PaNOSC
Photon and Neutron Open Science Cloud
H2020-INFRAEOSC-04-2018
Grant Agreement Number: 823852

Deliverable: D6.3 - Integration of the PaN AAI into the EOSC.

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:	D6.3
Deliverable Type:	Report and Demonstrator
Dissemination Level	Public
Contractual Delivery Date:	31 November 2021
Actual Delivery Date:	13 January 2022
EC project Officer:	Flavius Pana

Document Control Sheet

Document	Title: Integration of the PaN AAI into the EOSC
	Version: DRAFT
	Available at: https://github.com/panosc-eu/panosc
	Files: 1
Date	30 December 2021
Authorship	Written by: Jean-François Perrin
	Contributors: Christos Kanellopoulos, Jayesh Wagh
	Reviewed by: Erwan Le Gall, Jamie Hall
	Approved: Jordi Bodera

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

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

Integration of the PaN AAI into the EOSC

Table of Content

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

Summary

The Photon and Neutron community (PaN) gathers scientific analytical facilities such as Synchrotrons, Lasers, Free Electron Lasers and Neutron sources and their user communities. The Federated Identity Management (FIM), commonly called Authentication and Authorisation Infrastructure (AAI), is a critical topic in the PaN community due to the federated nature of the services provided by the different and on many aspects independent Research Infrastructures (RIs) that compose the community. One of the key objectives of the PaNOSC project is to transform the community AAI, UmbrellaID, for reaching full compatibility with the European Open Science Cloud (EOSC) architecture.

Since its inception in 2012, UmbrellaID (https://umbrellaid.org/what.html) has been serving the AAI needs of the PaN community. Within the PaNOSC project, work package 6 "EOSC integration" aims to integrate the PaNOSC cluster services within EOSC. The EOSC AAI[1] is currently not finalised and further developed inside the EOSC Future project. Nevertheless, the general architecture[2] has already been defined following the AARC projects recommendations. The AARC Blueprint Architecture 2019[3] and the AARC Interoperability Guidelines being the foundation of the EOSC AAI model, they underpinned all developments in this domain.

This integration activity significantly benefited from the long-standing collaboration between UmbrellaID and GÉANT. GÉANT team brought in this work its long-established technical expertise in AAI technology and organisation to achieve the integration of PaNOSC services in EOSC. In this document, we describe the current achievements of UmbrellaID done in PaNOSC WP6 and the roadmap to join other communities in an interoperable, trustworthy and secured authentication and authorisation framework inside EOSC AAI.

Brief history of UmbrellaID

UmbrellaID.org started out in 2012 as a collaboration between the main European analytical user facilities in the field of research using photons and neutrons (PaN facilities). The European PaN facilities have more than 50000 users. 40% of these users use different PaN facilities in Europe to do their experiments. The initial objectives were and still are to:

1. Provide a common Authentication and Authorization Infrastructure that allows users of all PaN facilities to connect seamlessly to digital facility services with a single and unique ID.

2. Authorization should remain in full control of each facility providing the services.

3. Users should be in full control of their personal data.

[1] https://data.europa.eu/doi/10.2777/8702
[2] https://op.europa.eu/s/twJE
[3] https://doi.org/10.5281/zenodo.3672784

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

The umbrellaID collaboration connects the photon and neutron communities, Hence, it is essential to preserve this visibility and keep ownership and governance within the community.

In 2015, the PaN facilities[4] signed a memorandum of understanding (MoU) to establish an efficient long-term collaboration between the PaN facilities in order to facilitate authentication and authorisation procedures to access trans-facility user services. This MoU allowed PaN facilities to jointly develop, implement and operate the UmbrellaID, a unique and persistent identity for users of the European Analytical Facilities.

The UmbrellaID collaboration comprises a Steering Committee (SC)[5] which is in charge of organisational and managerial aspects, and a Technical Team (TT)[6] in charge of the technical aspects of the collaboration. Each of the participating PaN facilities appoints one representative on the SC and the TT. TT meet fortnightly to discuss operations and projects implementation, while SC meetings are organised once a year to discuss and vote on strategic questions.

Technically we have always tried to keep the UmbrellaID infrastructure at a high level of reliability but also as simple as possible to minimise the maintenance operations. Before PaNOSC, only 1 geographically replicated Identity Provider (IdP) and 1 Resource Registry (RP) were in operation to serve the community. Distributed user's metadata was limited to the ID of the user.

Collaboration with GÉANT.

Having all PaN RIs willing to participate in the EOSC construction, it appeared early during the construction of the PaNOSC proposal that UmbrellaID had to evolve technically but also in terms of users' metadata proposed and policies in order to match the AARC BluePrint Architecture (BPA) prerequisites[7]. We drafted specifications that we shared with potential EU partners in order to help us address this necessary evolution.

1. EOSC integration

UmbrellaID users can link their eduGAIN account with their umbrellaID account, allowing them to maintain a single account. We need to go further and allow complete interoperability with EOSC services in terms of access, support and security.

2. Security policies

Policies regarding password changes, audit and user vetting must be introduced. Two factors of authentication should be proposed to the users. It is necessary to look at security issues at a global level and participate in international efforts.

[4] https://umbrellaid.org/where.html
[5] https://umbrellaid.org/sc.html
[6] https://umbrellaid.org/tt.html
[7] https://aarc-project.eu/architecture/

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

3. Attributes

We need to extend the set of attributes (e.g. email address, affiliation, ...) in full compliance with the different regulations regarding personal data protection.

4. Operation

The current operations of the IdPs are supported by a team of 2 engineers not solely dedicated to this task. We need to extend the support team in order to have a constant cover and surveillance of the services' operations.

5. Innovation

FIM and authentication technologies are constantly evolving; we would like to be at the forefront with these advancements. This would be easier to achieve in a larger technical collaboration.

For instance, we would like to solve proxy authentication use cases (typical of REST web Services) by introducing technologies like OpenID Connect.

Solutions for non-web-based access are still emerging and either not really user friendly or limited to few use cases. We hope that a larger collaboration could provide real advancements for the user communities.

6. Training and service support

FIM technologies are still relatively complex and demand specific training and support, especially for the service providers' engineers, in order to enable the rapid growth of the number of services in a secure environment.

At the same time GÉANT was developing the eduTEAMS[8] offer for communities that largely match the needs of the PaN community. Due to the numerous positive collaborations that occurred between GÉANT and UmbrellaID over the years of existence of UmbrellaID, we easily established a partnership between the two organisations based on trust and mutual understanding.

Within PaNOSC we have built a stronger relationship with GÉANT and adopted the eduTEAMS service operated by GÉANT as the core infrastructure. This decision has largely facilitated the integration with the EOSC AAI model which is still in development and helped us not only to closely follow the development but also to participate in this evolution by presenting our use cases and needs.

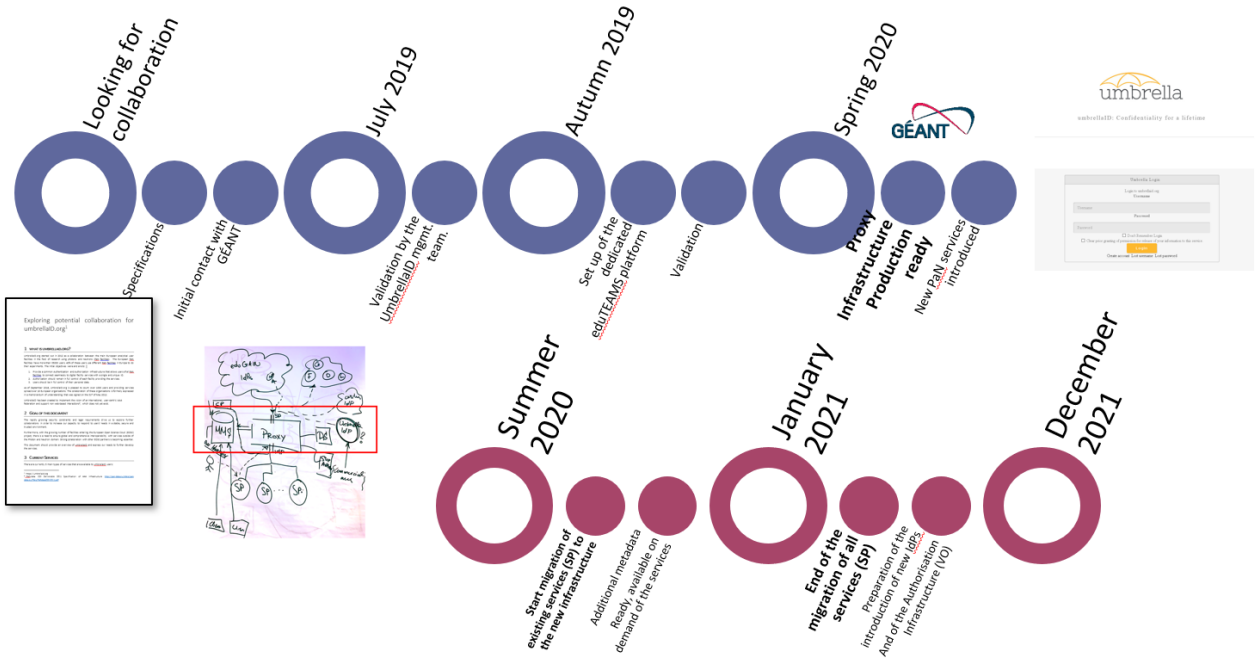
Migration of the UmbrellaID infrastructure to the eduTEAMS platform.

The following diagram illustrates the different steps of the migration that took place in the course

[8] https://eduteams.org/

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

of PaNOSC.



We have organised a series of technical meetings between the two partners, to analyse the specificities of UmbrellaID and tried to address smoothly the integration with eduTEAMS:

- The ID of the users should remain the same over the career of the scientists, which is not the case with eduGAIN when a scientist moves from one affiliation to another one. This feature could be addressed by introducing some public alternative IdPs like ORCID alongside eduGAIN IdPs. This would ensure that a scientist will not loose its UmbrellaID when he moves from one affiliation to another one.
- In the original UmbrellaID model, the users' authorisation, typically regarding data sets access, are managed locally by the services of the RI, no information is available to help authorisation decision at the level of the community AAI. This legacy model is not really compatible with the spirit of EOSC where a user should be able to compose data services from different communities. It has been to increase the number of metadata in order to meet REFEDS[9] Research and Scholarship (R&S) recommendations and to propose an authorisation mechanism at the level of UmbrellaID based on AARC-G002[10] "Expressing group membership and role information" standardisation principles.

Set up of the core infrastructure

The eduTEAMS infrastructures for UmbrellaID was set up during the autumn 2020.

[9] https://refeds.org/category/research-and-scholarship
[10] https://doi.org/10.5281/zenodo.5502533

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

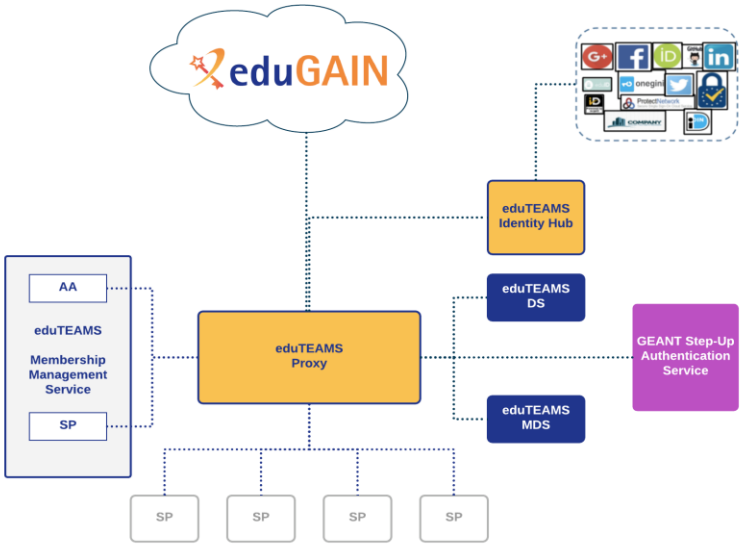


Figure 2 eduTEAMS typical architecture

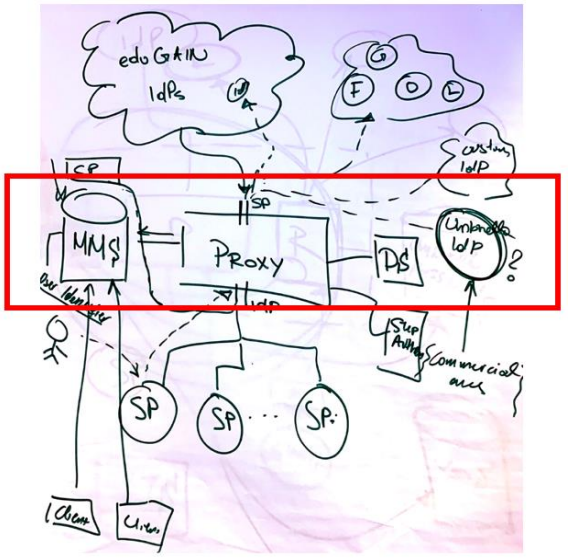


Figure 1 Sketch of the UmbrellaID setup on eduTEAMS

Two infrastructures, one dedicated to production and one for acceptance tests have been setup. Each of them proposes the proxy component at the core with a Metadata service (MDS) and the Membership Management service (MMS).

Migration of the existing Service Providers

After a period of tests on this new infrastructure, we started to migrate the existing Service Providers (SP) of UmbrellaID to this new configuration.

In order to facilitate this migration, we organised a 1-day training for the IT specialists of the community. This technical training workshop[11] was organised in February 2021. The aim was to introduce and to share the recent technical and organisational developments. It was necessary to ensure that the IT professionals of the PaN community are at ease with the concepts, processes and technologies in use, and can eventually actively participate in the evolution of our community AAI.

The workshop included presentations, hands-on session and live demonstrations. Over 40 IT professionals from PaN community attended this workshop remotely. A second edition of UmbrellaID training workshop will be organised in the first quarter of 2022.

All the different SP were migrated by the end of March 2021.

[11] https://indico.psi.ch/event/10773/

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

UmbrellaID Facility	Services Url
DESY	https://door.desy.de/door/
ALBA	https://useroffice.cells.es/
Elletra	https://vuo.elettra.eu/
ESRF	https://smis.esrf.fr
ILL	https://userclub.ill.eu/userclub/
MAX IV	https://duo.maxiv.lu.se/duo/
PSI	https://duo.psi.ch/duo/merge_accounts.php
SOLEIL	https://sun.synchrotron-soleil.fr/sunset/bridge/sunset/
Community Services	
PaN E-Learning	https://pan-learning.org/moodle/login/index.php
PaN training portal	https://pan-training.hzdr.de/
WayForLight	https://wfl.elettra.eu/
PaN Software catalogue	https://software.pan-data.eu/

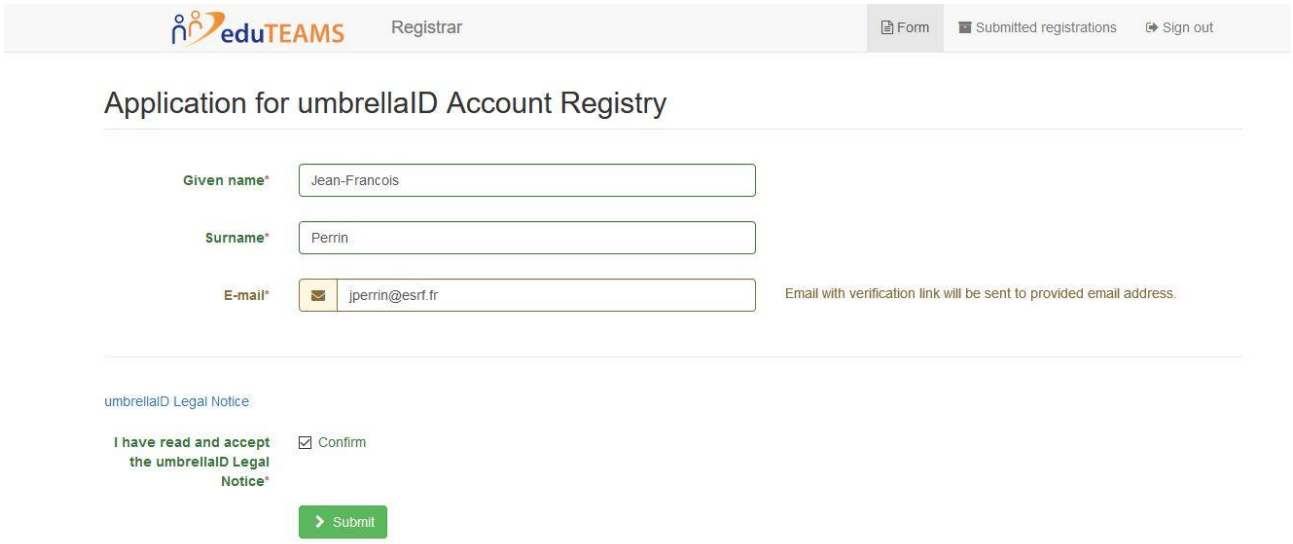
This represents a major achievement for WP6, as from this stage we are part of the EOSC AAI. Since then all new community services have used this new infrastructure, ready for EOSC.

Users' Metadata

We have then extended the list of metadata available at UmbrellaID to match the EOSC requirements and to be able to solely based authentication and authorisation on UmbrellaID for some services. The first service to benefit from it was the PaN software catalogue. The technical work has been simplified by using the MDS service of the eduTEAMS platform.

Currently only the minimum bundle requirement from the REFEDS R&S is in place (shared user identifier, person name, email address) we would like to extend it to the user's affiliation and ORCID. A decision is in preparation for the next UmbrellaID steering committee that will take place in January 2022.

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

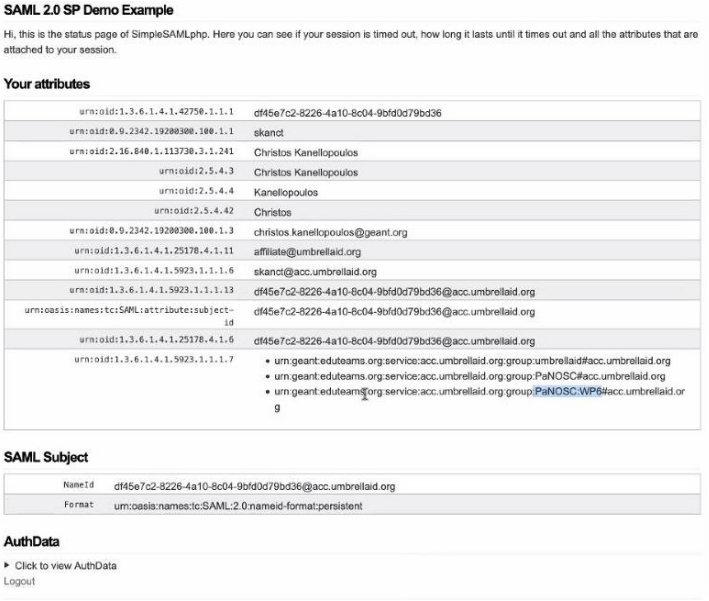


Currently these metadata are only available for SP that request it, following the 2022 steering committee decisions it should become the default for all services.

Future plans

In the course of 2022 we have planned two major actions because the end of the PanOSC project.

We would like to introduce the notion of Authorisation in UmbrellaID, this feature is necessary for users that want to consumes data services hosted by other communities or providers. Technically we would like to base it on the eduTEAMS MMS service. Tests have already been performed and a successful implementation realised with the EGI DataHub service. We now have to obtain the validation for the UmbrellaID SC and decide on the VO organisation.



We are also planning the Integration of other IdPs, once again this is technically ready thanks to the eduTEAMS platform. By Integrating eduGAIN as Identity providers, UmbrellaID users will be able to use their home organisation credentials for benefitting of the PaN services. By integrating

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

ORCID, RIs will be able to collect automatically ORCID identifiers of the users (if authorised by the user). Other public IdPs could also be integrated to simplify the access for the PaN users and ensure that they will continue to benefit from the same authentication even if they move from one affiliation to another. Decision on the list of IdPs to be integrated will be taken during the next UmbrellaID SC.

Conclusion

By engaging with GÉANT and adopting the eduTEAM platform, UmbrellaID has been able to evolved drastically and be ready for the EOSC AAI in the course of the PaNOSC project. All new services developed by the community, even beyond PaNOSC partners, are able to integrate EOSC and be ready in terms of AAI with very few efforts. This is an important milestone for PaNOSC. By continuing to work with GÉANT and other communities in the scope of EOSC-Future and AEGIS[12], we should be able to establish an EOSC AAI were interoperability between the communities will not anymore be a challenge.

[12] https://wiki.geant.org/display/AARC/AEGIS

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