05-10-2015

Deliverable DJRA1.1: Analysis of user community and service provider requirements

Deliverable DJRA1.1

Contractual Date:	31-08-2015
Actual Date:	05-10-2015
Grant Agreement No.:	653965
Work Package:	JRA1
Task Item:	JRA1.1
Lead Partner:	EGI.eu
Document Code:	DJRA1.1
Editors:	Christos Kanellopoulos, Nicolas Liampotis, Niels van Dijk, Peter Solagna

The research leading to these results has received funding from the European Community's Horizon2020 Programme under Grant Agreement No. 653965 (AARC).

Abstract

This document, produced by JRA1 Task 1 "Analysis of user community requirements", identifies the requirements of user communities and services providers building upon the outcomes of previous activities such as the TERENA AAA Study and the FIM4R workshop series. The requirements identified by these activities have been updated and enriched with new requirements that the team collected through a survey of user communities as well as a set of targeted interviews. These requirements are analysed here and will be provided as input for upcoming activities in AARC.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

This document provides an analysis of the requirements the team has collected up to now from scientific communities and project stakeholders. It is an outcome of the "Requirement Analysis" activity in the JRA1 work package, "Architecture for an integrated and interoperable AAI". For the requirement-gathering process, we followed a three-step approach. Section 2 elaborates on the requirement extraction methodology adopted by AARC to ensure that the proposed integrated AAI framework can meet the needs of all stakeholders.

As a first step, we returned to the results of previous activities such as the "TERENA AAA Study" and the FIM4R workshop series and analysed their outcomes in order to produce an initial set of requirements, presented in Section 3.1.

During the second step, this set of requirements was enriched through the results of a survey that was run by AARC during the summer period. The goal of this survey was to identify the barriers communities are currently facing to adopting and using federated access, and to capture their requirements for an interoperable AAI. A total of 10 scientific communities have responded to this survey and the results are presented in Section 3.2 of the document. In addition, a set of one-to-one discussions were held with key representatives from EGI, ELIXIR, EUDAT and GÉANT, in which we discussed their plans, their successes and the barriers they are facing. The results of these discussions are presented in detail in Section 3.3.

As part of the third and final step, the extracted requirements have been extensively processed, homogenised, merged, filtered, extended and classified to produce a refined set of requirements formulated into tables, shown in Section 4. We have identified 25 distinct requirements, 18 relating to tools and architecture and 7 relevant to policies and best practices. The document's conclusions are drawn in Section 5.

Finally, the questionnaire used for the AARC internal survey is presented in 0, while a brief description of the main AAI use case for each participating user community is provided in Appendix B.

2 Introduction

Controlling access to research-related resources and collaborative tools is challenging, particularly when dealing with research communities that can be geographically dispersed across Europe and the globe. National identity federations for Research and Education (R&E), whereby the user's identity is verified by the institution that issues the user's credentials, enable users to access different services with the same credentials, while allowing research e-Infrastructures to offer resources in a more controlled and consolidated way.

While eduGAIN serves as a global interconnection framework for national identity federations, multiple interoperability gaps still exist between the different Authentication and Authorisation Infrastructures (AAIs) that are operated by the national federations and the various research collaborations and e-Infrastructures. Besides interoperability issues, there are also functional aspects that have yet to be addressed, such as identity attribute aggregation from multiple providers, Single Sign On (SSO) support for non-web applications, delegation capabilities, and credential translation services. In addition to addressing these technical aspects, integration of AAIs requires significant efforts to define a common policy framework covering the necessary legal and operational practices for all entities involved in the AAI ecosystem, including identity providers, attribute providers, resource providers and federation operators.

The AARC project aims to build on the developments and deployments that have led to the AAI frameworks used to date to deliver a common framework for authentication and authorisation that will allow research communities to share information resources and services easily and effectively across e-Infrastructures in a secure, well-controlled fashion, while at the same time reducing operational burden. AARC brings together different e-Infrastructure providers, as well as libraries and research communities, in order to capture and analyse the requirements for a Pan-European identity federation for researchers, educators and students. This approach ensures user-community engagement, avoids duplication of efforts by reusing existing AAI frameworks, and creates the conditions to enhance these frameworks to address community requirements.

The goal of this document is to collect the needed requirements to enable the design and pilot of an integrated Federated Authentication and Authorization Architecture for e-Infrastructures.

3 Methodology

Before AARC, several other initiatives have explored the requirements for federated AAI capabilities. The work presented in this document builds upon these efforts in order to capture and analyse today's requirements for a Pan-European identity federation for researchers, educators and students.

The findings from previous studies provided an insight into the different user communities involved and the current status with respect to the penetration of federated identity management. AARC researchers then collaborated to devise questions that were circulated to many of these communities. It should be noted that the project used a questionnaire (see 0) as part of its methodology in order to gain access to a larger sample of user communities than could be feasibly achieved through direct contact. However, to gather more detailed information about the requirements described in the survey's responses, a series of interviews were also conducted with a subset of the participating communities. These were held in the form of informal conversations and semi-structured interviews with key users of the selected communities who have an extensive understanding of the needs and issues of their AAI.

More specifically, the methodology used for this deliverable to analyse requirements comprises the following steps:

1. gathering of initial requirements from a multitude of external sources, including input from previous activities such as the "TERENA AAA Study" and the documents produced by FIM4R,
2. evaluation, analysis, filtering and refinement of initial requirements based on collected feedback from AARC internal survey and interviews with selected but broad representation of user communities,
3. harmonisation, classification and ranking of requirements.

The described methodology will also serve as the basis for extracting use cases in a continuous and iterative process the results of which will be recorded on the project's wiki[1]. In addition, the information gathered will allow for an assessment of the authentication and authorisation technologies adopted by the user communities. The latter results will be documented in a different report to be shared among AARC partners (MJRA1.1).

[1] https://wiki.geant.org/display/AARC/AARC+Home

4 Input from stakeholders

4.1 Past activities

4.1.1 FIM4R

FIM4R (Federated Identity Management for Research) is an international cross-domain activity which organised a series of workshops with the purpose of discussing and proposing solutions to the problems experienced by research infrastructures that need FIM capabilities in order to operate their facilities and serve their user communities. Through these workshops, FIM4R produced a set of common requirements and recommendations for the uptake of FIM, which are documented in [FIM4R2012]. This report considers information originating from a diverse selection of research communities:

- High Energy Physics, represented by the WLCG collaboration and CERN;
- Life Sciences, represented by the ELIXIR ESFRI;
- Humanities, including infrastructure projects: CLARIN ESFRI, DARIAH ESFRI, CESSDA ESFRI, DASISH, and Project Bamboo;
- European Neutron and Photon Facilities;
- Climate Science, including projects CEMS, ESGF and CMIP5, Metafor, IS-ENES, CORDEX, Exarch, Climate Data Exchange, GENESI-DEC.

After initially extracting the requirements specific to each community, FIM4R summarised the common requirements that emerged, assigning them two levels of priority: High and Medium. A list of the identified requirements grouped by priority level is presented below:

High Priority FIM Requirements

- **User friendliness**: The Federated AAI framework should provide simple and intuitive tools that are able to address the needs of users with different levels of ICT literacy.
- **Homeless users***:* Users without a federated institutional IdP should be supported. Such users include citizen scientists and researchers without formal association to research laboratories or universities.
- **Browser & non-browser based federated access**: FIM capabilities are required for a wide range of applications used by the communities, regardless of whether a web browser front-end is available or not.
- **Implementations based on open standards and sustainable with compatible licenses***:* Open and standards-based AAI technologies should be used by the different communities to allow for interoperability by means of suitable translation services.
- **Different Levels of Assurance with provenance**: Credentials issued under different policies and procedures will need to include the provenance of the level under which they were issued.

- **Authorisation under community and/or facility control**: Communities should be able to manage the assignment of attributes to their members for authorisation purposes.
- **Attributes must be able to cross national borders**: The different data protection policies need to be considered for flows of identity attributes between countries.

Medium Priority FIM Requirements

- **Bridging communities**: Efficient identity attribute mapping mechanisms are required in the case of communities in different research fields, commercial sections and social groupings.
- **Multiple technologies with translators including dynamic issue of credentials**: Translators between different technologies will be required to allow credentials from one community to be used by services offered by another community and this translation will often need to take place in a dynamic fashion.
- **Well-defined semantically harmonised attributes**: Authorisation across the service providers residing in different communities requires open and/or standardised identity attributes.
- **Flexible and scalable IdP attribute release policy**: Given that bilateral negotiations between all SPs and all IdPs is not a scalable solution, a flexible negotiation mechanism is required to govern the release of identity attributes.
- **Attribute aggregation**: Appropriate mechanisms will be required to support the aggregation of attributes originating from different sources of authority, including federated IdPs and community-based attribute authorities.
- **Privacy and data protection to be addressed with community-wide individual identities**: The release of personal data should be managed in a way that satisfies all legal requirements for the protection of user privacy.

FIM4R also analyses several operational aspects that are important for the adoption of federated identity management in the services of a production infrastructure. These aspects include risk analysis, traceability, incident response and easy integration with existing SPs.

FIM4R spawned the Federated Identity Management Interest Group (FIMig) at RDA (Research Data Alliance) to align their requirements with the works undertaken in RDA and to increase the co-operation with Non-European activities that had limited representation in the FIM4R workshops.

4.1.2 TERENA AAA Study

In 2012, TERENA (now GÉANT Association) published a study[2] on AAI Platforms for Scientific Data in Europe. This study was led by TERENA and carried out by a consortium composed of four partners and a number of external experts.

[2] https://www.terena.org/publications/files/2012-AAA-Study-report-final.pdf

The study was commissioned to address the recommendations of the High-Level Expert Group on Scientific Data that indicated that an authentication and authorisation system should be set up by integrating existing AAA infrastructures in order to allow distributed and collaborative AAA for scientific data.

The goal of the study was to evaluate the feasibility of delivering an integrated Authentication and Authorisation (and, possibly, accounting) Infrastructure (AAI) to help the emergence of a robust platform (Scientific Data Infrastructure (SDI) for access to and preservation of scientific information.

The targeted actors in the study were the research and education communities, information service providers (data centres, libraries) and e-Infrastructure providers.

The output of the study consisted of a set of recommendations (of technical, policy, funding and legal nature) for the delivery of an integrated AAI to be used for SDI. Some of these recommendations confirmed the results of the FIM4R study.

The recommendations highlight the following priorities:

- The general assumption confirmed by this study is that an AAI for SDI should be built on standard technologies, using mechanisms to translate between various authentication and authorisation technologies, and that federated access plays an important role;
- To fully benefit from federated access, more funding is needed to improve the reach of national identity federations in research and education;
- Further research is needed to enhance authorisation and accounting mechanisms;
- A common policy and trust framework for identity management is needed, as well as clarity on data protection laws – these should be coordinated at a European level.

The tables below provide a more detailed description of these technical and policy recommendations, many of which confirmed the findings of the FIM4R document.

Technical Recommendation	Action required	Main Stakeholder(s)
Support the use and standardisation of federated technologies for network, service and application access across Europe	Specific support should be given to inter-federation to achieve cross-disciplinary and cross-boundary requirements and to create a common, but distributed AAI for research and education	e-Infrastructures, national federations and international research collaborations.
Enhance existing AAIs to address the demand of the research communities to access different type of services in a secure way.	a. Enable federated AAI for mobile access and mobile devices. b. Support for federated access for non-web application; c. Develop security translation services to enable the interoperability of different AAIs; d. Provide guest identity providers for users that cannot rely on institutional IdPs;	National identity federations, eduGAIN and international research collaborations.

Technical Recommendation	Action required	Main Stakeholder(s)
	e. Enable the support of persistent identifiers f. Develop tools to allow for effective accounting across the highly distributed, heterogeneous infrastructures envisaged for global research data. g. Support social identities and groups as both identity providers and attribute providers for the SDI.	
Enhance authorisation in inter-federation scenarios by providing support for distributed attribute management	Enable identity federations to consume attributes provided by third trusted parties.	National identity federations, eduGAIN and international research collaborations.
Phase out IP-based authentication in favour of federated access	Provide support for those institutions relying on IP-based authentication to migrate to federated access	National identity federations, service providers and funding bodies

Table 1: Summary of Technical Requirements gathered from the AAA Study

Policy Recommendations	Action Required	Main Stakeholder(s)
Facilitate the development of a common policy and trust framework for Identity Management that involves, identity federations, research communities, libraries and data centres.	Use existing frameworks to coordinate the creation of best practices.	e-IRG, REFEDS, IGTF, ESFRI, e-Infrastructures, LIBER and libraries
To expand the coverage of national federations	Allocate national and international funding to train communities to join federations.	National funding bodies and EC
Implement scalable policy negotiation mechanisms	Define ways to simplify the negotiations of service agreements	REFEDS, eduGAIN, national federations
Identity federations to harmonise their policies	Define guidelines to harmonise policy federations	REFEDS, national federations

Policy Recommendations	Action Required	Main Stakeholder(s)
Lower the entry level for using a AAI	Consider ways to offer ready to use solutions that hide technical complexity from the users	e-Infrastructure (AAI) providers and national federations

Table 2: Policy Requirements gathered from the AAA Study

4.2 AARC Internal Survey

In addition to the information gathered from external sources, AARC conducted a survey of the communities directly involved in the AARC project with the aim of verifying their requirements to date and determining which issues have been addressed since the FIM4R and TERENA AAAI studies were carried out.

The Survey included ten pan-European Research Communities and Resource Providers. The statistics presented in this section are based on the responses collected.

Out of 10 responders, 8 indicated their community currently had some sort of AAI solution in place. These responders were then asked to rate the level of user experience within their community with their existing AAI solutions (where “0” indicated no experience whatsoever and “3” indicated good experience), to which the average rating response given was “2”. Investigating this question in more detail, it emerged that the communities that used some kind of web-based solution, including FIM, had given a positive usability rating. Certificate-based access to resources, on the other hand, was considered not to be user friendly and to be excessively bureaucratic.

4.2.1 FIM benefit and barriers

100% of the surveyed communities responded in the affirmative when asked whether they saw benefits in Federated Access. Clearly, FIM is viewed as a way forward for enabling access to shared resources. However, several impediments were also identified. The lack of adequate information about FIM and the need to improve it was noted as a factor by 75% of respondents. Other barriers were also highlighted, as shown in Figure 1: Perceived barriers to joining a federation. Lack of funding and the excessive bureaucracy when joining a federation were noted as the main barriers, followed by the lack of clarity on benefits within the organisation.

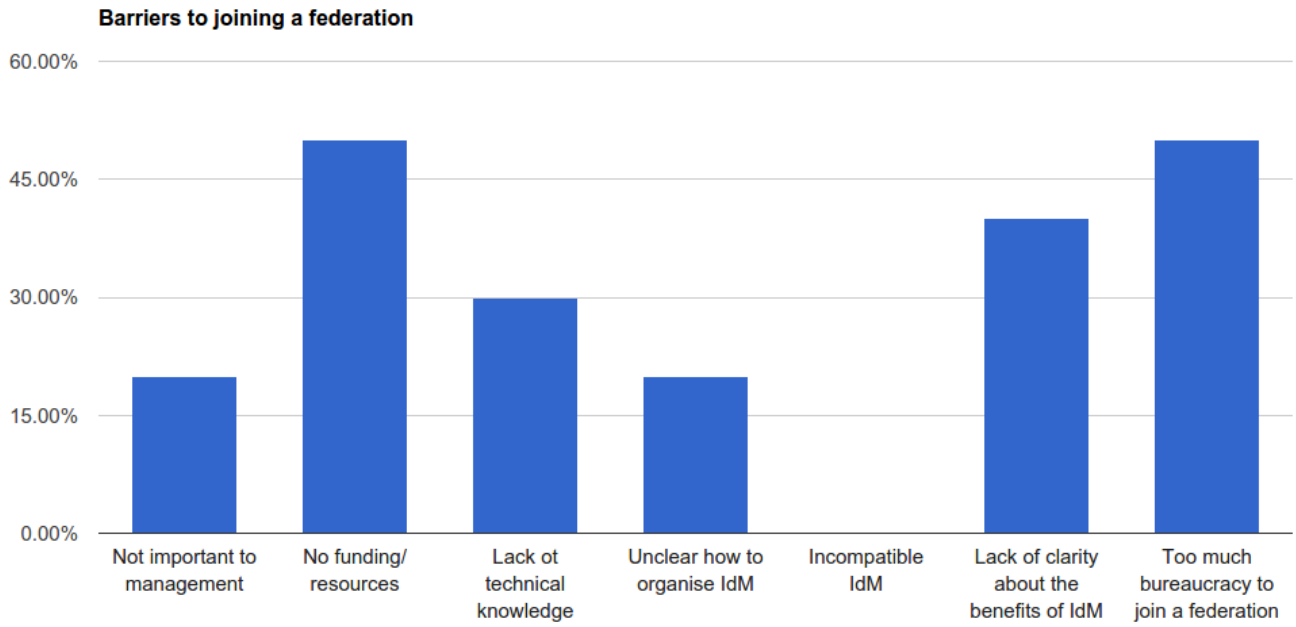

Figure 1: Perceived barriers to joining a federation

The survey went on to investigate how the level of penetration of AAI solutions in the community could be improved. Respondents were asked to indicate what kind of information should be provided to facilitate this. Figure 2 shows their responses, with "Online resources" being most often mentioned followed by “Materials for management and decision makers”. The findings of the survey have been passed on to the AARC NA2 activity.

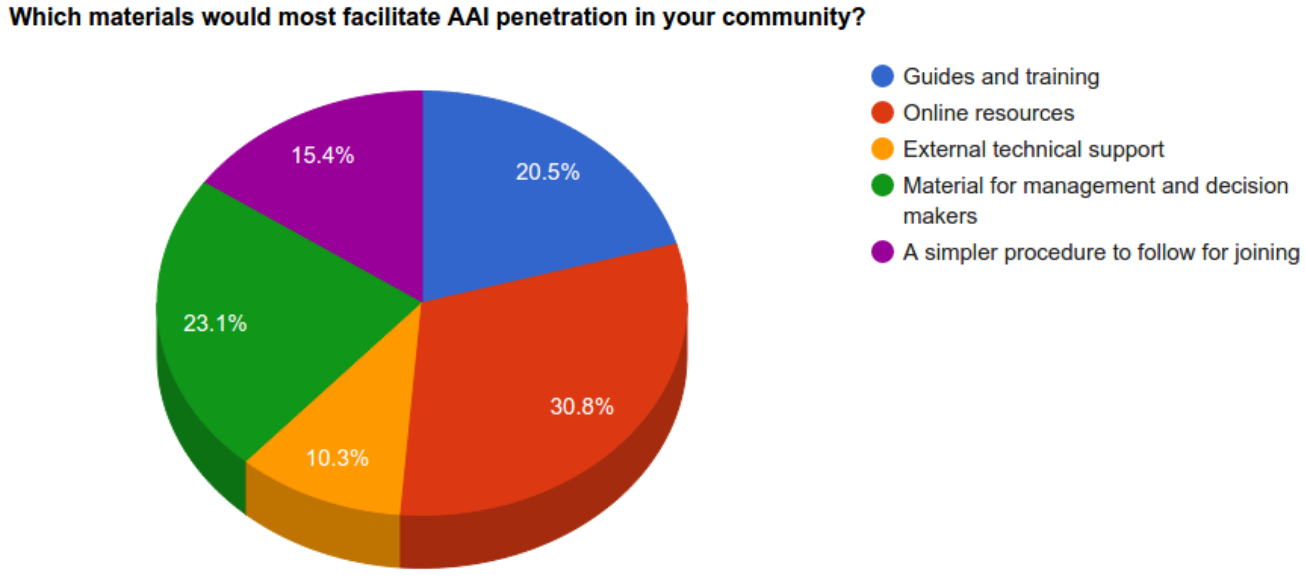

Figure 2: Materials needed to facilitate AAI penetration in your community

4.2.2 FIM technology

The survey next focused on the technical solutions already adopted by communities. Figure 3 shows responses with regard to protocols in use. From the survey it is clear that SAML2-based solutions are already popular. Several X509 and OAuth based solutions are in place or being deployed (especially for OAuth).

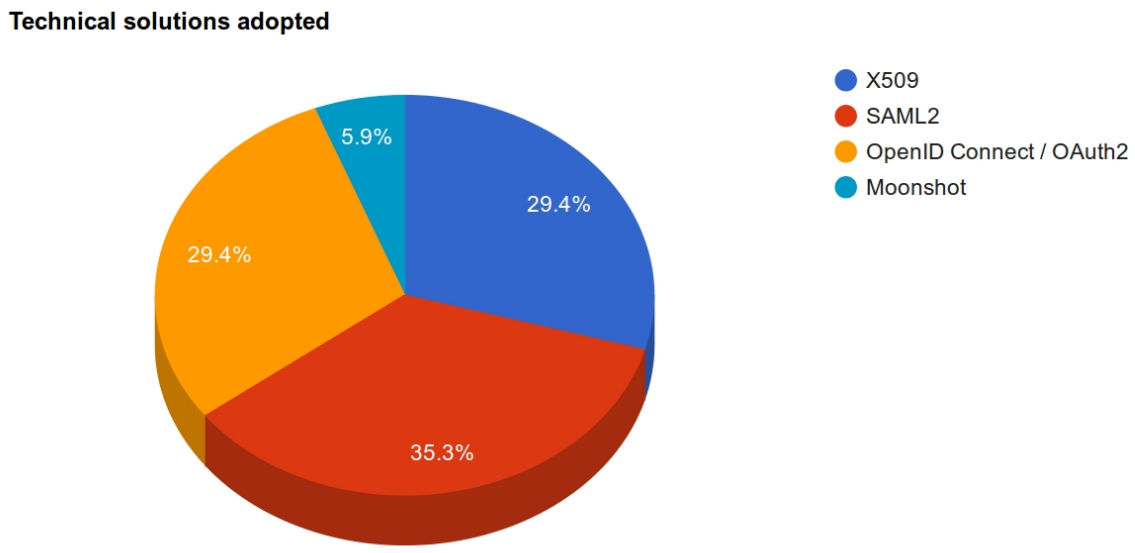

Figure 3: Technical solution adopted

The scale of deployments varied considerably between the participants of the survey. Table 3 shows the reported number of users and connected identity providers. Zeroes indicate there was no production level service in use, whereas a dash shows that the communities in question provided no data on connected IdPs and users. In some of the responses it is not clear whether the numbers refer to actual or potential IdPs and users. This will be clarified in the interviews that are going to follow in the period until the end of the first project year.

Community	Number of IdPs	Number of users
BioVEL	0	-
Clarin	1000	100000
MoBrain (EGI Competence Centre)	-	-
D4Science	1	-
DARIAH	30	3000
EISCAT	0	0

EUDAT	0	0
FMI	98	1600
PSNC	2	750
Umbrella	-	700

Table 3: Users and IdPs in AARC communities

4.2.3 Community requirements

The respondents were also asked to indicate their high-level requirements for Federated AAI. Some common elements were evaluated, including types of Identity Providers, preferred authentication technology, and other high-level requirements.

Figure 4 shows the Identity Providers used by respondents, with home institutions clearly preferred for authentication, though it is also noted that many VO participants still remain outside of Academia.

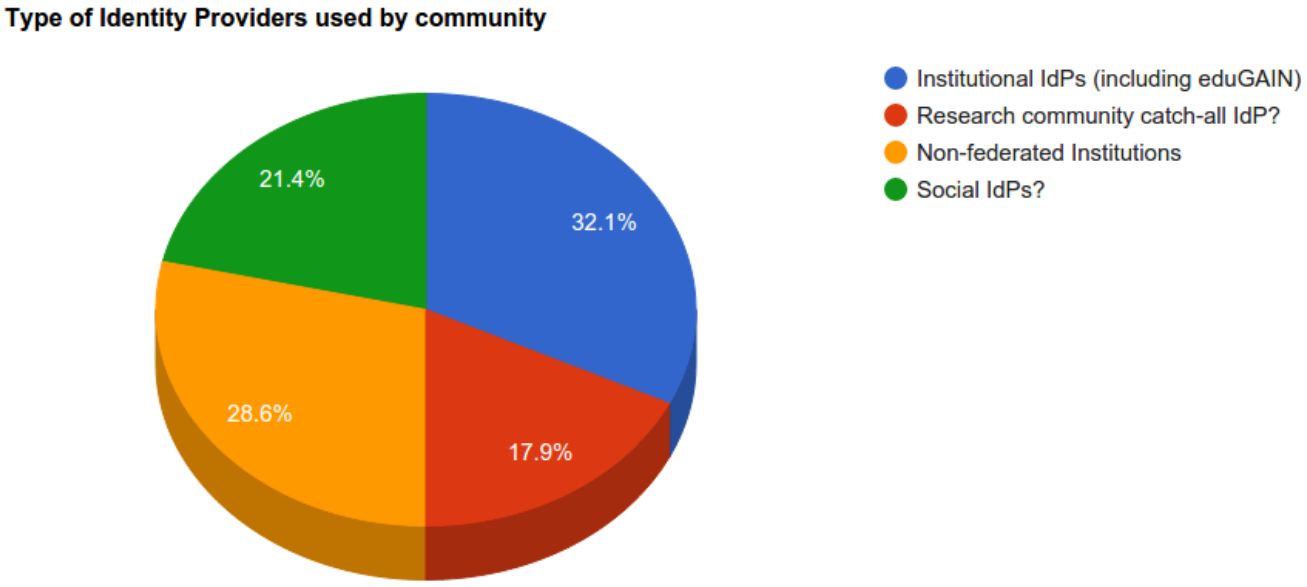

Figure 4: Type of Identity Providers used by community

When communities were asked their preferred Authentication method, web-based and non-web-based authentication scored equally, as shown in Figure 5. This clearly confirms that web SSO alone will not solve the AAI challenge for VOs.

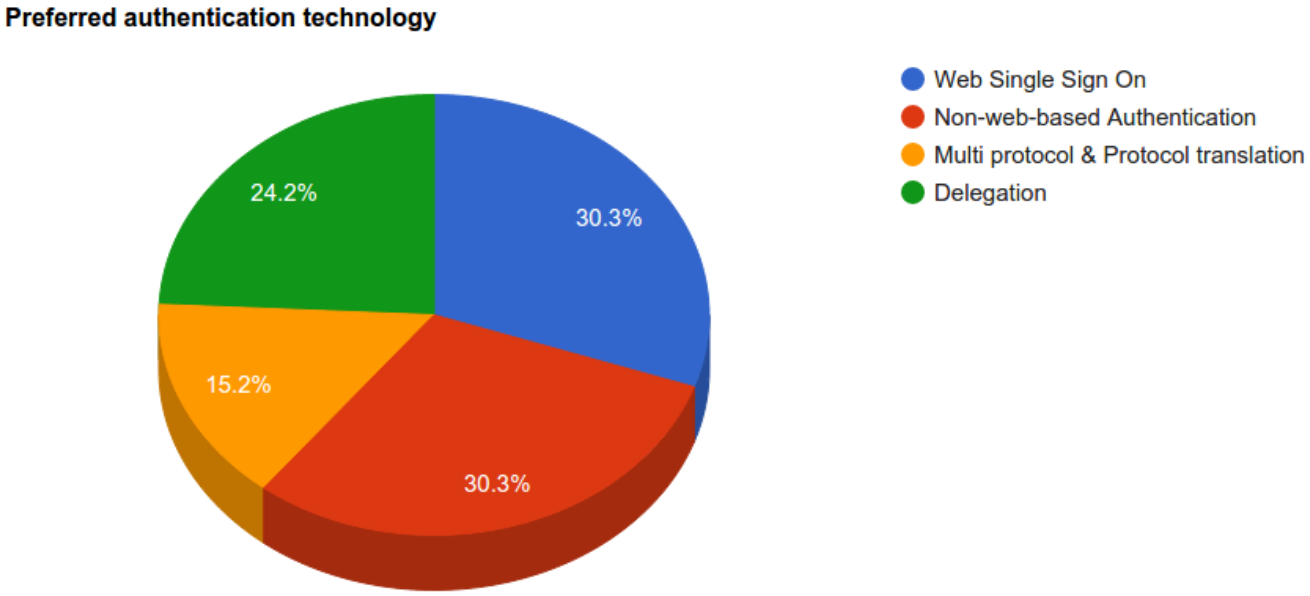

Figure 5: Preferred authentication technology

Other high-level requirements we also queried, including the need for scalable IdP attribute release, the need for persistent identifiers, support for different levels of assurance and the need for community-level authorization. Communities were also asked if currently the identity federations provide sufficient coverage in terms of users. Results are presented in Figure 6.

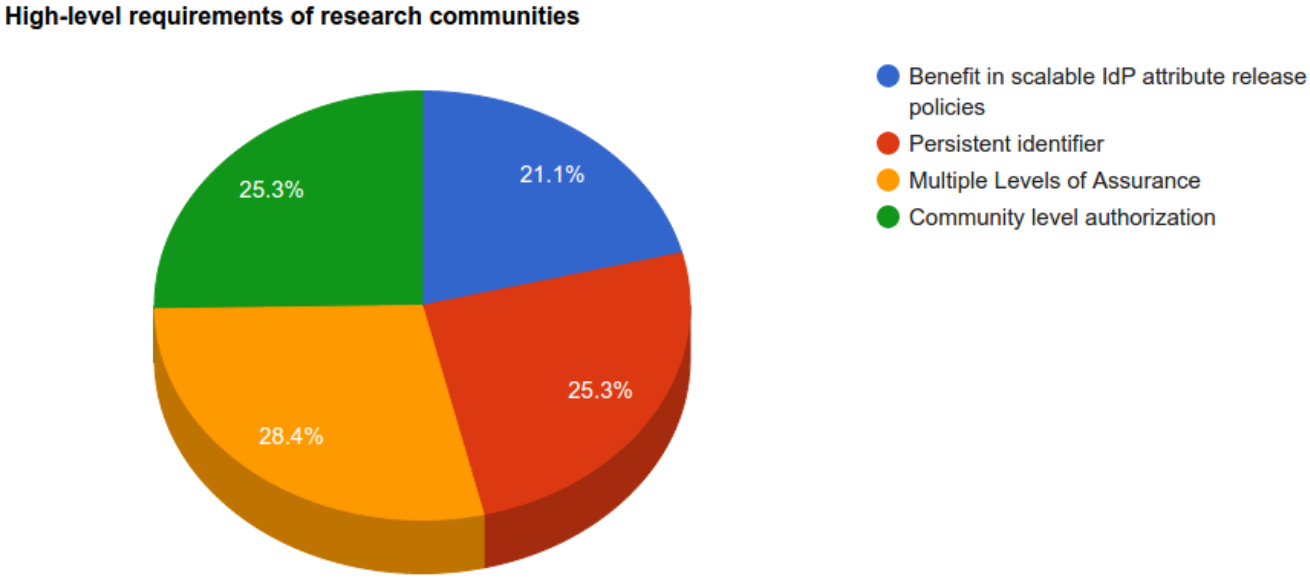

Figure 6: High-level requirements of research communities.

Persistent identifiers, multiple levels of assurance and the ability to perform community-level authorization were indicated as important by most communities. While (lack of) attribute release was seen as an impediment, some also considered the situation to be workable.

Most communities reported that coverage from Identity federations for their collaboration is poor. Less than half of the communities reported that this coverage is sufficient for their activities. One community

noted, on the success of eduGAIN: "In eduGAIN, countries with Opt-In are not well covered and are not sustainable in this respect. Moreover, 6 countries are behind > 1000 IdPs in eduGAIN (July, 2015) leaving ca 300 to the rest of the world, so it would be nice to see IdP coverage in eduGAIN improve."

4.3 AARC interviews

In order to further understand AAI requirements, a series of interviews with selected user communities was conducted in September 2015 via teleconference. These interviews provided room to discuss and gain a much deeper insight into the communities' AAI needs and issues. In view of the usefulness of these results, more interviews are scheduled to take place up to the end of JRA1.1. A summary of the findings from the interviews is presented hereafter.

4.3.1 EGI

EGI is a highly distributed, multi-disciplinary resource infrastructure, integrating more than 300 resource centres (service providers) and almost 20,000 users grouped in 200 user communities called Virtual Organizations (VO). Currently, authentication and authorisation within EGI is enabled through a X.509-based Public Key Infrastructure (PKIX), based on the Interoperable Global Trust Federation (IGTF) and EUGridPMA Certification Authorities federation.

The EGI requirements have been discussed with Peter Solagna, Senior Operations Manager at EGI.eu.

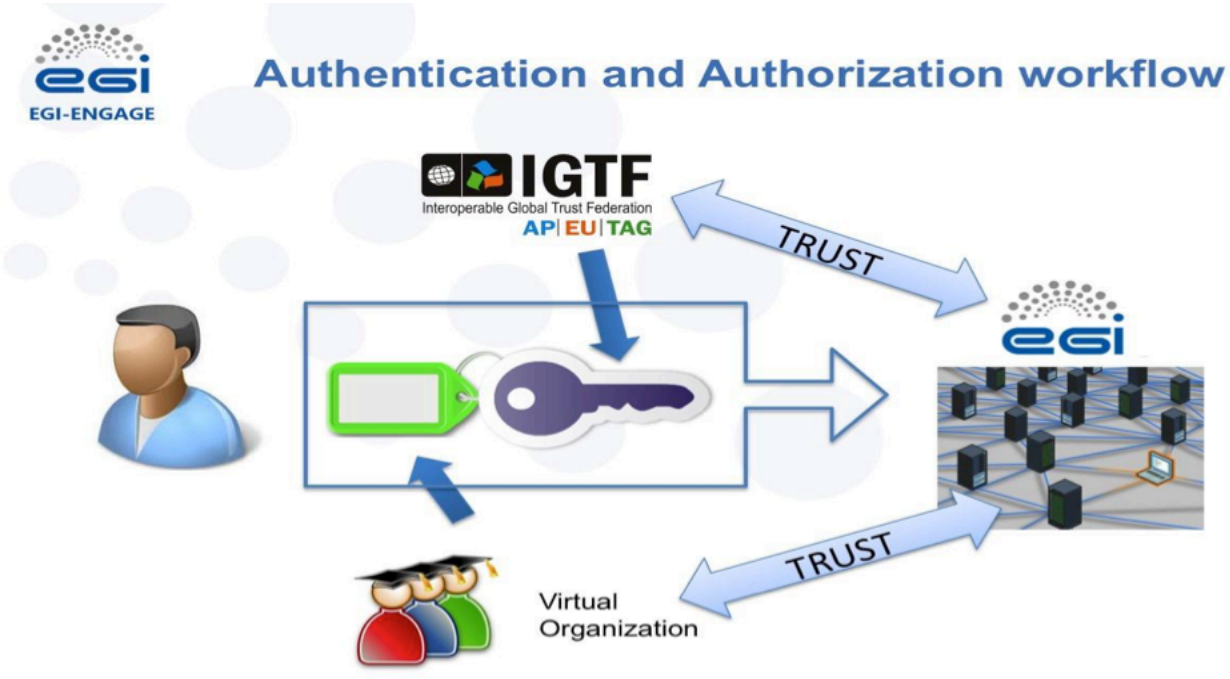

Figure 7: Current EGI authentication and authorization workflow

The process to access EGI services, as illustrated in Figure 6, is described below.

The user, shown on the left, obtains a personal certificate from a certification authority recognised by EGI, then adds the information about their VO to the certificate, generating a X.509 proxy. To be able to do this, the user must be authorised by the VO manager (who can be the Principal Investigator of the collaboration) who controls the membership of the VO they are managing.

All the needed information (user identity, VO membership, additional roles and capabilities within their VO) is shipped to the EGI services that use this information to authorize access to the services. Generally, access to services is regulated by VO membership: a service provider supports a number of VOs and users can search for the services enabled for their VO. Finer-grained user authorization is possible but not often applied in view of the scale of the infrastructure. The AuthN/AuthZ process is based on the trust between the CAs' federation and the service providers' federation, as well as the trust between the service providers and the user communities (VO).

The key feature for EGI is collaboration. The scale of EGI does not allow having a single entity responsible for the authentication and authorization of the users. Service providers, user communities and identity providers must be able to work together, contributing to the operations of the AAI workflow, to enable the users to access EGI services.

X.509 technology, and in particular X.509 proxy certificates, enables some of the most important capabilities for EGI: scalability, command line access and delegation. The user actions are usually initiated by submitting a request with an attached X.509 proxy certificate. This means that, for example in case of bulk submissions of thousands of computing tasks, the services do not need to contact the IdP or the attribute authority of the users thousands of times, since all the needed information is contained in the proxy certificate. The X.509 credentials work with no issues with command line commands, and the proxy certificate implements a form of delegation (impersonation) that allows a service to perform a defined set of actions on behalf of the user.

From a technical point of view, X.509 authentication has proven to be scalable and to work for almost any use case. However, new user communities prefer to use other technologies for authentication, for example username/password-based authentication.

There is a capillary network of certification authorities and registration authorities, distributed among the EGI partners, which can be contacted by users to obtain a certificate. EGI runs a catch-all CA to support users who – for any reason – cannot access another existing CA.

To bridge different authentication technologies with X.509, the EGI partners and the user communities are deploying science gateways and portals where users can authenticate with username/password, and access the resources through web-based tools and interfaces. The portals are then generating short-lived X.509 credentials that are used to access resources.

The most common mechanism used to bridge between IdPs and X.509 are the robot certificates, which can generate programmatically short-lived X.509 proxy certificates that then can be used to access EGI services. One of the drawbacks of this solution is that the real user identity is hidden behind the robot certificate. To partially address this issue, EGI is implementing an extension of the X.509 proxy certificate that contains an ID that can identify the user if needed. This is particularly useful for accounting purposes and, at the same time, improves the overall security of the implementation.

EGI critical requirements:

1. **Single sign on**. Users should be enabled to use their institutional credentials to access EGI services. One barrier for new users is that they have to obtain a new credential to access the e-infrastructure. In some cases, this is just an inconvenience, yet another credential to manage, but for some users – those outside institutions or the major IdP federations – it may be not

possible to obtain such a credential. User friendliness is of course a major feature for any SSO capability.

2. **Community attributes-based authorization**. Community-based authorization has been implemented in EGI from the beginning, and is at the basis of the collaborative nature of EGI. It is fundamental for EGI that every AAI technology and architecture enables the communities to manage the capabilities and the roles of their users, and to let these attributes be used by the services to regulate the authorization. Given the scale of the EGI service, providers cannot implement per-user authorization, but must authorize a user based on the attributes associated to that user.

3. **Non-web access**. EGI services are accessible natively by command line clients, implementing the services' standard APIs. While there are several options for users to use web-tools to access the resources, authentication must consider both web and non-web access scenarios.

4. **Delegation**. EGI services allow complex workflows, and users often submit thousands of computational tasks that need to access other resources (e.g. storage) on their behalf. The authentication technology must support delegation, either natively or through credential translation to another technology. Without delegation EGI users cannot get the full potential of the services.

5. **Scalable policies for attribute release.** EGI is a highly distributed infrastructure, with hundreds of service providers, hundreds of communities, and tens of thousands of users. In this scenario, it is critical that the policies and the procedures are scalable with the number of actors involved. Attributes are important to reduce the effort for user management falling on the communities or the service providers. If trusted IDPs can release adequate information to the service providers, credentials can be used to access the most complex workflows in the infrastructure without the need for additional vetting of the user, even in a scenario where the IdP releases a minimal set of attributes, policies must scale. For example, services must be able to store and share with other services the unique identifier of the user provided by the IdP. Service provider federations should be seen by the IdPs as trusted entities, and policies, once agreed with the federation, should be valid for all the service providers within the federation.

EGI non-critical requirements:

1. **PID for users**. EGI foresees the need for a user unique and persistent identifier. One use case is to map multiple credentials to a single user. A second use case, and in particular for an EGI-specific unique identifier, is to share authorization assertions between EGI services, which may not be possible when using an IdP-provided user UID.

2. **LoA management**. EGI supports a very diverse set of use cases. Open data is a typical use case where a very large community of users can access a data set, but there is need for a lightweight authentication, for example to account for the number of actual users using a service. In this example, EGI needs to enable users without the need for 'expensive' high assurance credentials. Clearly service providers must be able to extract information about the LoA from the attributes associated to the user identity. LoA definitions should be standard and simple, so as not to over-complicate the service provider's decision as to whether or not to allow the user task.

3. **Credential translation services.** While the plan is to push forward with the direct support of federated identity technologies for the central tools and the new service types, some services, for example those offering HTC resource, will likely continue to use X509 technologies,

therefore credentials translation capabilities are necessary to allow users with federated identity credentials to access the full set of EGI services.

How EGI AAI will evolve in the future.

EGI has strong requirements for extending the credential types that can be used to access services, in particular from new communities. Some communities also asked for basic AAI tools not necessarily related to resources access, but to be used for their internal workflows.

EGI is eager to implement and use the outcomes of the AARC project, and at the same time to implement federated AAI where necessary as soon as possible.

The current plan for EGI is to provide an IdP proxy to integrate external IdPs with the EGI services. The IdP proxy will provide the following capabilities:

- Integrate new heterogeneous IdP. EGI will integrate in the proxy-IdP the identity provider requested by the (new and existing) communities, making it easier to configure with the EGI service providers. Services will only have to consider the IdP proxy as the source of identity information.

- Aggregate attributes provided by community attribute management services. The aggregation makes it easier to configure additional attribute authorities, and at the same time allows EGI to certify authoritative sources of authorization data.

- Allow users to use multiple credentials to access EGI resources. The proxy IdP will associate a single UID to multiple credentials, either an EGI UID or an externally provided ibe, to allow services to identify the users if they use credentials.

As previously mentioned, it is a priority for EGI to provide AAI services that are aligned with the AARC architecture and, more in general, with the project recommendation.

4.3.2 ELIXIR

ELIXIR is a Pan-European infrastructure for data storage, management and distribution, which aims to support biological research. ELIXIR resources are available to all life science researchers. An ELIXIR service may comprise a number of individual services provided by more than one institution, potentially geographically dispersed across Europe. However, such a service should be perceived by the end-user as a single service, whereby authentication and authorisation take place only once.

The purpose of the AARC interview with ELIXIR representatives was thus to assess the ELIXIR service requirements for federated AAI, as well as to discuss barriers to its implementation and respective solutions. A list of the collected requirements follows:

- **Unique identity:** Each ELIXIR identity should represent a single natural person and should be uniquely identified by two independent identifiers:

    - a unique, permanent, opaque, non-reassignable ELIXIR identifier, and
    - a unique, human-readable, user-defined nickname that may change. In the case where the nickname is updated by the user, the old nickname should not be assigned to other users.

- **Up-to-date affiliation information:** Each ELIXIR identity can be associated with one or more affiliations, known as *home organisations*, such as research institutions or private companies. To describe a user's affiliation with their home organisation(s), it is recommended to use one of the predefined values, i.e. faculty, member, or affiliate. The up-to-dateness of the affiliation information should be guaranteed through a verification process recurring every 12 months.

- **User-managed identity attributes:** A user should be able to self-manage some of their attributes through a web-based UI. Depending on the attribute type, special update procedures should be supported (e.g., to update their email address, users need to demonstrate ownership of the address through a handshake protocol).

- **Multiple authentication providers:** A user should be able to link one or more external authentication providers to their ELIXIR identity. To link a new authentication provider, the user should first log into the ELIXIR AAI through an existing authentication provider and perform a subsequent login using the new authentication provider. The ELIXIR AAI must provide a web-based self-service management UI for this. The following authentication providers should be supported:

    - IdPs managed by the users' home organisations to which ELIXIR AAI connects via the eduGAIN interfederation service or otherwise. Users can use the authentication credentials managed by their home organisation to log in (SAML 2.0 technology);
    - IdPs managed by common social media providers, such as Google, Linkedin or ORCID. ELIXIR users can register a social identity to log into ELIXIR (typically, OAuth2 technology);
    - the ELIXIR authentication credential (password) managed by the ELIXIR AAI.

- **Level of Assurance:** To qualify as an authentication provider, the eduGAIN IdPs must provide sufficient LoA meeting the following requirements:

    - The accounts must belong to individual users;
    - The home organisations must have a standard identity-proofing mechanism for issuing user credentials;
    - The home organisations should either deactivate the account of a departing user or update their affiliation information accordingly.

- **Common attribute policy framework:** All participating entities in the AAI ecosystem (IdPs, AAs, SPs) should commit to a common policy framework for the processing of personal data. This framework should incorporate at least the GÉANT Data Protection Code of Conduct.

- **Step-up authentication:** The ELIXIR AAI should provide a step-up authentication service covering both *strong identity proofing* (face-to-face) and *strong authentication* (two-factor) at the time of login. For the latter case, the following second-factor authentication mechanisms are suggested:

    - SMS-OTP (one-time password delivered to the user’s registered cell phone number);
    - a smartphone application (e.g. Tiqr, http://tiqr.org/); and
    - a hardware token generating one-time passwords, (e.g. Yubikey, a token emulating a keyboard).

    In the future, the ELIXIR step-up authentication service should be able to integrate with external strong authentication services, such as government eID or eIDAS.

- **Groups and roles:** Each user can belong to one or more groups. A group member can have arbitrary roles in a given group, such as “member”, “owner”, “secretary” or “chair”. The AAI should provide a web-based service for managing ELIXIR group members and roles. The group owner needs to periodically confirm that the group is still active. Integration of the ELIXIR group management service with external group management systems (e.g. VOOT or SCIM technology) is also foreseen.

- **Distributed access control:** ELIXIR AAI requires a distributed interface for delivering access rights from the system component where they are granted and archived (Policy Decision Point - PDP) to the system component that enforces access control (Policy Enforcement Point - PEP). Revocation of access rights distributed across multiple PEPs should be supported in a timely fashion.

- **Browser & non-browser based federated access:** The ELIXIR services that require federated access can be either web-based or non-web-based, e.g., SSH, FTP, etc.

4.3.3 EUDAT

The B2ACCESS service is a bridge between technologies. It is based on Unity software with the addition of the CA Server from the Contrail project. Unity is an IdP/SP proxy, which also works as attribute provider, asking the users to add the missing attributes. In the current implementation of B2ACCESS, EUDAT can add extra attributes to the users, but the users have no control over the EUDAT-specified attributes. B2ACCESS also has support for social IdPs; the policy says that depending on what users authenticate with, they get assigned a different LoA that gives them different access rights. B2ACCESS will go in production by the end of September 2015. EUDAT has already registered B2ACCES with the DFN-AAI and the CLARIN SP Federation.

Account linking is not yet available in Unity. If this functionality becomes available from Unity, then EUDAT will consider it.

At the time Unity was chosen as the preferred IdM for EUDAT, it was the only solution that supported bridging between different authentication technologies. Unity is an open source middleware, developed in Poland and currently well supported. Unity is used in Unicore and by the Human Brain Project. Unity also gives control deciding what information is sent to different services, giving the possibility to delegate different sets of attributes to different services. Attribute release happens after user consent.

Regarding eduGAIN, EUDAT does not see any technical issues that might hinder its future integration. There are some aspects to clarify in the policy space, especially how to deal with the GÉANT Data Protection Code of Conduct, as B2ACCESS will effectively act as a proxy for a number of SPs, which will be hidden behind it. Jens Jensen is dealing with the Code of Conduct from the EUDAT side.

EUDAT is interested in having B2ACCESS join eduGAIN as an SP and not as an IdP. As mentioned earlier, B2ACCESS has joined the DFN-AAI federation, so it can already consume credentials coming from the German federation.

EUDAT sees a role for AARC to harmonise the attributes, but it was clarified that the semantic harmonisation is a lost battle and that AARC will not work to that extent. However, some work on the harmonisation of which attributes should be released by the IdPs and which could be offloaded to attribute providers is in scope. B2ACCESS was already designed with the notion that a core set of attributes will be retrieved from the institutional IdPs as well as a richer set of attributes from community-

specific attribute providers. Given the low number of attribute provider services operated by the communities, EUDAT could add and manage attributes in B2ACCESS for the communities it supports.

- **Requirement:** There should be harmonization of the attributes that will be used by the scientific communities for cross e-Infrastructure collaboration.

Would B2ACCESS work as attribute provider, for instance to offer attributes to access services outside EUDAT? In general, they are not so interested in this as there are also legal implications (in the Data Privacy Statement it is stated that they will not release user information to entities external to EUDAT), but they would be willing to state that a user is also an EUDAT user. Nevertheless, their requirement is to be able to collaborate with other e-Infrastructures, like PRACE and EGI, immediately. In order to circumvent obstacles such as the lack of available community-based Attribute Providers or globally unique user identifiers, EUDAT may also have to act as an Attribute/Identity Provider for cross-infrastructure collaboration. EUDAT does not want to become an IdP (though this functionality is there for practical reasons, they can offer pure EUDAT identities), therefore will not join eduGAIN in this capacity, but may have to operate B2ACCESS as an IdP offering some attributes for collaboration purposes, at least until a better solution is found.

- **Requirement:** Scientific Communities should be able to provide attributes for their members in an interoperable way. Due to the lack of community-based attribute providers, e-Infrastructures have to provide this functionality for the collaborating communities.

The plan for the immediate future is to integrate with two communities, INET and CLARIN. Planning further ahead, more communities and e-Infrastructures will join.

As for unique IDs, a truly unique Id for users would be desirable but is not planned, as EUDAT map the user authN with an EUDAT ID.

If B2ACCESS joined eduGAIN as an SP, it would trust all IdPs that come from eduGAIN. As for authentication, EUDATwill implement some kind of assurance framework based on the source of authentication (eduGAIN vs Social IDs). The real issue for EUDAT is the absence of an LoA framework for Attribute Providers. LF noted this work is in scope with AARC; clearly we should aim for the R&E community to agree on a common LoA framework and use that.

- **Requirement:** A Level of Assurance framework both for authentication and for attribute release.

Another point made by EUDAT was the need for harmonization of Data Protection/Privacy Statements for collaborating Service Providers. A set of templates for such policies would be of great help to Service Providers.

- **Requirement:** what type of privacy statement should a service use? AARC should offer best practice terms and conditions that service providers (operated in the R&E) should use.

4.3.4 GÉANT Project

4.3.4.1 *GN4-1 SA5 VOPaaS*

The GN4-1 project coordinated by GÉANT started a work area within SA5 with the purpose of offering infrastructure services to help Pan-European Virtual Organisations (VOs) leverage federated Authentication and Authorisation Infrastructures (AAI) for their collaborations. The proposed service portfolio GÉANT could offer for such VOs is referred to as Virtual Organisation Platform as a Service (VOPaaS).The project started in May 2015 and aims to deliver a first set of services by the end of the

GN4-1 project in May 2016. This section presents the results of a discussion with Ann Harding, Activity Leader of the GN4 SA5 activity.

To identify the requirements for the VOPaaS platform, the project conducted a survey among a number of small and large VOs with collaborations across Europe. The survey was conducted in Q4 of 2014. In addition, some VOs already had an operational infrastructure at the time of the investigation as part of the GN3plus project. For these VOs, both the requirements documents and the existing infrastructure were investigated in a desk study. Finally, the results of both the survey and the desk study were verified against existing requirements documents, such as the FIM4R paper.

Figure 8 shows an overview of the generalised requirements, sorted by occurrence over all investigated VOs combined. The Y axis represents the topics mentioned, while the X axis represents the frequency with which these were mentioned by various Virtual Organisations.

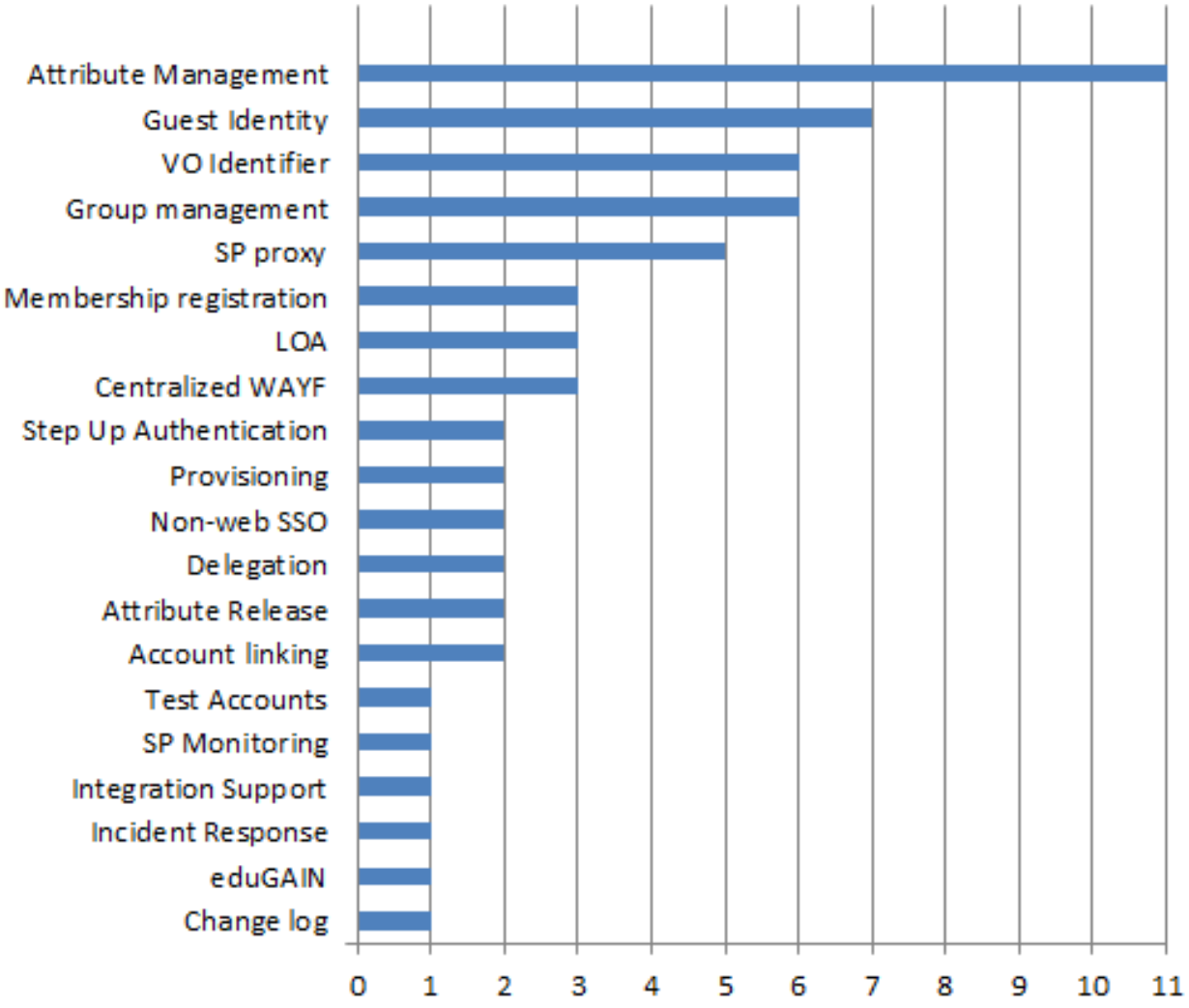

Figure 8: Generalized requirements for VOs, as reported in the GN44-1 SA5 VOPaaS Market Analysis.

With the requirements defined, a team of experts from seven NRENs, with considerable expertise in both developing infrastructures for collaboration as well as in directly supporting VOs, reviewed the combined requirements in a series of workshops. The result of the workshops yielded a market analysis, which lists in functional terms the most needed elements for a VOPaaS platform. Table 4 lists these requirements, grouped by “Basic” or “Advanced” Service offering.

VOPaaS aims to support both those VOs that are already aware of the capabilities of federated AAI and may already be able to operate such an infrastructure, and the 'long tail' of VOs which are just beginning to discover the benefits of federated AAI for collaborations. To this end, the VOPaaS platform aims to offer two sets of services, labelled “Basic” and “Advanced” where the former would support emerging use of AAI in collaborations and the latter would support more advanced use cases.

The table below shows the capabilities for both the Basic and Advanced offers.

Platform capability	Basic services	Advanced Services	Deployment tenancy	Data ownership
Persistent Identifier	y	Y	Multi	VOPaaS Operator
VO Membership management	y	Y	Multi	VOPaaS Operator
External ID provider	y	Y	Multi	VOPaaS Operator
Basic Groups	y	-	Multi	VOPaaS Operator
Basic Provisioning	y	-	Multi	VOPaaS Operator
Advanced Groups	-	Y	Single	VO operator
Attribute Management	-	Y	Single	VO operator
Advanced (de-)Provisioning	-	Y	Single	VO operator
SP proxy, attribute aggregation	-	Y	Multi	VO operator

Table 4: Platform capabilities in Basic and Advanced service scenarios as described in the GN4-1, SA5 Market Analysis

It is expected that the VOPaaS platform will work towards delivering the Basic services first, rapidly followed by the Advanced services. The VOPaaS market analysis also investigated the way the services could be deployed, and who should retain ownership of the (end user) data stored in the platform.

4.3.4.2 *GN4-1 Cloud Activity (WIP)*

The GN4-1 Service Activity Support to Clouds (SA7) enables NRENs to deliver cloud and mobile services to their communities with the right conditions of use. Through this activity, European NRENs collaborate to overcome challenges and to enable and facilitate higher education and research to adopt the cloud on a large scale. The primary objective of this activity is to ensure that GÉANT and the NRENs are optimally positioned and play an active role with respect to this rapidly developing paradigm, so that the benefits of the cloud can be fully realised and the attendant risks appropriately managed.

In September 2015, AARC JRA1 had the opportunity of discussing the main points described below with Panos Louridas, leader of the Cloud Integration task.

During the two years of the GN3plus project, and now in GN4-1, the SA7 activity has had the opportunity to interact with several commercial Cloud Service Providers (CSP) in order to make their services available on the GÉANT platform. CSPs were invited to register their services in an online Cloud Catalogue. In the Cloud Catalogue they provided specific details of their offerings by responding to a questionnaire drawn up by GÉANT that covered both technical and policy-related aspects of cloud services. The list of CSPs that are present in the Cloud Catalogue includes Google, Amazon, Microsoft, Box, CloudSigma, Netskope, Code42, Edu Zone, Advania, Ultimum Technologies, GRNET, and CARNet.

Federated authentication is one of the key assets of our community, so we need to ensure that this middleware is "cloud ready" by gathering and improving available components (from within the community) and encouraging vendors to do the same for their services. One of the initial goals of the activity was to consult with the CSPs and guide them through the process of making their service offerings available through eduGAIN. During this exercise we gained much experience and many insights relating to the challenges international service providers are facing in the process of joining eduGAIN.

- **Awareness about eduGAIN:** Although eduGAIN is widely understood within the academic and research communities, it is not clear to the commercial world what eduGAIN exactly is and what its actual benefits are. Even in the cases in which the CSPs were aware of eduGAIN, their understanding of it was incorrect. For example, in most of the cases in which eduGAIN was already known, the understanding was that eduGAIN is a Pan-European federation of academic Identity Providers (IdPs) and that by joining it the Service Provider (SP) would be able to automatically gain users from any of the academic institutions across Europe.

- **Need for development environment and guidelines:** In many cases the CSPs asked for guidance on how to implement a SAML SP interface and how to perform IdP service discovery. In all cases we were asked to provide a development environment against which the CSPs could test their implementation. A convenient solution for this was the OpenIdP service operated by Feide. The OpenIdP service allowed an SP to manually register its SAML endpoints and test its technical implementation. Unfortunately, this very useful service will be decommissioned by the end of 2015.

- **Registration procedure:** All the CSPs with which we engaged had a Pan-European scope for their operations. In addition to the issue mentioned earlier regarding the understanding of eduGAIN, most of the CSPs also expected that they would be registering to eduGAIN directly. The requirement to join one of the national federations in order to have their metadata published to eduGAIN is an important point of confusion. Furthermore, the fact that federations have different policies makes the registration process even more difficult. In GN4, this issue will be partially addressed as there is going to be a simple registration process relying on an existing

eduGAIN member federation that will act as the home federation for SPs that do not yet have a natural relationship with another eduGAIN federation.

- **Enabling eduGAIN SPs in eduGAIN IdPs:** When a CSP has properly registered with a federation and has had its metadata published in eduGAIN, it expects it will be automatically trusted by the IdPs that are participating in eduGAIN. Apparently this does not happen. In many cases the IdPs do not automatically trust the SPs they receive via the eduGAIN metadata service and as a result SPs still have to reach out to and negotiate with individual IdPs. This situation has a strong negative impact on the initial perceived value of eduGAIN.

- **Infrastructure services:** Most of the CPS were used to dealing with simple integration, in which they had to configure the trust relationship for one IdP per customer. By joining eduGAIN, they must be able to automatically handle a large number of IdPs and provide extra functionality, such as a Discovery Service. This has proven to be one of the biggest obstacles most of the CSPs faced and in most of the cases we had to provide a custom Discovery Service to them, or even connect them through an IdP/SP proxy. Having such services as Infrastructure Services provided centrally would be a very significant help to SPs (especially big international service providers who are not so flexible) to enable them to offer their services in eduGAIN.

- **Available Attributes**: The set of available attributes from the eduGAIN IdPs is too limited for delivering personalized services. Still, even this very limited set of attributes is not always available from the IdPs. One of the most basic requirements for delivering personalized services to the end users is the release of a globally unique identifier for each user, so that SPs can uniquely identify the users.

4.3.4.3 *GN3plus and GN4-1 SA5 Enabling Users*

The scope of the Enabling Users task is to be an expert partner for research projects that have a need for federated identity management in a global context and to carry out a selection of pilots in collaboration with these research communities to improve their ability to use federated identity. A related goal is thereby to increase the practical use of eduGAIN, the inter-federation service created by the GÈANT project.

The following considerations are the result of a discussion with Lukas Hämmerle, the Task Leader of the GN3plus and GN4-1 SA5 Enabling Users task, taken from the report delivered in December 2014 [GNSA52014].

The ideal scenario would be for GÉANT to be able to provide Pan-European research groups with a catch-all out-of-the-box solution which immediately addresses all their cross-border needs. However, the experience from pilots suggests that unfortunately this is often not possible. Although the requirements of different research groups at first glance may appear to be similar or at least related, the details and focus of the specific use cases often vary greatly. These differences are often what provides the greatest value to the end users of a particular service, and therefore care has to be taken to respect them. A possible promising alternative being considered involves additional services being layered on top of eduGAIN. Some of these, such as VO Platform as a Service or Affiliation checking services, are being developed with a view to operation at a GÉANT or federation level. The intent is to benefit from the experience of those federations that already offer these advanced services on a national level, but to scale the service to serve cross-border communities. In other cases, consultancy is recommended to continue to help research communities embed any highly specific niche requirements in their own systems under their own control. In this way, each level of the infrastructure preserves the trust relationship that is so important when delivering services to the end user.

For some research groups, a high Level of Assurance (LoA) within their limited group of users is fundamental (ELIXIR), while for others a wide coverage with a good attribute release is more important (DASISH); others still already have and require fairly advanced group management tools (DARIAH), while others may not need fine-grained group management at all (Umbrella). Applying all these features uniformly to the full infrastructure for the benefit of some would greatly increase the costs and complexity for others in the delivery chain and therefore damage the value proposition of eduGAIN as a general infrastructure. In particular, the costs of enhancing the service uniformly with advanced features can often fall in disproportionate measure on the campus IdP administrators, who typically have no clear access to research infrastructure funding to meet them. The business and cost recovery aspects of federated identity for eResearch therefore require further study.

A number of issues that are mentioned in the TERENA AAA Study and in the FIM4R paper were also confirmed by the experience of the Enabling Users team. These include:

- The need for stronger authentication and identity vetting (ELIXIR/CERN).
- Insufficient attribute release (DARIAH, CLARIN) by Identity Provider.
- The general lack of campus IdPs exported by federations into eduGAIN in the early years of the service.

As the AAA study also stated, no silver bullet solution or technology exists that could address all requirements or solve all current issues. However, promising progress is being made in these areas. The Enabling Users activity is gaining a greater understanding of practical needs in terms of assurance vs. federation capabilities. Progress towards attribute release is also being made in the deployment of the GÉANT Data Protection Code of Conduct (CoCo). Finally, coverage is improving, with the number of IdPs in eduGAIN doubling in one year alone. As several federations are considering moving to an opt-out rather than opt-in model, this number is likely to quickly increase further.

Some of the problems and requirements analysed cannot be addressed in the short term. For example, the issue of LoA requires the development of a common understanding of LoA tailored to the higher education sector, possibly also including LoA for organisations (and their identity vetting processes), plus LoA for individual users with a different LoA to that of their organisations. Based on the Enabling Users experience, the organisational and business case aspects of delivering such a solution will be examined during future GÉANT projects.

Many research communities also requested non-web authentication. Moonshot is one potential solution to provide this, but is still in the early stages and not widely deployed, and also offers a unifying solution for Single Sign-On that is too large for what some of the simpler use cases require. SAML ECP is an alternative method for supporting non-Web SSO and in some areas could lead to faster results. This option has been identified as requiring a deployment strategy within the eduGAIN environment.

As regards FIM, although it ultimately does reduce user management overhead and risks, it is a specialist field contained within a much wider systems administration and delivery skillset, so that the relevant know-how within research communities is often limited. Manpower dedicated to working on FIM within research projects is generally scarce, which made collaboration and planning sometimes difficult. Drawing up project plans with research communities is helpful in this respect, but the core purpose of the eResearch group is often a priority that overrides other concerns, and although FIM is important to research communities, their primary aim obviously remains the research itself. The ongoing availability of expertise and consultancy from the federation community on a national and Pan-European basis is therefore critical to prevent wasting limited resources on reinventing the wheel or their being confused by the various options available.

Our experience is that the assistance, expertise and services that Enabling Users provided was generally very welcome, needed and appreciated by the research communities. Sufficient use cases have been identified for Enabling Users to continue work on collaborative pilots in future projects, complementary to the AARC project. It is also recommended that the provision of basic eduGAIN consultancy on request to research communities should be continued, as it often results in positive and fast results given a relatively small effort. Consultancy also enables research communities to become familiar with the topic without too much upfront effort and investment being required on their part prior to their making work plans. Receiving basic consultancy from the Enabling Users team could therefore be the first step on the way to developing more advanced FIM use cases later on. In addition, the greater number of results generated by collaborative pilots increases the ability of the consultancy team to help more research communities more quickly.

4.3.5 Dutch Consortium of the National and the University Libraries

In 2014, the UKB, a consortium of thirteen university libraries in The Netherlands, carried out a thorough investigation to identify the pros and cons of introducing federated identity management for access to library resources. In September 2015, AARC JRA1 had the opportunity to have an in-depth discussion with Paul van Dijk from SURFnet about the main findings from this investigation in relation to the current situation.

To date, IP-based access control is regarded as the standard to obtain access to library content, but this method involves a number of limitations, including the considerable administrative effort required to maintain correct IP ranges and lack of support for home use by students and staff. In this light, moving to federated access management seems to be an obvious choice, as more and more providers of licensed material (e.g. ScienceDirect, EBSCO host, OCLC PiCarta, Project MUSE, Springerlink and, recently, KluwerNavigator) are supporting federated access, which is already supported by all universities in The Netherlands. By implementing federated access, IP address administration becomes unnecessary, and “access anywhere” (at the campus, at home, on the road) is available. At the same time, features such as SSO access and personalisation are made available. The UKB, however, identified a number of potential stumbling blocks to the implementation of federated access. These are, for example:

- The coverage of service providers with SAML support is still far below 100%. A mixed approach including IP-based access and federated access could be used but is considered confusing for users, as it means it is often unclear which resources are accessible through their educational account and which are not.
- So called “walk-by users”, such as citizen scientists, are not able to access academic content. With IP-based access, they are able to access content as long as they reside at the campus.
- SAML-based AuthN [itself] does not provide functionality to preserve the privacy of users.
- Single sign-off is not supported.
- Deep linking is not always supported (but depends on SP implementation),
- No seamless integration between providers.
- Lack of standardisation in the use of labels to login (e.g. “institutional account”, “shibboleth login”, “educational login” …).
- It is often confusing or time-consuming to deal with users that already have existing accounts at service providers. Not all providers provide functionality for account mapping.

Given these issues, the full adoption of federated access as a suitable alternative to IP-based access is not recommended by the UKB at present. However, it is willing to explore solutions for the abovementioned issues and collaborate with the AARC consortium.

5 Requirements Analysis

The comparisons made in the previous chapter between the requirements gathered in different projects, from different disciplines, and over a period of several years, reveal striking similarities.

Research communities, resource providers and research libraries alike are all seeing the benefits of Federated Identity Management for sharing resources with users across Europe and beyond. Equally, most struggle with internal and external factors to implement federated AAI in their communities. Lack of knowledge and time, and of a firm understanding of the benefits are general internal factors, whereas federation bureaucracy and lack of attribute release remain problematic as external factors. Nevertheless, several communities actively try to move forward, incorporating strategies to deal with new challenges as they arise. A sign of increased adoption is also seen in the rise of new areas of interest, including investigations into leveraging multiple Levels of Assurance and the activities around federating incident response handling.

It is fair to state that the findings of the AARC internal Survey firmly echo requirements as voiced in the other resources listed in this document. This is both good and bad news.

In the positive, this means that it is reasonable to assume the communities within the AARC project are a good representation of the e-Science communities throughout Europe. The work within AARC will therefore likely not only benefit communities participating in AARC directly, but will also be relevant for communities not participating.

That said, it is also clear that many areas still need significant improvements before federated AAI can become a catch-all out-of-the-box solution for Pan-European research groups which immediately addresses all their cross-border needs.

The requirements presented in section 4 have been extensively processed, i.e. homogenised, merged, filtered, extended and classified to produce a refined set of requirements, presented below. More specifically, each requirement is associated with a unique ID, title and short description. The type is also specified, along with the source(s) from which the requirement was extracted. It should be noted that the classification of requirements is based on the FURPS+ model [FURPS+], where the F letter stands for **Functional** requirements, while the rest of the letters further classify non-functional requirements into the following categories:

- **Usability:** Ease of use, effectiveness of training and documentation available
- **Reliability:** Frequency/severity of errors, ability to recover from failures
- **Performance:** Efficiency covering aspects such as speed, resource consumption, throughput, and response time
- **Serviceability/Supportability:** Technical support, ease of installation process

Finally, the “+” symbol is used to identify additional categories, such as Design, Implementation, Interface and Physical requirements

5.1 Architectural and Technical Requirements

ID	R1
Title	**User and Service Provider friendliness**
Description	The Federated AAI framework should provide simple and intuitive tools that are able to address the needs of users with different levels of ICT literacy and enable more Service Providers (commercial and non-commercial) to connect.
Type	Usability
Source(s)	FIM4R, EGI, AARC Survey, GN4 Cloud Activity

ID	R2
Title	**Homeless users**
Description	The Federated AAI framework should support users without a federated institutional IdP, such as citizen scientists and researchers without formal association to research laboratories or universities
Type	Functional
Source(s)	FIM4R, TERENAa AAA, GN4-1 SA5 VOPaaS

ID	R3
Title	**Different Levels of Assurance**
Description	Credentials issued under different policies and procedures should include the provenance of the level under which they were issued
Type	Functional
Source(s)	FIM4R, ELIXIR, EUDAT, EGI, AARC Survey, GN4-1 SA5 VOPaaS, GN3plus and GN4-1 SA5 Enabling Users

ID	R4
Title	**Community-based authorisation**

Description	The Federated AAI framework should enable communities to manage the assignment of attributes to their members for authorisation purposes
Type	Functional
Source(s)	FIM4R, EUDAT, GN4-1 SA5 VOPaaS

ID	R5
Title	**Flexible and scalable attribute release policies**
Description	Flexible negotiation mechanisms are required to govern the release of identity attributes
Type	Functional
Source(s)	FIM4R, EGI, AARC Survey

ID	R6
Title	**Attribute aggregation / Account linking**
Description	The Federated AAI framework should support the aggregation of identity attributes originating from different sources of authority, including federated IdPs and community-based attribute authorities
Type	Functional
Source(s)	FIM4R, TERENA AAA, EUDAT, EGI, GN4-1 SA5 VOPaaS

ID	R7
Title	**Federation solutions based on open and standards-based technologies**
Description	Open and standards-based AAI technologies should be used by the different communities to allow for interoperability by means of suitable translation services
Type	Implementation
Source(s)	FIM4R, TERENA AAA, EGI

ID	R8

Title	**Persistent user identifiers**
Description	The Federated AAI framework should reference the digital identities of users through long-lasting identifiers
Type	Design
Source(s)	TERENA AAA, AARC survey, EGI, ELIXIR, GN4-1 SA5 VOPaaS

ID	R9
Title	**Unique user identities**
Description	Each user should have a single digital identity to allow SPs to uniquely identify their users
Type	Design
Source(s)	AARC survey, EGI, ELIXIR, GN4-1 Cloud Activity

ID	R10
Title	**User-managed identity information**
Description	A user should be able to self-manage some of their attributes, e.g., through a web-based UI. Depending on the attribute type, update restrictions should be imposed.
Type	Usability
Source(s)	ELIXIR

ID	R11
Title	**Up-to-date identity information**
Description	The up-to-dateness of identity attributes should be guaranteed through an on-demand and/or recurring verification process
Type	Reliability
Source(s)	ELIXIR

ID	R12
Title	**User groups and roles**
Description	The Federated AAI framework should support the assignment of groups to users, as well as the assignment of roles to users within their groups
Type	Functional
Source(s)	ELIXIR, GN4-1 SA5 VOPaaS

ID	R13
Title	**Step-up authentication**
Description	The Federated AAI framework should provide an additional factor or procedure that validates a user's identity for high-risk transactions or according to policy rules
Type	Functional
Source(s)	ELIXIR

ID	R14
Title	**Browser & non-browser based federated access**
Description	The Federated AAI framework should provide federated access to both web-based and non-web-based services/applications
Type	Functional
Source(s)	FIM4R, TERENA AAA, EGI, ELIXIR, GN3plus and GN4-1 SA5 Enabling Users

ID	R15
Title	**Delegation**
Description	The Federated AAI framework should provide the capability for the users to delegate third parties, mostly computational tasks or services, to act on their behalf. This allows users to run thousands of actions in parallel without the need for interactive access, for example to save output data.
Type	Functional

Source(s)	FIM4R, ELIXIR, EGI, AARC Survey

ID	R16
Title	**Social media identities**
Description	The Federated AAI framework should support common social media providers, such as Google and LinkedIn, but also the researcher ID providers, such as ORCID, to act as authentication providers and/or attribute authorities
Type	Interface
Source(s)	TERENA AAA, AARC survey, ELIXIR

ID	R17
Title	**Integration with e-Government infrastructures**
Description	The Federated AAI framework should support broader cross-domain collaboration including e-Government infrastructures
Type	Interface
Source(s)	AARC survey, ELIXIR

ID	R18
Title	**Effective accounting**
Description	The Federated AAI framework should support effective accounting across distributed, heterogeneous data infrastructures
Type	Functional
Source(s)	TERENA AAA, ELIXIR

5.2 Policies and Best Practises

ID	R_P_1

Title	**Policy harmonisation**
Description	All participating entities in the AAI ecosystem (IdPs, AAs, SPs) should commit to a common policy framework regarding the processing of personal data. This framework should incorporate at least the GÉANT Data protection Code of Conduct.
Type	Supportability
Source(s)	ELIXIR, TERENA AAA, EUDAT, GN3plus and GN4-1 SA5 Enabling Users

ID	R_P_2
Title	**Federated incident report handling**
Description	A common procedure should be adopted for reporting security incidents that involve federations spreading across multiple administrative domains
Type	Supportability
Source(s)	FIM4R, AARC survey

ID	R_P_3
Title	**Sufficient attribute release**
Description	The set of attributes released to SPs should be extended, primarily, to allow consuming services to operate and, also, to allow for more advanced features, such as personalisation of services
Type	Supportability
Source(s)	FIM4R, AARC survey, EGI, GN4-1 Cloud Activity, GN3plusand GN4-1 SA5 Enabling Users

ID	R_P_4
Title	**Awareness about R&E federations**
Description	The benefits offered by R&E federations should be promoted to all stakeholders, such as (commercial) service providers and identity providers that have not joined a federation yet
Type	Usability

Source(s)	AARC survey, GN4-1 Cloud Activity

ID	R_P_5
Title	**Semantically harmonised identity attributes**
Description	A common set of vocabularies should be used by the different communities to denote identity attributes managed by identity providers and attribute authorities
Type	Supportability
Source(s)	FIM4R, EUDAT

ID	R_P_6
Title	**Simplified process for joining identity federations**
Description	The bureaucracy involved in joining identity federations should be reduced
Type	Usability
Source(s)	AARC survey

ID	R_P_7
Title	**Best practises for terms and conditions**
Description	AARC could offer guidelines for describing the terms and conditions that service providers (operated in the R&E) should use
Type	Serviceability/Supportability
Source(s)	EUDAT

6 Conclusions

In this document requirements were investigated from many sources in different projects, from different disciplines, and using documents written over a period of several years, including the surveys produced and collected in the first 6 months of the project. Out of these requirements, 25 high-level items were identified and categorised using the FURPS+[3] method. This classification of requirements is illustrated in Figure 8, where it can be seen that the majority of them are non-functional. This supports our initial assumption that, in addition to functional gaps, federated AAI requires significant efforts towards the definition of common policies covering the necessary legal and operational practices for all entities involved in the AAI ecosystem.

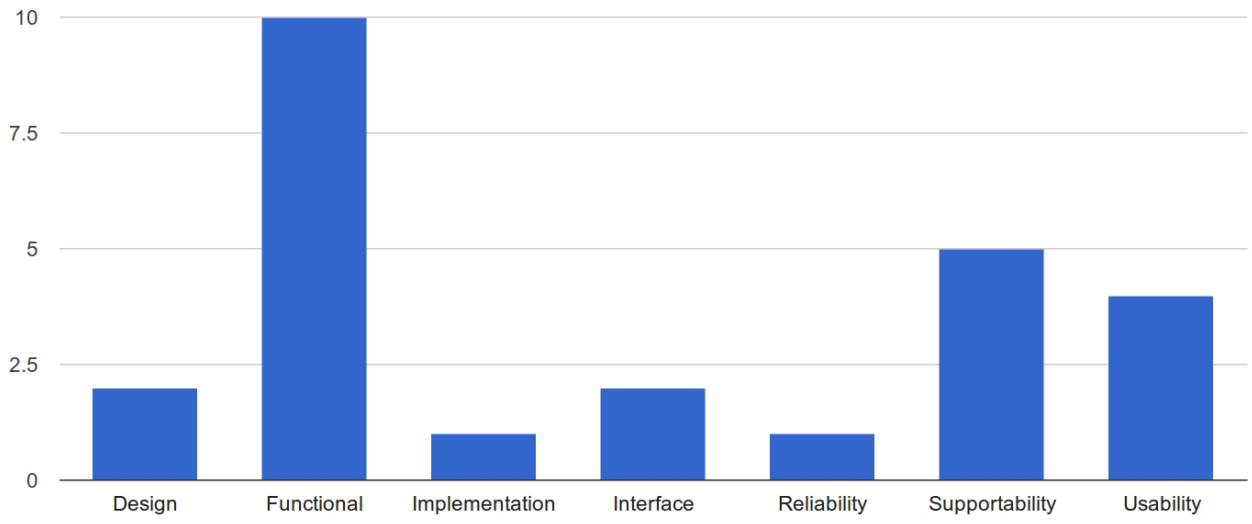

Figure 9: Distribution on the requirements presented in this document grouped by FURPS+ category.

[3] FURPS+ at IBM: http://www.ibm.com/developerworks/rational/library/4706.html#N100A7

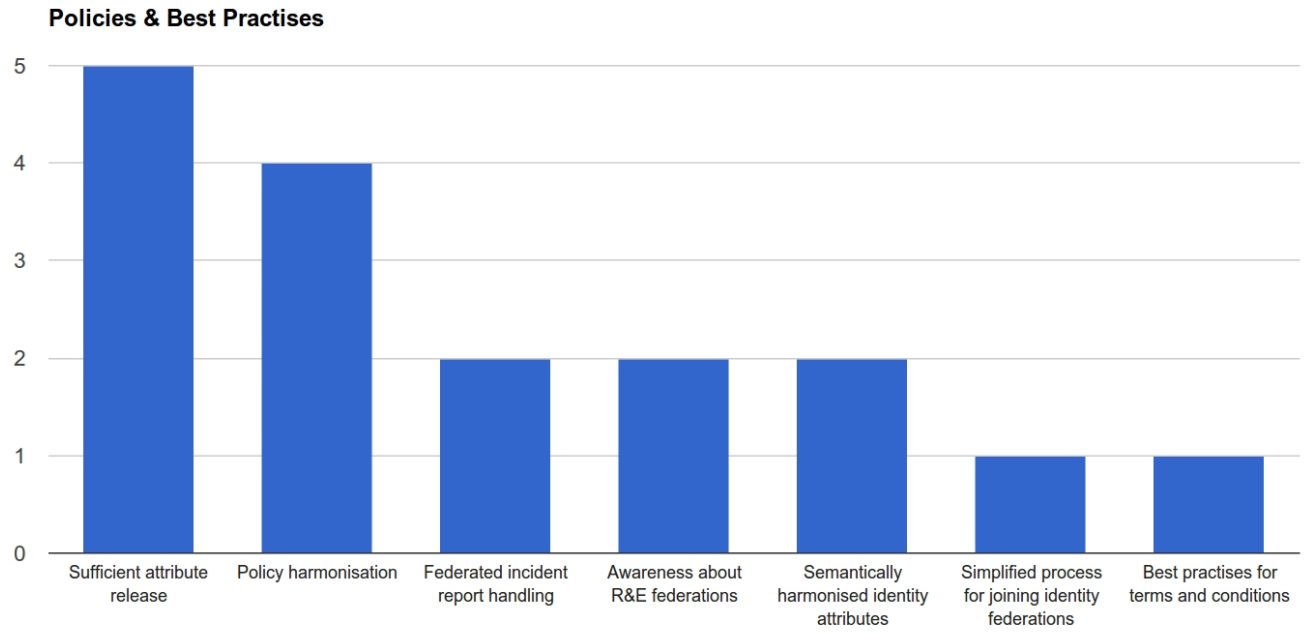

Figure 10: Number of communities interested in the policy and best practices topics.

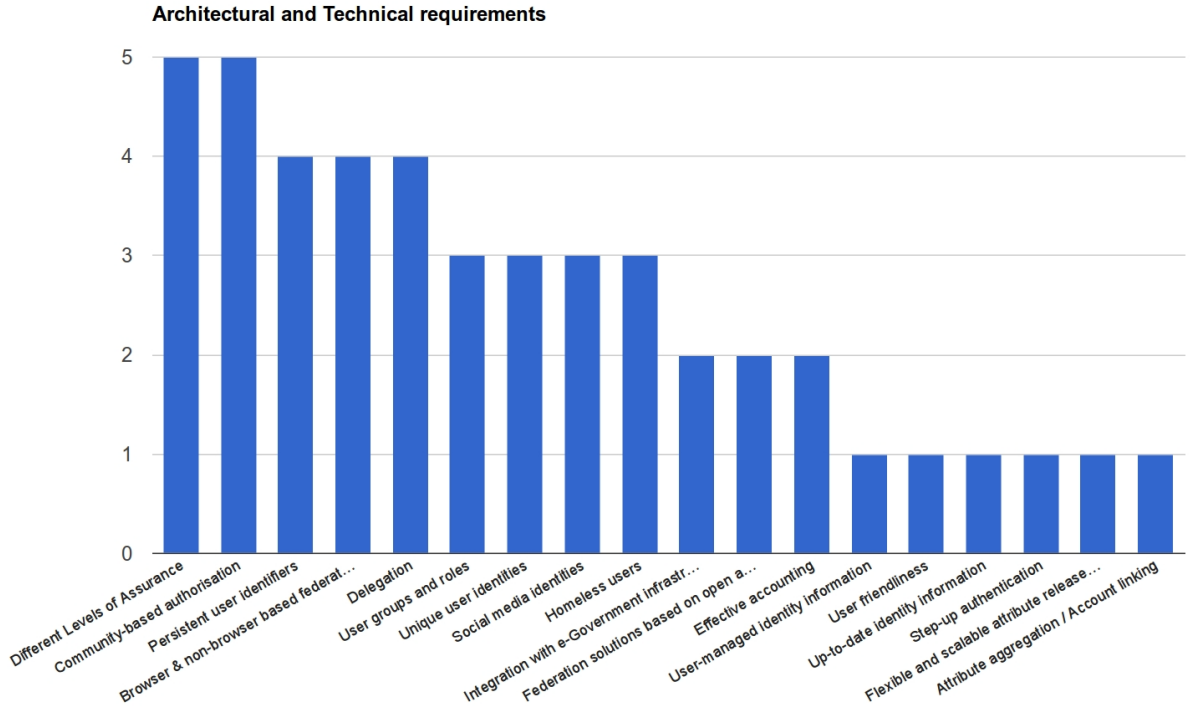

Figure 11: Number of communities expressing interest for the architectural and technical requirements.

The requirements related to policies and best practices are illustrated in Figure 10, where they are ranked based on the frequency of occurrence in various sources. In the figure, the information is presented in such a way as to highlight the possible areas that should be given the greatest attention within the AARC project. For example, sufficient attribute release and policy harmonization are clearly of interest to most of the communities participating in the requirement gathering process. Similarly, regarding the identified architectural and technical requirements, Figure 11 shows that support for different LoA and community-based authorization are required by the majority of the communities.

AARC will continue the discussion with the project stakeholders in order to understand and evaluate their requirements. Over the next months, until the end of the first project year (April 2016), we are going to work on prioritizing these requirements taking into account actual use cases, their horizontal applicability and their technical feasibility. The outcome of that work along with the information included in this document will be important inputs for the first internal draft version of the high-level AAI blueprint architecture, which is planned for the November 2015.

Complementary to this work, in MJRA1.1 "Existing AAI and available technologies for federated access", we will capture and analyse the tools and technologies which are available today for building federated AAIs, and will identify any technology gaps that might exist. This document will be available at the end of 2015.

Appendix A JRA1 Requirements Survey

This appendix shows the surveys used to gather the requirements.

A.1 Requirements gathering form for the AARC Project

The purpose of this survey is to gather further requirements and use cases for AAI from s user communities and research infrastructures.

If your community has already produced a similar document with the same topics (AAI requirements), please provide a link to the document and focus on the questions that have not been answered, and the new developments not included in the referenced document.

A.1.1 Part 1: Brief high-level description of the use case

Please, provide a brief, high-level description of the use cases for your community.

A.1.2 Part 2: Current AAI status of your community/research infrastructure

What is your community's current experience with AAI?

- Does your community/research infrastructure already use AAI solutions for their use case?
- What benefits do you see for Federated Access?
- What barriers do you see in joining a federation?
    - The Management doesn't consider that important
    - Not enough funds/resources
    - Lack of technical knowledge
    - It is unclear how to organise an Identity Management (IdM)
    - It is unclear whether the IdM should be internal or external
    - There is no clarity in the organization about the benefits of using an IdM
    - We already have an IdM but it is not completely compatible with SAML/OpenID or other industry-standards
    - Too much bureaucracy to join a federation
    - Other
- What is the user experience in the interaction with the available AAI solutions?

- Expectations fulfilment
- User friendliness
- Quality of service delivered by the tools

How can the penetration of AAI in your community be improved?

- Do you think that your organization is lacking in information about federated identity management?
- What material do you need to inform your organization about federated access? Do you see the need for training to better inform representatives of members in your research area?
- If your organization is not yet part of a federation, what could be helpful towards its joining one? For example:
  - More informative material for management and decision makers
  - Technical personnel dedicated to IdM
  - Guides and training on how to implement an IdM and an Identity Provider
  - Solutions that are ready to use (such as Virtual Machines)
  - More resources
  - External technical support
  - A simpler procedure to follow for joining
  - Other

What are the technical solutions adopted?

- Can you describe the solutions you have adopted, highlighting as applicable?
- Technology or technologies adopted:
  - X509
  - SAML2
  - OpenID Connect/OAuth2
  - OpenID
  - Kerberos
- Identity Providers (IdP) federations integrated (e.g. eduGAIN) or:
  - Approximate number of individual IdPs integrated

- Approximate number of users

- Solution for homeless users (users without a federated institutional IdP)
- Solutions to handle user attributes

A.1.3 Part 3: Requirements for federated AAI

Which type of Identity Providers are relevant to your community?

Which IdPs would your users use?

- Their institutional IdP, part of national federations and eduGAIN.
- Non-federated institutional IdPs
- Research community catch-all IdP
- Social media

What is their preferred authentication technology?

What are the requirements for the authentication technologies to be used in your use cases?

- User friendliness, single sign-on (SSO)
- Web browser and non-browser applications support
- Support multiple technologies and credential translation services (e.g. SAML -> X509 translation)
- Support for delegation (e.g. execute complex workflows on behalf of the user)

Could your community benefit from Scalable IdP attribute release policies?

In your use case, do you foresee the need to get attributes (e.g. institution, email address, name…) from the IdP of the user? If the users will use many different IdPs, coming from different institutions, the service providers supporting the community need to access these attributes, therefore there is need for a set of policies that make scalable the negotiation between SPs and IdPs.

Does your community need persistent identifiers for users?

Do you foresee for your use cases the need to have a persistent identifier to be associated to a user's identity? Supporting a persistent identifier allows the users to more easily change IdP while preserving their identity.

Is support for different levels of assurance relevant to your use cases?

Does your community need to allow users to use credentials with different LoA, and properly communicate the information to the service providers about the LoA of the used credential?

If yes, whom can we contact to ask further questions on your LoA needs? There is a dedicated task in AARC that investigates LoA.

Does your community need community-level authorization?

Authorization is separated from authentication and iis controlled by the community. To implement this community-driven authorization, user communities must manage a set of attributes associated to the users, which need to be provided to the service providers together with the identity attributes provided by the IdP.

How is the current coverage of IdP federations?

Are the current identity federations (e.g. IGTF or eduGAIN) covering enough identity providers/institutions to be a feasible option for your users? What are the use cases where the coverage is not sufficient to reach all the users involved?

Other requirements

Please feel free to add more requirements or topics to be discussed within the AARC project.

Appendix B Organisations that responded to the survey

This section describes the user communities that participated in the Survey and outlines their basic AAI use cases.

B.1 BioVel

BioVel (Biodiversity Virtual e-Laboratory) supports researchers in the domain of ecology, biodiversity and ecosystems science. BioVeL supports scientists to carry out research on biodiversity by offering computerised tools ("workflows") to process large amounts of data from users' own and cross-disciplinary sources, as well as tools for designing and running workflows. BioVeL's e-infrastructure for data analysis of biodiversity research is one of a number of European projects contributing to LifeWatch, which is the research infrastructure program initiated in the European Strategy Forum on Research Infrastructure (ESFRI) in order to construct and operate facilities, including virtual laboratories for biodiversity research. BioVeL's e-infrastructure also aims to support GEO BON, the Group on Earth Observations Biodiversity Observation Network, which coordinates activities relating to the Societal Benefit Area (SBA) on Biodiversity of the Global Earth Observation System of Systems (GEOSS).

In this context, BioVel users need access to data repositories in order to search and upload data, and to access computing services for analysing the data and then storing the results of their analysis. The BioVel community is leveraging on both internal service providers, for example for the data repositories, and on the European multidisciplinary e-infrastructures for accessing computing resources. SSO capabilities are thus fundamental to enable scalable workflows, together with a uniform authorisation infrastructure. As the BioVel community can also include citizen scientists, integration with low Level of Assurance credentials such as social media credentials will need to be supported.

B.2 DARIAH

The Digital Research Infrastructure for the Arts and Humanities (DARIAH-EU) is a Pan-European Research Infrastructure Consortium (ERIC) formed as part of the ESFRI strategy. The blocks composing the research infrastructure build on national initiatives. Digital research methods are becoming a fundamental part of the mainstream of humanities, arts and social sciences research. The digital arts and humanities are at a critical point in the transition from a specialty area to a full-fledged community with a common set of methods, sources of evidence and infrastructure. All of these are necessary for achieving academic and data-driven scientific recognition. Information and data-intensive, distributed, collaborative and multidisciplinary research is now becoming the norm in many research areas. In this context, the goal of DARIAH is to provide a connected network of people, information, tools, and methodologies for investigating, exploring and supporting work across the broad spectrum of the digital humanities.

Currently, the DARIAH community has more than 3000 active users, who profit from the SAML and LDAP-based DARIAH AAI. Most of the users are not affiliated with an institutional IdP, sothe DARIAH AAI includes a community IdP that is integrated into the DFN AAI and is therefore already part of eduGAIN. The DARIAH user accounts are managed in a delegated distributed fashion. The DARIAH AAI also manages VO attributes (memberships of privilege groups) for the DARIAH accounts as well as for the users who can already connect with their campus IdP. Although all DARIAH SPs conform to the GÉANT Data Protection Code of Conduct DARIAH has found that campus IdPs are not yet willing to release a minimal set of attributes to DARIAH SPs, which is an additional reason for operating the homeless IdP..

B.3 EISCAT

EISCAT, the European Incoherent Scatter Scientific Association, was established to conduct research on the lower, middle and upper atmosphere and ionosphere using the incoherent scatter radar technique. The EISCAT Scientific Association is funded by six research councils. The operation of the facilities is divided into two programmes, one common programme for joint activities, and another run by different associates depending on funding. The lower levels of data gathered are available only to the associate countries, while in the non-common programme, each associate has exclusive rights for one year. In recent years, a programme for smaller organisations was started to operate the facilities at relative small costs. These affiliate organisations have the right to access data for one year after the date of observations.

Thus, the main use case for authentication and authorisation in EISCAT is to facilitate access to datasets to the institutions/users who are eligible to download the data. Access control has so far been based on IP addresses but, with the inclusion of affiliates, this becomes more and more complicated to manage. Therefore, federated AAI capabilities could significantly improve this process and, at the same time, enable the monitoring of resources consumed by individual users for accounting purposes.

B.4 WLCG

The World-wide LHC Computing Grid project (WLCG), is a global collaboration, led by CERN, linking hundreds of computer centres worldwide. It was launched in 2001 to provide a global computing resource to store, distribute and analyse the data generated by the LHC. The infrastructure, built by integrating thousands of computers and storage systems, enables a collaborative computing environment on a scale never seen before. The WLCG serves a community of more than 10,000 physicists around the world with near real-time access to LHC data, and the power to process it.

In most HEP distributed infrastructures, a high level of trust and fine-grained authorization are required and this is currently implemented using x509-based personal and host certificates. Non-browser-based applications are also a fundamental component of the WLCG use cases.

B.5 EPOS

The European Plate Observing System (EPOS) aims to create a Pan-European infrastructure for solid Earth science to support a safe and sustainable society. The mission of EPOS is to monitor and understand the dynamic and complex Earth system by relying on new e-science opportunities and

integrating diverse and advanced Research Infrastructures in Europe for solid Earth Science. EPOS will enable innovative multidisciplinary research for a better understanding of the Earth's physical and chemical processes that control earthquakes, volcanic eruptions, ground instability and tsunami as well as the processes driving tectonics and Earth's surface dynamics.

EPOS is now entering its Implementation Phase (EPOS-IP). One of the main challenges during the implementation phase is the integration of multidisciplinary data into a single e-infrastructure. Multidisciplinary data are organized and governed by the Thematic Core Services (TCS) and are driven by various scientific communities encompassing a wide spectrum of Earth science disciplines. TCS data, data products and services will be integrated into a platform, "the ICS system", that will ensure their interoperability and access to these services by the scientific community as well as other users within the society. This requires dedicated tasks for interactions with the various TCS, as well as the various distributed ICS, such as High-Performance Computing, high-throughput computing, and cloud.

B.6 Photon and Neutron community (Umbrella)

Umbrella is an identity system designed by the European Photon and Neutron source facilities (PaNs'). It aims to make life easier and science more productive both for the facilities and their users. Umbrella first of all provides any PaN-user (and effectively anyone interested in scientific discovery) with a unique identity, the UmbrellaID. Equipped with such an ID a user can virtually access the facilities with a single sign-on. Since the same Identity is known at each of the facilities, a user can more simply access or share data, manage administrative processes or make use of services and infrastructures provided by the PaNs'. Umbrella is a joint project of the PaNs' and other facilities with similar needs for an Identity Management System. The joint nature of this undertaking is the major benefit for the facilities. It permits to share the efforts for developing and maintaining the Umbrella system. Services offered by one of the facilities can be used by any of the users, allowing sharing of services within the Umbrella federation, which not only reduces the overall maintenance efforts but also leads to a richer eco-system of services for the user communities.

Future user operation at large-scale facilities will result in users' needing a unique persistent user identification to have unified access to the following functionalities: a) 40% of the users do experiments at different facilities and need trans facility access, b) need for access to and management of experimental data, c) online entry mode: remote experiment access, d) access to efficient data analysis tools, e) remote file access, f) minimal administration load for users.

Umbrella is part of several FP7 projects namely: EuroFEL- ESFRI project Free Electron Lasers of Europe, PaNData-Europe & PaNData ODI- FP7 projects, CRISP – Cluster project of different ESFRI projects, CALIPSO – I3 synchrotron community, NMI3 - I3 neutron community, BioStruct-X –structural biology with synchrotron radiation.

B.7 ELIXIR

ELIXIR is building a sustainable European infrastructure for storing, analysing and sharing biological data, supporting life science research and its translation to medicine, agriculture, bio industries and society. ELIXIR unites Europe's leading life science organisations in managing and safeguarding the massive amounts of data being generated every day by publicly funded research. It is a Pan-European research infrastructure for biological information.

The infrastructure is organized in vertical platforms, one of which is the ELIXIR compute platform. The authentication and authorization infrastructure ("AAI") is part of that platform and will provide user identity and access management services for the ELIXIR services ("Relying services").

Actors in the ELIXIR AAI are currently natural persons. The ELIXIR AAI can be later extended to also cover machine accounts (e.g. servers and portals).

ELIXIR AAI will be deployed gradually in the ELIXIR EXCELERATE project starting in September 2015.

B.8 CLARIN

CLARIN, (Common Language Resources and Technology Infrastructure) aims to provide easy and sustainable access for scholars in the humanities and social sciences to digital language data and advanced tools to discover, explore, exploit, annotate, analyse or combine them, wherever they are located. CLARIN is building a networked federation of language data repositories, service centres and centres of expertise, with single sign-on access for all members of the academic community in all participating countries.

At the time of writing, the CLARIN infrastructure is still under construction, but a growing number of participating centres are already offering access services to data, tools and expertise. The goal is to facilitate access to restricted language resources available within CLARIN.

Thus, rather than requiring academic users to register a new username and password for each individual web application, users should be able to login with their existing institutional credentials by leveraging on IdP federations.

B.9 EGI

Over the last decade, EGI has built a federation of long-term distributed compute and storage infrastructures that support research and innovation. This international e-infrastructure has delivered unprecedented data analysis capabilities to more than 38,000 researchers from many disciplines. The federation brings together more than 350 data and compute centres worldwide. EGI is governed by EGI.eu and funded through a combination of membership fees and national and EC funding. Today, EGI provides both technical and human services, from integrated and secure distributed high-throughput and cloud computing, storage and data resources to consultancy, support and co-development. The research supported by EGI is diverse. Examples include supporting the LHC experiments, bioinformatics, astronomy and astrophysics and climate simulations.

B.10 EUDAT

EUDAT's objective is to build a Collaborative Data Infrastructure (CDI) as a Pan-European solution to the challenge of data proliferation in Europe's research communities. EUDAT will allow researchers to share their data within and between the communities. The expectation is that the services offered through the CDI will foster innovative, multidisciplinary research. EUDAT aims to provide a data management solution that will be affordable, trustworthy, robust, persistent and easy to use.

The EUDAT infrastructure is implementing an AAI solution to bridge various AAI technologies used within the communities and various technologies used by the service providers within the infrastructure. Users of the infrastructure must be able to use their home organisation's Identity Provider to log into the EUDAT infrastructure. The user identity contains attributes provided by the community Identity / Attribute provider. This identity, provided by the community, is mapped to a EUDAT identity and augmented with attributes required by EUDAT if needed. If a user has no access to an Identity Provider, EUDAT can provide the means to create an account and act as an Identity Provider. This is currently being implemented in the B2ACCESS service, which will be released soon. Communities currently participating within EUDAT come from the domains of linguistics, climate, earth observations and human physiology.

B.11 D4Science

D4Science is an e-infrastructure – developed over time by several initiatives,including the iMarine EU project and the forthcoming BlueBridge EU project – providing data and services to efficiently support activities related to the fisheries sector. The data and services provided are accessible through a web Portal and SOAP-based web services. All these access points need Authentication and Authorization; Identity Federation to trusted external infrastructures should be possible. SOA3 (Service Oriented Authentication, Authorization and Accounting) is the framework providing Authentication and Authorization (and Accounting) services for both the web portal and the web services; it also supports SAML-based Identity Federation.

B.12 PSNC

The Poznań Supercomputing and Networking Centre runs the IDP service (https://sso.man.poznan.pl), which is a member of eduGAIN and PIONIER.Id federations. This gives its users access to a number of services federated in these two federations.

As a service provider, PSNC runs:

- Filesender service for PIONIER consortium (accepts PIONIER.Id users)
- Web Conference Service for PIONIER consortium (accepts PIONIER.Id users)
- Plans to enable federated access also to other browser-based services:
- Citizen Science Registry of Resources (https://registry.civic.psnc.pl/ , which is required to give a wide group of volunteers controlled and not anonymous write access to the Registry to populate it with data)

Considering non-browser services, PSNC is also interested in enabling federated access to our Popular Archival Service.

B.13 FMI

The Friedrich Miescher Institute (FMI) is an internationally recognised centre for fundamental biomedical research. As an affiliated institute of the University of Basel and the Novartis Institutes for BioMedical Research, the FMI is actively involved in both academic research and biomedical application. The FMI has established cutting-edge technology support platforms that develop and exploit new tools for biomedical research. These platforms, ranging from functional genomics to microscopy and imaging, focus on the development and optimisation of critical technological solutions in areas such as neurobiology, epigenetics and mechanisms of cancer.

FMI has about 1600 users, using more than 90 different IdPs. FMI research is carried out by independent but highly interactive teams across institutions and countries. Researchers are often not federated in national federations, for example those working in hospitals, therefore services need to federate with the IdPs involved individually. The need to support social logins is also foreseen.

B.14 Libraries and education

The community consists of Students, Researchers, Professors in the university system, and researchers accessing libraries to access and share documents and data. In particular, for the scope of this document, information has been gathered from Italian Universities and LIBER partners focusing on the Dutch libraries, respectively by GARR, the Italian NREN, and LIBER in collaboration with SURFnet, the Dutch NREN.

Examples of use cases from this community are:

- Each university is entitled to run an IdP for their users. Until now 42 universities out of 80 in total have their own IdP in place and running. One major issue is to try to speed up the IdP adoption.

- Behind the IdP it is necessary to have the Identity Management in place. This means to automatically provide identities from authoritative databases. This operation should be standardized. Provisioning and deprovisioning have to be considered.

- Libraries need to better integrate their services with the IdP authentication and go beyond the IP-address based authorization.

An open question is how to provide federated identity to educational institutes in the national education system. A key issue in this area is whether we should start from a bottom-up approach and build a new community of trust step by step – to be progressively extended to the highest possible number of schools – or to wait for the full establishment of the national identity system (SPID, in Italy), and take over from there, including institutes in the eGov-based trust model and ensuring interoperation with the national trust federation for Research.

References

[FIM4R2012] Broeder, D., Jones, B., Kelsey, D., Kershaw, P., Lüders, S., Lyall, A., Nyrönen, T., Wartel, R., & Weyer, H.J. (2012). Federated Identity Management for Research Collaborations. Report number CERN-OPEN-2012-006. http://cds.cern.ch/record/1442597/files/CERN-OPEN-2012-006.pdf

[FURPS+] Grady B.R. (1992). Practical Software Metrics for Project Management and Process Improvement. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

[TERENA AAA Study] Advancing Technologies and Federating Communities, A Study on Authentication and Authorisation Platforms For Scientific Resources in Europe. FINAL REPORT A study prepared for the European Commission DG Communications Networks, Content & Technology, 2012. https://www.terena.org/publications/files/2012-AAA-Study-report-final.pdf

[GNSA52014] Towards Horizon 2020 – The Enabling Users Experience. http://www.geant.net/Resources/Deliverables/Documents/D9-4_DS5-5-1_Towards-Horizon-2020_The-Enabling-Users-Experience %283%29.pdf

Glossary

AAI	Authentication and Authorisation Infrastructure
eduGAIN	International interfederation service interconnecting research and education identity federations
IdP	Identity Provider
IGTF	Interoperable Global Trust Federation – A body to establish common policies and guidelines that help establish interoperable, global trust relations between providers of eInfrastructures and cyber-infrastructures, identity providers, and other qualified relying parties.
LoA	Level of Assurance – degree of certainty that the user has presented a credential that refers to that user's identity
SAML	Security Assertion Markup Language
SP	Service Provider
VO	Virtual Organisation – a dynamic set of individuals or institutions defined around a set of resource-sharing rules. Resource sharing is, necessarily, highly controlled, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs.