%	31/08/2019
1	Deliverables submitted late (as per EC Portal)	N/A	3	30/11/2019
1	% deliverables submitted late (as per EC Portal)	N/A	30.00%	30/11/2019
1	Milestones submitted late (as per EC Portal)	N/A	1	30/11/2019
1	% Milestones submitted late (as per EC Portal)	N/A	10.00%	30/11/2019
2	number of instruments on which Data Policy is implemented	16	21	30/11/2019
2	number of techniques / instruments for which metadata are defined	9	11	30/11/2019
2	Number of instruments available	97	97	30/11/2019
2	Percentage of techniques for which metadata is defined	33%	33%	30/11/2019
2	number of metadata parameters defined	500	1,070	30/11/2019
2	number of raw data DOIs	500	589	30/11/2019
2	number of user defined DOIs	5	12	30/11/2019
2	number of downloads	5	10	30/11/2019
2	number of citations of DOIs	5	10	30/11/2019
2	number of datasets	5	10	30/11/2019
2	volume of data archived	5	10	30/11/2019
2	number of datasets cited in publications	0	2	30/11/2019

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on
3	number of PANOSC facilities represented in the NeXus advisory committee	2	3	30/11/2019
3	number of PANOSC facilities offering public metadata via OAI-PMH	0	1	30/11/2019
3	number of facilities offering a PANOSC catalogue search endpoint	0	0	30/11/2019
3	number of datasets with FAIR data leaving embargo period in the coming year	0	2	30/11/2019
4	Percentage of facilities offering JupyterHub or remote desktop for analysis services	50%	67%	30/11/2019
4	Number of unique users making use of JupyterHub or remote desktop for analysis services at partner facilities	0	156	30/11/2019
4	number of techniques available through remote services	2	8	30/11/2019
5	Number of contributors to ViNYL	0	6	18/11/2019
5	number of users of ViNYL services at ESRF	0	0	18/11/2019
5	number of users of ViNYL services at ILL	0	0	18/11/2019
5	number of users of ViNYL services at XFEL	0	0	18/11/2019
5	number of users of ViNYL services at ESS	0	0	18/11/2019
5	number of users of ViNYL services at ELI	0	0	18/11/2019
5	number of users of ViNYL services at CERIC	0	0	18/11/2019
5	Number of modules included in ViNYL service	0	0	18/11/2019
5	Number of partner infrastructures that have used ViNYL service	0	0	18/11/2019
5	Number of DOIs for simulated data (by counting datasets with "ViNYL" labels on open-access repositories like Zenodo).	0	1	18/11/2019
5	Number of openPMD standard domain extensions merged into mainline openPMD	0	0	18/11/2019

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on
	repository			
6	Service Providers connected to eduTeams Umbrella	0	0	06/11/2019
6	Services connected and accessible through eduTeams Umbrella	0	0	06/11/2019
6	Users that have used at least one service	0	0	06/11/2019
6	Number of partners that have set up data transfer	0	0	06/11/2019
7	WP7 has no KPIs for now			
8	Number of facilities adapting the e-learning platform for their own needs	1	1	30/11/2019
8	Number of programs integrated into the e-learning platform	2	2	30/11/2019
8	Instrument techniques considered in the e-learning platform	6	6	30/11/2019
8	Number of staff participating in e-learning workshop (T8.4)	0	0	30/11/2019
8	Number of staff participating in data stewardship course (T8.5)	0	0	30/11/2019
8	Number of users completing e-learning course (T8.7)	0	0	30/11/2019
8	Number of students participating in summer school(s) (T.8.8)	0	0	30/11/2019
9	Number of social media posts mentioning PaNOSC	0	116	10/11/2019
9	Number of followers on Twitter	0	256	10/11/2019
9	Number of user meetings in which PaNOSC is presented	0	0	10/11/2019
9	Number of invited talks as PaNOSC representatives on topics of relevance for the project and its partners	0	11	10/11/2019

WP	KPI description	Value before PaNOSC	1st Value obtained	1st Value obtained on
9	Number of visitors of the PaNOSC website	0	1186	10/11/2019

The actual values of these KPIs will be updated at least once a year in order to show the progress made by the project.

Conversations with ExPaNDS have also taken place and the KPIs listed above shared, supporting ExPaNDS and enabling the possibility that both projects share many KPIs.

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	
D1.7	Mid-year summary 3	31-May-2021	
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	
D2.4	Integration of the policy	30-Nov-2021	
D3.1	API definition	31-May-2020	Submitted
D3.2	Demonstrator implementation	31-Mar-2021	
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	

Milestone or Deliverable Id	Name	Due Date	Status
D5.1	Prototype simulation data formats	30-Nov-2019	Submitted
D5.2	Documented simulation APIs	30-Nov-2020	
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	
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	
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	

Milestone or Deliverable Id	Name	Due Date	Status
MS2.3	Implementation of PaNOSC DP framework	30-Nov-21	
MS3.1	Survey of Catalogue APIS and Roadmap to EOSC Integration	30-Nov-19	
MS3.2	Anthology Feedback to API Tasks	30-Nov-19	Achieved
MS3.3	Catalogue Integration Best Practices Meeting	31-May-21	
MS4.1	Prototype data analysis services completed	31-May-20	Planned
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	Planned
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	Planned
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	
MS8.5	e-learning virtual facilities	30-Nov-21	
MS9.1	PaNOSC's Website Ready	31-May-19	Achieved

Next steps

The Second Annual Meeting is being planned, with the dates (9th to 11th November) and location (ELI Beamlines, Dolní Břežany, Czech Republic) confirmed. This event will be a joint event with ExPaNDS, further cementing the good collaboration between the two projects. The final decision to hold a physical or virtual event will be taken end of August 2020 and depend on the travel restrictions in vigor then. The web site for the event is: <https://indico.eli-beams.eu/event/369/page/1>.

PaNOSC aims to continue delivering for the remaining of the project as it has been doing so far, however several challenges remain:

- The definition and implementation of the European Open Science Cloud (EOSC)
- Engagement of certain stakeholders (e.g. users) on the project
- Collaboration with other cluster projects

In order to address these challenges

- PaNOSC representatives attend regular EOSC meetings and conferences to provide

feedback about the EOSC, help giving it shape and engage with other cluster projects.

- PaNOSC is planning to attend user meetings and raise awareness of the EOSC and PaNOSC among the research institutes' user communities.