31-10-2024

Milestone 3.1:

Use-case collection and Analysis

Abstract

This document, which is the draft version of the deliverable D3.2 "Use-Case Collection and Analysis" due in February 2025 provides an overview of the approach taken to conduct a survey within the Research Infrastructures and e-Infrastructures participating in the AARC TREE, AEGIS and in FIM4R to identify emerging requirements and use-cases that should be supported by the enhancement of the AARC BPA framework.

The results of the survey will serve as input to produce two deliverables due in WP3, but also for the technical (WP1) and policy (WP2) work in Y2, as well as for the "Compendium & Recommendations" work package (WP5) which will deliver recommendations for a common long-term strategy for AAI services in pan-European Research Infrastructures in Europe.

Table of Contents

1. Introduction

The Work Package 3 (WP3) "Use Cases Collection and Analysis" of the AARC Tree project, leads the use-cases and emerging requirements gathering across the Research and e-Infrastructures that participate that operate an AARC BPA compliant AAI, which gather in AEGIS, the AARC Engagement Group for Infrastructures; many of them participate in the AARC TREE Project.

The aim of the work page is twofold: on one hand the work package aims to asses the deployment of AARC BPA compliant AAIs as well as the adoption of the AARC guidelines that accompany the BPA; on the other hand the work package aims to collect new requirements that should be supported by the AARC BPA. These are two different aspects but complementary that will be reflected in two separated deliverables:

- Landscape analysis of AARC BPA adoption (D3.1) due in December 2024
- Use-Case Collection and Analysis (D3.2) due in February 2025

This document, which formally is a milestone, sets the ground for both deliverables and presents the approach taken to gather both use-cases and new requirements, as well as to assess the adoption of the AARC guidelines across the communities that have deployed an AAI that follows the AARC BPA.

The results of the work conducted in WP3 serves as an essential input to Architecture (WP1) and Policy (WP2) work packages to steer their work on revising the AARC BPA and its guidelines with respect to the gathered implementation experience and further requirements.

The body of work is significant. More than 8 years have passed since the initial AARC BPA and the initial guidelines were published. In the meanwhile there have been a large number of its implementations within the Research and e- Infrastructures in Europe and beyond.

Over the years more guidelines have been produced and endorsed by AEGIS; AEGIS, gathers the operators of AAIs that are compliant AARC BPA, was established at the end of the AARC2 (2019) project to maintain and evolve the AARC guidelines.

AARC TREE provides an excellent framework to review the adoption of these guidelines.

In the beginning of the AARC Tree Project (March 2024), in collaboration with WP1, WP2 and WP5 "Compendium & Recommendations", a questionnaire was developed in order to capture the relevant questions for the Research and e-Infrastructures about their AAI use-cases and requirements, and adoption of the AARC BPA and its guidelines. The second half of 2024 was devoted to conducting comprehensive interviews with Research and e-Infrastructures.

This document starts with describing the methodology of collecting the relevant data, and then describes how the questionnaire results are being elaborated. As the responses to the questionnaires are still being collected at the time of writing this document, they are not included in this draft version.

2. Use-case collection methodology

During the kickoff meeting of the AARC Tree project, it was agreed that the use-cases collection and analysis is an essential input for the majority of Working Packages. Developing a methodology for collecting relevant data, along with ensuring content accuracy, was deemed crucial, and required feedback from all project participants. Accordingly, the following methodology was collaboratively established:

-   **Interview questions template** - a template with relevant questions was developed, so that the answers could be captured for each Research and e-Infrastructure in the same way. This provided a basis to later produce a comprehensive and meaningful analysis, resulting with indispensable information guiding the further development in AARC TREE project.

-   **Interview organisation** - The interviews were conducted in online meeting sessions. To ensure consistent answers to the questions, it was decided that the best approach would be to engage with the Research and e-Infrastructures through conversations, allowing context to be explained and relevant responses to be captured more effectively. Before the interview, a list of questions was shared with the Research and e- Infrastructures which provided a basis for them to know which information to prepare making the meeting effective. Since it takes significant time to organise and hold a relatively high number of the interviews, this phase was stretched from June 2024 to November 2024, taking into account holiday season.

-   **Research and e-Infrastructures to interview** - The primary focus of the interviews were the Research and e-Infrastructures implementing an AAI that is compliant with the AARC BPA. Since the contributors to Work Package 3 are representatives of several Research and e-Infrastructures (R/e-infras) that run an AARC BPA compliant AAI, they were the first group interviewed. In addition to the R/e-infras represented in AARC TREE, WP3 also targets R/e-infras that participate in AEGIS and will make an effort to also include some that at the moment do not adopt this framework but have expressed an interest in doing so.

-   **First analysis and feedback** - After the first two interviews were held, which helped refine the questions, the questionnaire was shared with the AARC TREE participants, collecting the first round of feedback once few real data was in. Based on this the questionnaire was updated and all of the interviews progressed as planned. After all of the data was collected, a first analysis of the result will be produced and shared with AARC Tree participants through a workshop. This will establish a feedback loop which will help to finalise the analysis report.

-   **Engagement with relevant stakeholders** - This work is not happening in isolation and engagement is a key aspect. This includes: a presentation at the FIM4R meeting that will take place during TechEX 2024 (in Boston, USA); a presentation for the OSCARS workshop in October 2024 (OSCARS is the EC funded project that brings together world-class European Research Infrastructures in the ESFRI roadmap and beyond); providing continuous updates

for AEGIS during the monthly calls; and engagement with the EOSC Technical and Semantic Interoperability Task Force which re-started in October 2024.

The developed interview questions template is presented in Appendix A. It contains a number of qualified questions that provide a basis for statistical analysis. It also contains a number of open ended questions which are to be analysed to provide meaningful information. To keep an manageable overview of the questions, the questionnaire is broken down into the following sections:

- **Section A - General information** - meant to capture basic information about the Research Infrastructure e.g. services offered, user audience etc.

- **Section B - AAI solution** - the information about used AAI solution and implementation of AARC BPA and Guidelines, providing comments from the perspective of relevance and implementation experience

- **Section C - Policy for Access Management** - the information about used policies for access management, required user attributes and employed mechanisms to perform user authorization

- **Section D - Security** - the information about security incident practices in place

- **Section E - Workflow** - description of the researcher workflows with the regard of gaining access to services

- **Section F - Requirements** - further requirements, gaps and challenges when it comes to implementation of AAI/access management

The finding of the interviews will be handed over to WP1, WP2 and WP5 as an input for their further work, especially with regards to describing the emerging requirements and what is needed to address them in the context of the AARC BPA and related policy and technical guidelines. In collaboration with the Architecture, Policy and the "Adoption and Validation" work package (WP4), some of the use cases (those that are cross research collaborations) will be selected and used to drive the implementation of a validator that will allow for automatic tests in a real environment of the proposed technical and/or policy guideline(s). The validator is meant to offer a test suit for similar use-cases.

3. Overview of interviewed infrastructures

During the use-cases and requirements collection phase of this work, 18 research and e-infrastructures were interviewed, as well HPC communities. In addition to that two main large service providers that are also participating in AARC TREE were interviewed. The image below shows the distribution of the interviewed Infrastructures per the field of science they serve.

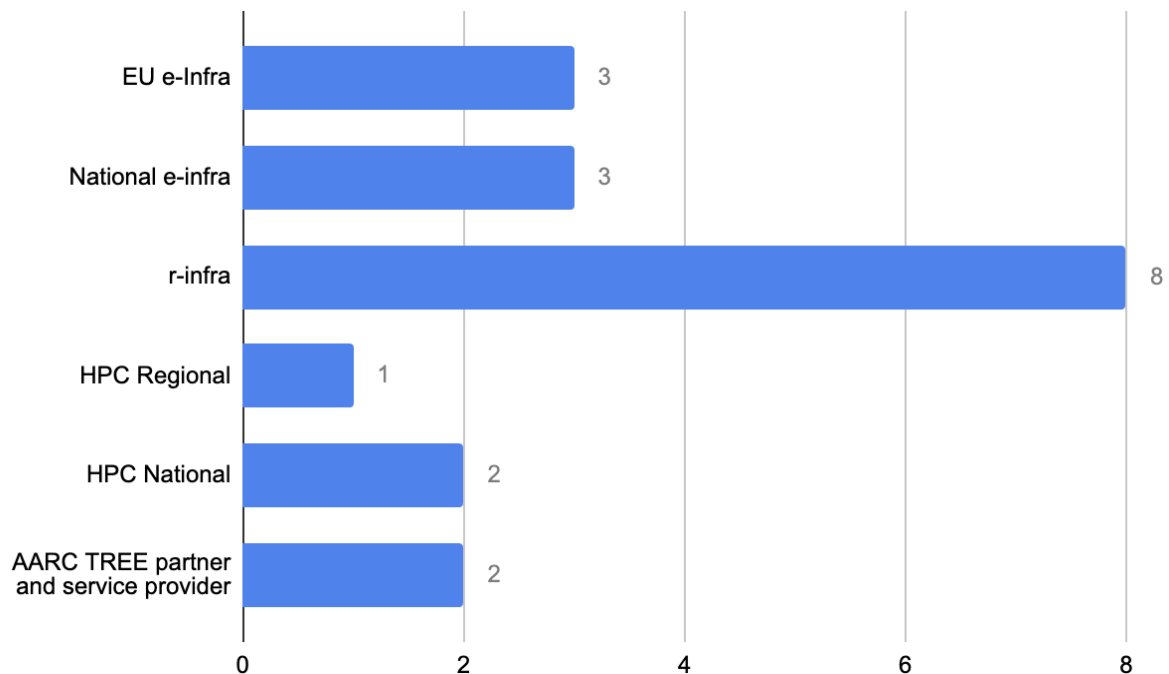The list of R/e-infras interviewed is listed in the table below.

Figure 1 - Distribution of Infrastructures per field of science they serve.

Name R/e-Infra	Short description	Type
GÉANT	GÉANT is a fundamental element of Europe's e-infrastructure, delivering the cutting edge pan-European network as well as identity services to support scientific excellence, research, education and innovation. In AAI context, GEANT provides several implementations of the AARC BPA, namely MyAccessID (for the HPC community), MyAcademicID (for the education community), and the EOSC AAI.	EU e-Infra
EGI	EGI is a distributed e-infrastructure established to provide advanced computing services for research and innovation across Europe. In the context of the AAI, EGI provides the EGI Check-In, an implementation of the AARC BPA to serve different communities.	EU e-Infra
EUDAT	EUDAT is a collaborative Data Infrastructure. In the context of the AAI it offers B2Access, an implementation of the AARC BPA.	EU e-Infra
SURF (SRAM)	SURF is the National Research and Education Network in the Netherlands. They offer the SRAM as a national research access platform that enables federated access to services provided by SURF and other SURF members aimed at the research community.	e-infra (National)
NFDI (National)	NFDI the German national research infrastructure (Nationale Forschungsdaten Infrastruktur)	e-Infra (National)
IRIS (National)	IRIS, that UK national digital research infrastructure.	e-infra (National)
European Photon and Neutron Community	The European Photon and Neutron Community (PaN)	r-infra
ESCAPE	European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures)	r-infra
SeaDataNet	The distributed Marine Data Infrastructure for the management of large and diverse sets of data deriving from in situ of the seas and oceans.	r-infra
NEIC (PUHURI)	PUHURI is a resource allocation system for computing services. It also provides reporting and access management to compute services. LUMI is the main user of PUHURI services.	HPC (Regional)
KOMONDOR HPC	KIFÜ HPC	HPC (National)

(national)		
Julich	Forschungszentrum Jülich - Jülich Supercomputing Centre	HPC (National)
ELIXIR	A distributed infrastructure for life-science information. Elixir is part of the Eosc Life Cluster.	r-infra
ELI	The Extreme Light Infrastructure ERIC (ELI ERIC)	r-infra
EISCAT	EISCAT Scientific Association/EISCAT3D	r-infra
DARIAH	DARIAH is just one of the RIs in SSHOC cluster. https://www.sshopencloud.eu/news/sshoc-ssh-open-cluster This interview will focus on Dariah, but we can have a bit of feedback about all the social and humanity sciences.	r-infra
LHC (CERN)	European Organisation for Nuclear Research (CERN) and Worldwide LHC Computing Grid	r-infra
CINECA	CINECA, FENIX	AARC TREE partner and service provider
NIKHEF	NIKHEF	AARC TREE partner and service provider

Table 1: List of initial R/e-infras interviewed

Most of the interviewed Infrastructures are members of EOSC Association, while a handful of them are also participating in citizen scientist programmes. In the remainder of this section, the list of interviewed Infrastructures and the basic information about them is given.

4.   Deployed AAI Solutions

Since the AARC2 project, it became clear that for many research infrastructures operating an AARC BPA AAI is a demanding task. The Life Science community paved the way to a different approach: they argued that e-Infrastructures are best positioned to offer these type of services that enable research infrastructures to carry out their science.

This model was successfully tested in AARC2 and after that deployed during the EOSC Life project.

Other research infrastructures have taken a similar route; by offloading the operations to an e-Infrastructure the research infrastructure still can retain control over the policies that the AAI should support.

5.   Conclusion

This is an initial snapshot of the work that is taking place. The first deliverable will be published in December 2024 and more insight will be provided.

Annex A - Interview template

Section A - General Questions

A.1 Which Research Infrastructure (RI) are you representing?

A.2 Which field of science are you serving ? (Frascati manual of Fields of Research and Development (FORD)) (

| Optional comment: |
| Please choose ▼ |

A.3 Please provide description about the research infrastructure (e.g. which kind of infrastructure and related services are delivered and by whoom, is there a formalised collaboration etc.).

A.4 Please provide description of the user audience - type of users (research, citizen scientists, industry users), number of users, distribution over the globe and organisations.

A.5 Is the RI member of European Open Science Cloud (EOSC)?

| Optional comment: |
| Please choose ▼ |

A.6 Is the RI participating in Citizen Science Programmes or other initiatives or programmes?

| Optional comment: |
| Please choose ▼ |

Section B - AAI Solution

B.1.Describe the currently running solution for authentication and authorisation infrastructure (AAI). Which specific authentication methods being used to cater for different user audience (e.g Institutional accounts (eduGAIN), ORCID, Social media, Others - please specify)

B.2 Is your AAI solution compliant to AARC BPA (blueprint architecture)? Please add comment about BPA implementation (challenges in implementation, clarity, technical difficulties etc.)	
Please chose ▼	Optional comment:

B.3 Which AARC guidelines are you implementing? Please add comment about guidelines implementation (challenges in implementation, clarity, technical difficulties etc.)

Guidelines for expressing community user identifiers (AARC-G026)	Please choose ▼	Optional comment:
Guidelines on expressing group membership and role information (AARC-G002) (superseded	Please choose ▼	Optional comment:
Guidelines on expressing group membership and role information (AARC-G069)	Please choose ▼	Optional comment:
Specification for expressing resource capabilities (AARC-G027)	Please choose ▼	Optional comment:
Guidelines for expressing affiliation information (AARC-G025)	Please choose ▼	Optional comment:
Inferring and constructing voPersonExternalAffiliation (AARC-G057)	Please choose ▼	Optional comment:
Exchange of specific assurance information between Infrastructure (AARC-G021)	Please choose ▼	Optional comment:
Guidelines for evaluating the combined assurance of linked identities (AARC-G031)	Please choose ▼	Optional comment:
A specification for IdP hinting (AARC-G049) (superseded by AARC-G061)	Please choose ▼	Optional comment:
A specification for IdP hinting (AARC-G061)	Please choose ▼	Optional comment:
Specification for hinting an IdP which discovery service to use (AARC-G062)	Please choose ▼	Optional comment:
A specification for providing information about an end service (AARC-G063)	Please choose ▼	Optional comment:
Guidelines for Secure Operation of Attribute Authorities (AARC-G071)	Please choose ▼	Optional comment:

Section C - Policy for Access Management

C.1 Does the Research Infrastructure have an access policy? (governs who can access the infrastructure, under what conditions).

C.2 Is there a formalised procedure to manage access rights to services (e.g. cooperation agreement, call for application and evaluation, ad-hoc individual order/access, member of an organisation, etc.)?

C.3 What are the requirements for identification of the users (e.g. what is required information, are there requirements for LoA or streinght of the authentication method)?

Unique Identifier Attribute	Please choose ▼	Optional comment:
Name Attribute	Please choose ▼	Optional comment:
email Attribute	Please choose ▼	Optional comment:
Affiliation Attribute	Please choose ▼	Optional comment:
Organisation Attribute	Please choose ▼	Optional comment:
Level of Assurance Attribute	Please choose ▼	Optional comment:
Multi Factor Authentication	Please choose ▼	Optional comment:
Other	Please choose ▼	Optional comment:

C.4 How do you implement the policy for access management (e.g. how is the individual who can access the research research data/measurement data/your research instrument identified and authorised)?

C.5 Which AARC BPA authorisation model (see Implementing scalable and consistent authorisation across multi-SP environments AARC-I047) does your service currently implement or plan to implement?		
Centralised Policy Information Point (Proxy aggregates and provides user attributes to end-services)	Please choose ▼	Why do you use this? Provide Use case example:
Centralised Policy Management and Decision Making (Proxy informs end-services about authorization decisions)	Please choose ▼	Why do you use this? Provide Use case example:
Centralised Policy Management, Decision Making, and Enforcement (Proxy enforces decisions directly at the proxy level)	Please choose ▼	Why do you use this? Provide Use case example:
Centralised Policy Information Point (Proxy aggregates and provides user attributes to end-services)	Please choose ▼	Why do you use this? Provide Use case example:
Other (please specify)	Please choose ▼	Why do you use this? Provide Use case example:
C.6 Which authorisation in your Research Infrastructure based on?		
Group memberships/roles	Please choose ▼	
Resource capabilities	Please choose ▼	
Assurance	Please choose ▼	
Affiliation within Home Organisation	Please choose ▼	
Organisation domain	Please choose ▼	
Other (please specify)	Please choose ▼	
C.7 Is there a need to map external authorisation information (e.g., group memberships, resource capabilities from other Community AAIs) to resource capabilities understood by the end-services in your Research Infrastructure?		
Mapping is required?	Please choose ▼	Please describe the mapping process or any challenges you face:

Section D - Security

D.1 Do you have a designated Data Controller (in the context of GDPR)?

D.2 Do you have a designated security contact for AAI related security incidents?

D.3 Does the AAI adhere to SIRTFI or other recognised security frameworks?

Section E - User Access Workflow

E.1 Can you describe the research workflows? (especially from the perspective of how the user access)

Section F - Requirements

F.1 Can you describe further requirements, gaps and challenges your RI is facing when it comes to implementation of AAI/access management?

References

REF01   URL or doi reference information

Appendix A.1

Do you have appendicitis?