28-07-2015

Deliverable DNA2.1: Report on the identification of target groups and their requirements

Deliverable DNA2.1

Contractual Date:	31-07-2015
Actual Date:	28-07-
Grant Agreement No.:	653965
Work Package:	NA2
Task Item:	T1
Lead Partner:	GÉANT
Document Code:	DNA2.1
Authors:	Alessandra Scicchitano (GÉANT), Maria Laura Mantovani (GARR), David Group (Nikhef)
Contributors:	Licia Florio (GÉANT), Peter Gietz (DAASI)

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 reports on the work done by NA2 Task 1 "Learning Needs Analysis" in liaising with user groups and communities including libraries with the objective of understanding their identity management requirements and needs.

Acknowledgement

Many thanks to Lukas Hämmerle from SWITCH for his important contribution to this document.

Table of Contents

Executive Summary

This report on the identification of target groups and their requirements for training and outreach by NA2 task 1 provides the learning needs analysis fro the AARC project based on both targeted surveying and in-depth interviews of representative communities.

The objective of this work is to understand the identity management needs and requirements of the target groups as well as assessing their current level of knowledge and skills. The target groups include libraries, arts and humanities, bio-medical, high-energy physics and e-Infrastructures as well as NRENs (that as representatives of national federations provide feedback on the underlying organisations that participate).

To ensure sufficient coverage of potential target groups, we used a two-pronged approach:

- A survey (see Appendix A) with both open and multiple-choice questions sent to organisations not yet federated and that by their nature constitute a representative cross-section of the community.

- In-depth interviews, conducted with specific trans-national user communities selected from distinct areas of research and e-infrastructure, as well as with the library community.

To complement this data, the national research network organisations (NRENs or national R&E federations) were asked to contribute information on national outreach and training needs of their connected organisations.

The results of the survey provided a good opportunity to understand the main barriers and obstacles that an organisation faces when joining an identity federation. By talking to individual user communities we could learn about their specific needs and gain a deeper understanding of their existing knowledge and skill levels. One to one meetings also enabled more in-depth discussion and elaboration on how NA2 could help fill the existing gaps.
The major areas identified for training are:

- Materials addressing decision and policy makers at organisational level, to clarify the value proposition of federated AAI and the need for its timely deployment.

- An organised catalogue highlighting available resources and services accessible through federated AAI.

- Training addressing prevalent technological knowledge gaps so as to facilitate adoption.

- Development of guidance for both identity and service provider operators, especially for the trans-national (research) communities and libraries.

The requirements collected in this phase are the foundation for the work package NA2 “Training and outreach” to prepare material and trainings that address these specific needs.

The Learning Needs Analysis task (NA2 Task 1) will continue to work with the individual communities in an iterative process to assess the impact of the training and materials provided, and continue to adapt these to evolving needs.

1 Introduction

During the last decade, national Identity Federations for Research and Education have emerged across the whole of Europe and beyond.

The lack of seamless integration between the different Authentication and Authorisation Infrastructures (AAIs) operated by the various research collaborations and e-Infrastructures, the non-ubiquity of federated credentials and technical and policy challenges are ultimately hindering the sharing of knowledge across the European and global research collaborations.

In addition to interoperability, functional gaps also exist. These aspects include support for aggregating information (attributes) needed for authorisation purposes from multiple attribute providers, better support for Single Sign-On for non-web applications, and a diverse training package to cover both technical, legal and policy matters.

The goals of the Authentication and Authorisation for Research and Collaboration (AARC) project address these challenges of interoperability and functional gaps. [AARC]

The focus of NA2 “Training and Outreach” is on a set of key aspects that will be addressed based on these common challenges.

The specific focus of NA2 Task 1 “Learning Needs Analysis” is on identifying the knowledge and skills gaps among target groups, including libraries, in order to prepare effective training to meet their needs.

Organisation of this document

Section 2 of this document describes the process and methodology adopted for the requirements gathering work. Section 3 describes the findings of the requirements gathering. Section 4 discusses the implications of the findings and conclusions. Appendix A provides a copy of the survey used for the requirements gathering.

2 Process and the Methodology

The work of gathering requirements started with screening the work that has already been done in FIM4R [FIM4R paper] and the AAI workshop and consultation meeting [AAI workshop] held in Brussels in 2014 as well as the Study on Authentication and Authorisation Platforms for Scientific Resources in Europe conducted by GÉANT (formerly, and at the time, TERENA) [AAA report]. This screening provided an overview of the different communities and the status of their federation.

When reviewing the requirements for data gathering it became clear that a three-dimension approach was needed. The individual requirements of each of the different communities were an important factor but it was also important to consider the roles that people cover within the community itself in relation to federation. Different roles have different needs and different requirements. These first two dimensions, user communities and target roles are described in section 2.1 and 2.2.

The third dimension considered was whether an institution deploys federated access or not. The organisations could be then divided into federated and non-federated bodies. This distinction would highlight the obstacles that especially non-federated organisations face when moving towards federation.

The methodology followed used two main approaches to identify the requirements:

- A survey sent to the organisations belonging to different communities that are not yet federated. The questions in the survey were elaborated keeping in mind the different target roles.

- Meetings with specific user communities.

Both approaches are presented in more detail in section 2.3 together with the results of the survey.

2.1 User Communities

A number of user communities with which to engage have been previously identified in the project:

- Libraries: libraries have been a traditional intermediary between researchers and sources of research information. While the benefits of moving towards federated access are clear (data sharing, access to more technology to exploit data, replacement of IP address based authorization, etc), there are still a number of barriers - like access to publishers and licensed material in general, as well as

emerging requirements for access to support data management - that need to be addressed.

- Arts and Humanities: as highlighted by the FIM4R paper [FIM4R paper], in the Arts and Humanities community there are several research infrastructure projects that have identified the need for AAI and that have implemented SAML based infrastructures. Facing the challenge of not getting enough information from campus Identity providers, they had to find solutions to become more self-sustained, either by creating a Service Provider federation (CLARIN), or by operating a community based Identity Provider and integrating that into the eduGAIN federation (DARIAH).

- Life Sciences: the Life Science research community is extremely large and generates significant volumes of data. Like the libraries, the benefits for this community to join national federations are different depending on local context – from the challenges of storing such data to specific limitations on access requirements.

- High-energy physics: the High Energy Physics (HEP) community is already successfully using federated identity and SSO for the Worldwide LHC Computing Grid (WLCG) in the form of X.509 certificates. In most HEP distributed infrastructures, a high level of trust in the vetting of the users and fine-grained authorisation are required; these requirements are met by using X.509 certificates issued by certification authorities that are accredited by the International Grid Trust Federation (IGTF). Support for non-browser-based applications is also needed. Due to the high number of applications used by this community that do not run in the web space, the deployment of SAML-based federated access has been a big challenge. Access to traditional or simple services (like wikis or web portals) would benefit from the simplicity and ubiquity of federated identity. In recent years however there have been a number of projects to bridge the gap between the SAML and the X.509 approaches. These have resulted in portals where users log in using their federated credentials; via the portals users can request and obtain x.509 certificates.

- E-Infrastructures: the past decade has seen the emergence of computing and data infrastructures in support for research in Europe, encompassing resources from a large number of different providers and concurrently used by many different research communities. The provisioning of collective services in the e-infrastructure has been the driving force for one of the most characteristic aspects of the grid AAI.

NRENs as representative of national federations and a broad set of institutions were also queried with the intent to extract needs and requirements from within their constituencies. A broad variety of NRENs are involved in AARC, ensuring a representative geographical spread across Europe.

2.2 Target Roles

The target role groups, independent of domain and research community identified, are:

- Decision makers: The term “decision maker” refers to high-level managers in universities and institutions, and leaders of large (research) consortia. The idea is to address the people that decide whether a university or consortium should join a federation or not, make them aware of the benefits of federated identity and build a business case for the use of federated identity for large-scale research projects.

- Identity Provider (IdP) operators: The people in the organisation that set up and operate an Identity Provider service. Addressing the requirements of this role would ease the implementation of new IdPs and support existing IdPs struggling to make appropriate decisions regarding attribute release and legal requirements.

- Service Provider (SP) operators and Service developers: Service provider operators and service developers have a distinct role in the uptake of federated identity. A service developer is in general interested in knowing how to integrate SAML in an application so to connect it with the SP, while the SP operator is interested in integrating the SP into the federation, which is less technical (certificates) and much more organisational (federation contracts, data protection code of conduct, etc.). It is important to address the requirements of both categories.

- Endusers: ease of access in terms of consistent login approaches, seamless access without error messages and other barriers, and effective group management are all important aspects for end-users.

2.3 Approach

Keeping in mind the target communities and roles, a survey and meetings with the specific communities were exploited to collect requirements.

These two approaches are described below:

2.3.1 Survey

The questions prepared for the survey followed the three-dimension approach described above. Tailored to technical people, the questions were focused on collecting requirements from organisations not yet federated and belonging to different communities. This set of questions was circulated in a conjunct effort together with JRA1 (see Appendix A) as one survey among different communities, while NRENs were asked to distribute a modified

version of the same survey (where only the questions from NA1 Task 2 were modified and left) within their constituency (See appendix B).

The group of organisations involved in the survey that were not yet federated was then further divided in two groups:

1. Universities.

2. Other institutions (*National libraries and archives, Research institutions in bio-medical fields, Institutions for Arts and Music).*

This distinction was made on the assumption that the technical people in the Universities have a higher level of knowledge and are technically more skilled than the ones in the institutions. In fact institutions are generally focused on their specific missions overlooking the need of a broad technical understanding of the digital identity related matters.

The difference in results of the survey shows that this assumption and the classification were indeed correct.

2.3.2 Meeting with Specific Communities

The user communities identified as the community to target are (by virtue of construction of the AARC consortium) also well represented in AARC itself:

- Libraries: represented by LIBER and MKZ.
- Arts and humanities: DARIAH represented by DAASI.
- Bio-medical: ELIXIR represented by CSC.
- High-energy physics: represented by CERN.
- E-Infrastructures: represented by EGI and SURFsara (for PRACE and – jointly with CSC – for EUDAT).
- NRENs: A broad variety of NRENs is involved in AARC. They not only represent the national federations, but also the organisational communities that have been supported by NRENs for decades. Because of this long standing relationship, NRENs have a good understanding of the differing technical and organisational requirements of different roles within the community (e.g. library, IT staff, educators, students, researchers etc.).

The meetings - held either in person or via VC - were focused on gathering the requirements of these communities as well as understanding how to best match their needs. The meetings

allowed room for discussion, which provided a much deeper insight of the obstacles that should be addressed by NA2.

3 Results

3.1 Survey Findings

This section describes the results from gathering requirements via the survey. In our 3-dimension approach the questionnaire was sent to a group of technical people (target role) belonging to organisations that are not federated yet (status). At the same time, with this survey we targeted all the user communities. The questionnaire was sent to a group of technical people belonging to organisations that are not yet federated. The group of organisations was further divided as:

1. Universities.

2. Other Institutions (*National libraries and archives, Research institutions in bio-medical fields, Institutions for Arts and Music).*

The idea behind the survey was to build a clear picture on the obstacles that are faced when moving towards federation. Obstacles rather than the requirements were focused on as it is often difficult for non-federated organisations to articulate the specific requirements of implementing middleware approaches. It is easier for these organisations to describe existing problems and where they perceive challenges in using federated identity as a solution.

The particular focus of this study was to gain input from the organisations in their role as identity providers. Connecting and federating identity providers - and through these the users, researchers, educators and students - is a prerequisite in order to perceive benefit from joining a federated AAI. It is also worthwhile to note that the role responsible for identity management and provisioning teams in an organisation will be customarily disjoint from the service providing groups, since the latter are driven by research use cases and application developers, whereas identity provisioning is organically linked closer to human-resources processes and registrar functions.

25 organisations answered the survey. The results of the closed questions (expressed in percentage) are shown below.

Universities

15 Universities answered the survey. They all declared good understanding of Identity Management and SAML and around 70% of them confirm to know the eduGAIN service.

Main obstacles

- Lack of technical people that know/can learn how to setup an IdP (76%).
- Management of the organisation does not consider having an IdP as business priority (24%).
- Lacking of a well-organised catalogue of eduGAIN resources that attracts the interest of the organisation (18%).

Support required

- IdP as a Service (41%).
- More technical people to devote to identity management (IdM) and the setup and operation of an IdP(35%).
- Information material for decision-makers to show the economic advantages of identity federations (24%).

Other Organisations: National libraries and archives, Research institutions in bio-medical fields, Institutions for Arts and Music

10 Organisations answered the survey. While there is some knowledge of what Identity Management is, almost all of them have no knowledge about eduGAIN or SAML.

Main obstacles

- Lack of technical people that know/can learn how to setup an IdP (60%).
- Lack of knowledge/technical expertise (30%).

Support required

- More technical people to devote to IdM/IdP (50%).
- Information material for decision-makers to show the economic advantages of identity federations (40%).

- "How-to" and training documents / events to implement IdM and IdP (30%).
- IdP as a Service (30%).

Common Identified Requirements:

The requirements reported in this section also include the ones collected via open questions.

- Information material for decision makers: when asked for more information on why the organisation lacks the technical people required to implement an IdM the most prevalent answer was that decision makers do not understand the business benefit of federated identity. Lack of identifiable benefits leads to skepticism and reluctance to invest money in hiring or training personnel on the deployment and operation of federated identity services.
- Training to fill in the gaps in the knowledge of the technical people already in the organisation.
- A better organised eduGAIN service catalogue that shows the benefits that the resources can bring to the organisations.
- Information material about eduGAIN.

3.2 Meeting Findings

3.2.1 Libraries

In AARC, the library community is well represented by various organisations. In order to effectively collect their requirements we have involved:

- LIBER, which is the main research libraries network in Europe. It has over 430 members from national, university and other research libraries across 45 countries. LIBER is actively working to promote the role of libraries within the European research infrastructure [LIBER].
- MKZ, the Moravian Land Library in Brno which is a research organisation whose main purpose is to carry out basic research, applied research or experimental development, and to disseminate their results by means of education, publications or transfer of technologies [MKZ].

- 2 NRENs partners in AARC - GARR and SURFnet - managing university libraries and national libraries in their federations. Both NRENs had previously collected requirements from already federated libraries.

Through these organisations it was possible to talk to different libraries and to collect common requirements. 6 Czech and 2 Slovakian Libraries not yet federated were interviewed by MKZ while LIBER interviewed UKB, the Dutch consortium (representing thirteen university libraries and the National Library of the Netherlands) and KB, the National Library of the Netherlands.

Requirements:

- Training on how tools used by the libraries can be configured to fulfil the library needs and improve the user experience – e.g. effective use of discovery tools. Involvement of tool producers is probably needed

- Open e-learning course, guiding participants through the federated login workflow, using a less technical and more operational language and clearly demonstrating federated login and SSO benefits.

- Some short, easy to understand materials explaining how federated access works for librarians and for library users and trainings materials that librarians can reuse towards end-users.

- Information material for decision makers. Some short general information about federated access explaining the pros and cons involved.

3.2.2 CERN

At CERN [CERN], the European Organisation for Nuclear Research, physicists and engineers are probing the fundamental structure of the universe. They use the world's largest and most complex scientific instruments to study the basic constituents of matter – the fundamental particles.

Founded in 1954, the CERN laboratory sits astride the Franco-Swiss border near Geneva. It was one of Europe's first joint ventures and now has participants from 21 member states. Because of the significant effort needed to perform research in high-energy physics (HEP), the facilities offered at CERN are unique in the world – which in turn attracts a global researcher audience. The collaborations around the Large Hadron Collider (LHC) experiments at CERN typically consist of 10000 researchers, drawn from hundreds of institutions and from over 60 countries and economic regions. It is in itself a federated research effort – even if home organisation credentials are not widely used in the AA infrastructure. For authentication it leverages the Interoperable Global Trust Federation IGTF as a trusted third party, and each of the (4) LHC experiments (research collaborations)

operates and manages its own authoritative membership database and attribute authority, mostly independent of the user's home organisation. Researchers affiliation with the CERN experiments also typically outlasts any (usually transient) affiliation with any particular university or research lab.

Requirements:

- Very high level understanding of identity federation (benefits, what changes, why it is necessary) for all. The user and business benefits of a federated approach are still poorly understood by users and management– it is seen as costly, complicated and with unknown benefits.

- Operational security training for eduGAIN IdPs and SPs. It is perceived that IdPs and SPs in eduGAIN have very little/no experience of computer security incident handing and operational security in general. Training eduGAIN IdPs and SPs to be aware of and ready to tackle operational security challenges is absolutely essential.

- "Enabling the X.509 SPs" technical training, to enable SPs based on X.509 certificates to make use of identity federations. Enabling this federated mode of operation requires know-how and re-assurance of the SPs. Training in this area would have significant positive impact.

3.2.3 DARIAH

DARIAH [DARIAH], the Digital Research Infrastructure for the Arts and Humanities, aims to enhance and support digitally enabled research and teaching across the humanities and arts.

DARIAH will develop, maintain and operate an infrastructure in support of ICT-based research practices ("Digital Humanities"). By working with communities of practice, DARIAH-EU will bring together individual state-of-the-art digital Arts and Humanities activities across Europe. It will preserve, provide access to and disseminate research (including research data) that stems from these collaborations and ensure that best practices, methodological and technical standards are followed.

Requirements:

- More information material for management and decision makers.

- Technical material and resources for IdP operators and for service developers.

- Material on privacy legislation to help IdP-operators to release user attributes to DARIAH SPs.

- Support application developers on how to AAI-enable their application.

3.2.4 ELIXIR

ELIXIR [ELIXIR] aims at building a sustainable European infrastructure for biological information, supporting life science research and its translation to medicine, agriculture, bioindustries 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.

ELIXIR will provide the facilities necessary for life science researchers - from bench biologists to cheminformaticians - to make the most of our rapidly growing store of information about living systems, which is the foundation on which our understanding of life is built.

Requirements:

- Increase the uptake of federated access within different research communities by addressing the main technical and policy challenges that prevent implementing AAI.
- Enhance and supplement the existing technical material as needed.
- Training material should consist of a set of reusable training modules.
- Help the service providers who are proceeding with implementation of federated AAI.

3.2.5 EGI

The European Grid Initiative [EGI], is a high throughput and cloud infrastructure distributed across 40 countries in Europe and beyond. EGI provide resources and services to diverse research disciplines from high-energy physics to life sciences, computational chemistry and the humanities.

Today EGI, like the high-energy physics community around CERN, uses services that rely on credentials issued by Certification Authorities accredited by the IGTF, which act as a trusted third party. These credentials are issued using an end-user facing "PKI" (public key infrastructure) which although solving many of the problems of a distribute authentication and authorization framework, is not considered very user friendly, in particular by the new users who are approaching EGI. Many of these users already own credentials from R&E

federations, therefore EGI hopes to strengthen the collaboration with the IdP federations, in order to provide a better user experience, enabling either the direct use of such credentials or the use of federated identities in the credential translation services already in place within EGI.

The EGI-Engage project supports 8 research community-specific Competence Centres (CC) [EGI-CC] whose work aims to develop the 'Knowledge Commons' of the Open Science Commons [OSC].

At the EGI conference held in Lisbon, Portugal in May 2015, an AAI session was organized by AARC and EGI. All CCs were invited to attend.
6 were present at the session:

- LifeWatch: One of the main goals of the LifeWatch EGI CC is to capture and address the requirements of Biodiversity and Ecosystems research communities.
- EPOS: The task operates a CC to drive the future design of the use of grid and cloud for the integrated solid Earth Sciences research as part of the European Plate Observing System (EPOS).
- MoBrain: The main objective of the MoBrain CC is to lower barriers for scientists to access modern e-Science solutions from micro to macro scales.
- DARIAH: The DARIAH CC aims to widen the usage of the e-Infrastructures for Arts and Humanities research.
- Disaster Mitigation: The objective of this CC is to make available customised IT services to support climate and disaster mitigation researchers.
- ELIXIR: the ELIXIR CC selects key life science use case workflows that have high impact for scientist end-users and use these as drivers to develop demonstrators that will assess the use of EGI cloud resources for tool and/or data services recognised within the ELIXIR community.

The purpose of the session was to collect the requirements from these communities.
The CCs were prepared ahead of time on the feedback to provide during the session.

Requirements:

- Better coverage of federations to also encompass research and data/storage centres would be beneficial.
- Some of the CCs do not have enough experience in federation yet. Training can help increase the knowledge and improve the overall experience.
- There are concerns about how to treat personal data. Support and trainings in policies and legal aspects (including EU data protection laws [EU data protection]) on this matter would help bridge the knowledge gap and make operators feel more comfortable when treating data.

3.2.6 NRENs

For the sake of completeness some of the NRENs were surveyed in their role of representatives and federation operators of national federations. The objective was to understand and gather through the federations, the requirements of organisations inside their constituency that are already using federated AAI, but that could benefit from additional trainings and support to better service their users for research and collaborative use cases.

Obstacles that prevent a more pervasive adoption of federated SSO that can be addressed with training and outreach are:

- Hard interoperability between SAML and the pre-existing Identity Management system. This also complicates the set-up of a new, efficient and simple identity provisioning/deprovisioning system. Training on Identity Management setup as well as IdM/IdP as a service can mitigate this barrier.

- Management of the organisation (university, nationally based institute) does not consider identity management or the operation of an IdP as a priority. Information and materials extolling the benefits of federated identity would help lowering this barrier.

- Resource providers do not integrate SAML. There are still providers - especially around libraries - that do not offer federated access. Informational material here as well as above could help.

- Some identity management software (in particular Shibboleth) is considered hard to deploy and maintain. Training and material focused on the deployment and maintenance of identity management software suites (of which there are several to choose) can mitigate this barrier as well as offering IdM/IdP as a Service.

- Resources in eduGAIN are not considered interesting. This barrier needs to be read in a different light. The organisations do not find the resources interesting because they do not really know the added value. The information that is actually missing is the benefits the resources can bring when joining a federation. A clear and well organised catalogue of eduGAIN resources would be of great help. This could be done in collaboration with GÉANT.

The obstacles identified in these meetings are inline with the ones collected during the survey that the IDEM federation has performed in May 2015 [IDEM survey].

4 Conclusions

The requirements gathered during this work can be grouped as follows:

- Information material for decision makers and users. One of the main outcomes of the survey as well as of the meetings with the communities is that most of the management and of the users does not seem to understand Identity federation and often fails to grasp the key benefits. A business case focused approach should be used to support institutional decision making.

- Increase the uptake of federated access within different research communities by addressing the main technical and policy challenges that prevent implementing IdM.

- Standardised IdP and SP international set up guides, which can easily be modified for national federations.

- Operational security and technical training for IdP and SP operators and standardised approaches to incident response to support the assurance needs of research collaborations.

- High level training and material that can be reused towards end-users.

- Support and trainings in data privacy legal aspects. This could prove challenging in light of the upcoming EU data protection changes and move from regulation to directive within member states

- A better organised eduGAIN catalogue that shows the benefits that the resources can bring to the organisations.

- Dedicated training for libraries that cover the understanding of the workflow of their use case.

- IdM/IdP as service.

The data gathered from the survey has indicated the major obstacles that organisations face when federating their users and services. It also pointed to what type of support is needed from information material addressed specifically to decision makers to a better organised catalogue indicating the different accessible resources. NA2 will provide information material and trainings to mitigate the knowledge gaps and facilitate the adoption of identity federation.

The meetings with the individual communities highlighted their common requirements such as the need of set up guides or specific operational training for IdP and SP operators. The identified requirements will help shape the trainings and the material produced within the work package. It is worth mentioning that both training and outreach material will not be

designed from scratch where possible. Instead, existing materials created by the GÉANT project, NRENs and other e-Infrastructures will be reused and enhanced as needed.

The last requirements (IdM/IdP as service) will be covered by the results of the activities within SA1 and JRA1. The trainings provided in this area by NA2 will be shaped based on the outcome of these two activities.

We plan to continue interacting with the individual communities to further expand our knowledge of their needs and requirements as well as to observe and measure the impact of the support and trainings provided.

Appendix A The Survey with JRA1

Requirements gathering form for the AARC Project.

The purpose of this survey is to expand the set of requirements and use cases that the user communities and research infrastructures have for AAI.
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.
The questions will cover the technical requirements to design the architecture, but also requirements for outreach and dissemination activities within AARC, to make these activities more effective.

Part 1: Short and high level description of the use case
Please, provide a short and high level description of the use cases of the community you represent.

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

What is the current experience with AAI of your community?

- Has your community/research infrastructure already uses 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.
    - No enough funds/resources.
    - Lack of technical knowledge.
    - It is unclear how to organise an Identity Management (IdM).
    - It is unclear if the IdM should be internally or externally done.
    - There is no clarity in the organisation 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 fulfillment

- User friendliness
- Quality of service delivered by the tools

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

- Do you think that your organisation is lacking in information about federated identity management?

- What material do you need to inform your organisation about federated access? Do you see the need for trainings to better inform representatives of members in your research area?

- If your organisation isn't part of a federation yet, what can be helpful in order to join one? For example:
  - More informative material for management and decision makers
  - Technical personnel dedicated to IdM
  - Guides and trainings on how to implement an IdM and an Identity Provider
  - Something ready that can be already used (like 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

Part 3: Requirements for federated AAI

Which type of Identity Providers are relevant for your community?

- Which IdPs your users would use?
  - Their institutional IdP, part of national federations and eduGAIN.
  - Non federated institutional IdPs
  - Research community catch-all IdP
  - Social media

- What are the 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.

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

Is the support for different level of assurance relevant for your use cases?
Different LoA, allow users to use credentials with different level of assurance, and communicate properly 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 it is 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 involved users?

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

Appendix B The Survey to the NRENs

Do you know what Identity Management is?
Yes No

Do you know what is eduGAIN?
Yes No

Do you know what SAML (Simple Assertion Markup Language) is?
Yes No

What barriers do you see in joining a federation?
Choose max 3 answers

- The Management does not consider that important
- No enough funds/resources
- Lack of technical knowledge
- It is unclear how to organise an Identity Management (IdM)
- It is unclear if the IdM should be internally or externally done
- There is no clarity in the organisation 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:_____

If your organisation is not part of a federation yet, what can be helpful in order to join one?
Choose max 2 answers

- More informative material for management and decision makers
- Technical personnel dedicated to IdM
- Guides and trainings on how to implement an IdM and an Identity Provider
- Something ready that can be already used (like Virtual Machines)
- More resources
- External technical support
- A simpler procedure to follow for joining
- Other:_____

References

[AAA report]	https://www.terena.org/publications/files/2012-AAA-Study-report-final.pdf
[AAI workshop]	https://www.terena.org/activities/AAI-Workshop/22072014workshop-report-final.pdf
[AARC]	https://aarc-project.eu/
[CERN]	http://home.web.cern.ch/about
[DARIAH]	https://dariah.eu/about/mission.html
[EGI]	www.egi.eu/
[EGI-CC]	https://wiki.egi.eu/wiki/EGI-Engage:Competence_centres
[ELIXIR]	https://www.elixir-europe.org/
[EU data protection]	http://eurlex.europa.eu/LexUriServ/LexUriServ.douri=CELEX:31995L0046:en:HTML
[FIM4R paper]	https://cdsweb.cern.ch/record/1442597/files/CERN-OPEN-2012-006.pdf
[IDEM survey]	https://www.idem.garr.it/documenti/doc_download/409-quinto-convegno-idem-sondaggio
[LIBER]	http://libereurope.eu/
[MKZ]	http://www.mzk.cz/en/about-library
[OSC]	http://www.opensciencecommons.org/

Glossary

AAI	Authentication and Authorization Infrastructure
AARC	Authentication and Authorisation for Research and Collaboration
CC	Competence Center
CERN	The European Organisation for Nuclear Research
CSC	IT Center for Science
DAASI	Open source experts for authentication infrastructures, authorisation, encryption and databases
DARIAH	The Digital Research Infrastructure for the Arts and Humanities project
eduGAIN	International interfederation service interconnecting research and education identity federations
EGI	European Grid Infrastructure
ELIXIR	European life sciences infrastructure for biological information
EPOS	European Plate Observing System
FIM4R	A group of large research projects started to investigate Federated Identity Management (FIM) for research collaborations
GEANT	Pan-European data network dedicated to the research and education community
HEP	The High Energy Physics Community
ICT	Information and communications technology
IDEM	IDEntity Management for Federation Access - The Italian National federation
IdM	Identity Management
IdP	Identity Provider
IdP as a Service	Creation and the maintenance of an IdP in the cloud rather than locally
LIBER	Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries
LifeWatch	Science and Technology Infrastructure for Biodiversity date and Observatories
KB	National Library of the Netherlands
MKZ	Moravian Land Library
MoBrain	A Competence Center to Serve Translational Research from Molecule to Brain
NRENs	National Research and Education Networks
SAML	Security Assertion Markup Language
SP	Service Provider
SSO	Single Sign On
UKB	the Dutch consortium of the thirteen university libraries and the National Library of the Netherlands
WLCG	The Worldwide LHC Computing Grid
X.509	X.509 is an ITU-T standard for a public key infrastructure (PKI) and Privilege Management Infrastructure (PMI)