30-04-2016

Deliverable D9.5
User Community Consultation and Use-Case Support Experience

Deliverable D9.5

Contractual Date:	30-04-2016
Actual Date:	30-04-2016
Grant Agreement No.:	691567
Work Package/Activity:	9/SA5
Task Item:	Task 5
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	SWITCH
Document Code:	GN4-1-16-131DE1
Authors:	Lukas Hämmerle (SWITCH), Ann Harding (SWITCH), Wolfgang Pempe (DFN)

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 691567 (GN4-1).

Abstract

This report provides an overview of the work and results of the GÉANT Enabling Users task in the Trust and Identity Service Development activity (SA5) in the GN4-1 project. Enabling Users is one of several GÉANT tasks in the area of federated identity, supplementing the core eduGAIN service task. 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, to support research communities in improving their ability to use federated identity and generally to promote the use of eduGAIN, the interfederation service created by GÉANT.

Table of Contents

Executive Summary 1

1 Introduction to Enabling Users 2

2 The Enabling Users Approach 3

3 Collaborations with Research Communities 4

3.1 CLARIN 4

3.1.1 Description 4

3.1.2 Use Case 4

3.1.3 Collaboration Objectives 5

3.1.4 Expected Benefits 6

3.1.5 Achieved Results 6

3.1.6 Remarks 6

3.2 EIDA 7

3.2.1 Description 7

3.2.2 Use Case 7

3.2.3 Collaboration Objectives 7

3.2.4 Expected Benefits 8

3.2.5 Achieved Results 8

3.2.6 Remarks 9

4 Basic eduGAIN Consultancy 10

5 Additional Results 11

5.1 eduGAIN Access Check Service (EACS) 11

5.2 eduGAIN Connectivity Check Service (ECCS) 11

5.3 eduGAIN Attribute Release Check Service (EARCS) 12

5.4 SP Registration Process 13

5.5 AAI Comparison 15

5.6 Training Events 15

5.6.1 TNC 2015 Attribute Release Training 15

5.6.2 AAI Workshop for Service and Resource Providers 16

5.7 Knowledge Database 16

5.8 Dissemination of eduGAIN to Scientific Communities 16

6 Conclusion and Recommendations 18

References 19

Glossary 21

Table of Figures

Figure 3.1: Federated authentication using a terminal client in EIDA 9
Figure 5.1: Status of eduGAIN Identity Providers in ECCS 12
Figure 5.2: 7 steps to join eduGAIN via the UK Access and Management Federation 14

Executive Summary

In the past year, the number of federations and entities in eduGAIN has once again seen a considerable increase, as have the interest and the needs of scientific communities. However, using eduGAIN is still challenging for some institutions, especially where they may only have a basic understanding of it and of the concept of federated identity management in general. In this context, the main objective of the Enabling Users task is to act as an expert partner to assist research communities and other prospective users of the service.

Previously in the GN3plus project, the Enabling Users task had also operated a number of successful pilots. This aspect of enabling research has now been taken up by the Authentication and Authorisation for Research Collaborations (AARC) project, which started in May 2015. AARC is in a good position to work on pilot use cases with research communities, as many of these are funded partners in the AARC project. While AARC is involved in large-scale pilots of concepts in the context of user-driven development, Enabling Users is helping research communities individually, including by providing tailored consultancy. Also based on AARC's results, the task has been working to improve eduGAIN by extending existing tools and developing new ones to improve the overall quality of the service and mitigate some of its main issues (configuration problems and insufficient attribute release).

During GN4-1, the Enabling Users task has continued to directly and indirectly support research communities as it had done during GN3plus. Two new close collaborations were started with the European Integrated Data Archive (EIDA) initiative and the Common Language Resources and Technology Infrastructure (CLARIN) ESFRI project, which operate respectively in the fields of seismic/geoscience and language research. These collaborations focused primarily on integrating available solutions and included organisations that had directly approached GÉANT for help. In addition, basic support, consultancy, expertise and dissemination were provided for over a dozen other research communities, as well as other e-Infrastructures and the AARC project itself.

Cloud providers, publishers and research communities often have no clear relationship to the individual eduGAIN member federations through which they would have to register their services to access eduGAIN. To facilitate the registration process for these types of service providers, a new generic registration process was created that also includes a "federation of last resort" option that can be used where it is not clear in which federation a service should be registered. This process is currently being piloted with the support of the UK Access Management Federation.

1 Introduction to Enabling Users

eduGAIN is an interfederation service operated and developed by GÉANT that enables users to securely use services from another country and/or national identity federation. The eduGAIN service has grown significantly in coverage since it was launched during the GN3 project in 2011, and now represents 38 national federations worldwide (+10 federations since the last Enabling users report [EUE]), with an additional eight candidate federations currently seeking to join.

The Enabling Users task works directly with research communities and other eduGAIN services operators to address the challenges they encounter when using federated identity management, thereby promoting the increased use of eduGAIN.

All participants in the Enabling Users task are directly responsible for running their respective NRENs' identity federations, including IDEM in Italy, DFN-AAI in Germany, HAKA in Finland, FER in France and SWITCHaai in Switzerland, and are therefore experts in the field of federated identity management.

The main objective of the task is to act as an expert partner for large pan-European research projects and other service providers linked to GÉANT's services (e.g. cloud providers working with SA7) that have a requirement for AAI. Its underlying goal is to increase the practical use of eduGAIN, while extending interfederation technology and AAI functionalities based on real user requirements.

International research communities as well as cloud providers benefit from eduGAIN in particular because their members and users are based at different institutions located in different countries. Thanks to eduGAIN, users can access web services through their home institution user account wherever they happen to be. eduGAIN-enabled services also benefit from the existing legal and technical framework it provides.

Building on the experience of the pilots in GN3plus and AARC, another related objective is the creation of a knowledge database in which research communities and other potential users could find information on the technical and organisational aspects of running services in eduGAIN.

2 The Enabling Users Approach

The Enabling Users task collaborated with two international user communities in setting up two structured, replicable use cases aimed at benefiting the wider research and education community. Expertise was also provided to several other user communities, cloud providers and e-infrastructure providers that expressed an interest in using eduGAIN. To assist in this, a simple step-by-step registration process for Service Providers was developed specifically for cloud providers and research communities that do not have a relationship with any existing eduGAIN member federation.

Past experience (GN3plus) had shown that some of the issues with the use of eduGAIN stem from the way Identity Providers within the eduGAIN member federations are operated. These may include configuration errors, as well as difference in interpretations of policy and law, which were also addressed.

This report highlights the work of the GÉANT Enabling Users task in the GN4-1 project. The task's main objectives are:

- To act as an expert partner for research communities wishing to use federated login as provided by the eduGAIN interfederation service;
- To help enable several of GÉANT's own services for the use of eduGAIN, working with a dedicated task in SA5, AAI for GÉANT;
- To build an eduGAIN knowledge database for sharing technical documentation focused particularly on supporting the needs of user communities;
- To generally promote the increased use of federated login and eduGAIN.

Close collaborations with two European research communities, as well as its work providing basic support for several other communities and cloud providers, enabled the task to gain a better understanding of the technical and organisational needs of these eduGAIN users and the issues they face. Understanding these needs then influences platform innovation for eduGAIN and associated tools.

Work on these complex use cases highlighted that many of the issues encountered can technically be solved today, but that the deployment of the needed solutions, as well as the decision-making process preceding this, often take longer than expected. While identity systems operators must undertake to improve matters in this respect, many research communities also first need to familiarise themselves with the underlying concepts of federated identity management and Authentication and Authorisation Infrastructures (AAI) before they can gain a clear understanding of their advantages and limitations. Close collaboration with AARC is therefore necessary in this area.

3 Collaborations with Research Communities

3.1 CLARIN

3.1.1 Description

"*CLARIN is the Common Language Resources and Technology Infrastructure, which provides easy and sustainable access for scholars in the humanities and social sciences to digital language data (in written, spoken, video or multimodal form), 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. Tools and data from different centres are interoperable, so that data collections can be combined and tools from different sources can be chained to perform complex operations to support researchers in their work.*" (Source: [CLARIN])

CLARIN was one of the first international research projects to use federated identity management for its highly distributed community. It also built the "CLARIN Service Provider Federation" (SPF, cf. [CLARIN-SPF]) as an "organisational proxy" in order to facilitate the registration of the CLARIN SPs with all relevant federations. CLARIN operates its own SAML2 metadata stream for this purpose.

3.1.2 Use Case

At the time of writing, the CLARIN research community operates 28 SAML Service Providers in at least nine countries. The two main issues for CLARIN in the context of eduGAIN have been insufficient IdP coverage and attribute release. The number of IdPs in eduGAIN has substantially grown in the past few years, mostly due to an increasing number of federations changing their IdP opt-in policy for eduGAIN to an opt-out policy. This means that IdP coverage is now less of a problem for CLARIN than previously. To mitigate the IdP coverage issue CLARIN also initially joined several national federations with all its services. As a result, it has had to deal with several different eduGAIN member federations, which have different processes and policies. This means that increasing CLARIN's IdP coverage has also meant an increase in the overall complexity of its management.

While the IdP coverage issue for CLARIN has at least partially been resolved, the attribute release issue still persists. CLARIN services usually require some attributes for a user, such as a persistent identifier (i.e. eduPersonPrincipalName) and a user name. Unfortunately, Identity Providers often fail to release these attributes to services, which then will cause an error message to be displayed to the user, as the services require at least one unique identifier attribute to allow the user access.

This problem is not specific to CLARIN but affects almost all eduGAIN services, in some cases including services within the same federation. There are various reasons why Identity Providers may fail to release attributes. In some countries it may be due to data protection laws being applied very strictly, while in others IdP administrators may not be given the tools and instructions they need to manage attribute release to thousands of services in scalable ways.

Two available measures to improve attribute release in a scalable and secure manner are the adoption of the GÉANT Data Protection Code of Conduct and implementation of the REFEDS Research & Scholarship entity categories [GN4-1_D9.4]. 23 of 28 CLARIN Service Providers support one or both entity categories. It is also helpful to make federation operators and IdP administrators aware that their end users, i.e. researchers, sometimes struggle to access the services they need for their work. Additionally, when a user faces an error message due to insufficient attributes, this message can be used to monitor the problem and ideally inform the relevant persons about the issue so that it may be properly addressed.

3.1.3 Collaboration Objectives

The objectives agreed with CLARIN were mainly platform enhancements focusing on eduGAIN, including:

- Defining a long-term agreement/solution/process between CLARIN ERIC and eduGAIN member federations to publish the CLARIN SP Federation metadata in eduGAIN.
- Using CLARIN's experience and existing concepts to create an eduGAIN Attribute Release Check tool that allows eduGAIN IdP administrators to verify if their IdP is correctly releasing user attributes based on the GÉANT Data Protection Code of Conduct [CoCo] and REFEDS Research & Scholarship entity categories.
- Based on information provided by CLARIN, to notify the identity federations with the highest numbers of users that have a requirement to use CLARIN services but are unable to do so. The organisations that operate the IdPs that have many CLARIN users but fail to properly release user attributes to CLARIN services are then contacted through the federations and supported to enable attribute release.
- Evaluating concepts and solutions to better handle cases where a user with too few attributes attempts to access an eduGAIN service. These solutions should be piloted with at least two CLARIN Service Providers. A guideline should then be issued and published in the eduGAIN wiki providing general instructions and recommendations for eduGAIN service operators.

3.1.4 Expected Benefits

CLARIN ERICS's long-term agreement with eduGAIN provides it with a sustainable and efficient solution to manage metadata and membership for its many Service Providers via an existing eduGAIN member federation.

The remaining objectives on attribute release all not only target CLARIN's current main issue with federated login and eduGAIN by mitigating and actively reducing the attribute release problem but also provide improvements for the community at large. Many other similar services and users at campuses within eduGAIN therefore benefit from the devised solutions resulting from the collaboration with CLARIN.

3.1.5 Achieved Results

The German DFN-AAI federation volunteered to continue publishing CLARIN's SP Federation metadata to eduGAIN. At the time of writing DFN's management has just signed the agreement drawn up to this effect which CLARIN ERIC is also expected to sign soon.

Work on an eduGAIN Attribute Release Check service was started and a configuration guide already exists [Attribute Checker]. Feedback and suggestions for improvements from eduGAIN member federations and research communities are currently being incorporated.

Federation operators that are known within CLARIN to be affected by attribute release issues were informed and made aware of the problems in January 2016. They were provided with information about which Identity Providers were affected as well as instructions on what could be done to improve the attribute release for CLARIN. A follow-up email to these federations is planned to be sent out in April 2016 to request feedback on what has been achieved since the first contact. Additionally, at the time of writing, discussions have started on the introduction of a CLARIN-specific entity attribute category that would allow Identity Providers to create attribute release rules specifically for all of CLARIN's Service Providers and scale to future SPs they may introduce.

A proof-of-concept implementation as well as instructions on how to deploy the proposed solution for Shibboleth-based Service Providers (the large majority of CLARIN SPs use Shibboleth) were created. The concept for this solution also benefitted from experience gained by the DARIAH community that has a similar solution in place for attribute checking and mitigation for insufficient attribute release. CLARIN agreed to pilot the proposed solution with two of their Service Providers. Generic documentation (using CLARIN as an example) will be published on the eduGAIN wiki.

3.1.6 Remarks

The GÉANT Data Protection Code of Conduct and the REFEDS Research & Scholarship entity categories, as well as the upcoming eduGAIN Attribute Release Check service, are helpful tools towards solving the attribute release issue. Nevertheless, ultimately this problem can only be resolved with the help and support of the eduGAIN federation operators and campus identity provider operators. However, some of these operators (primarily campus) are understaffed or not yet sufficiently aware of the issue.

3.2 EIDA

3.2.1 Description

“The European Integrated Waveform Data Archive (EIDA) is a distributed data centre established to provide a secure archive for seismic waveform data and related metadata gathered by European research infrastructures [EIDA]. It provides the geoscience research community with transparent access to these archives via a single web portal. EIDA is part of the Earth Plate Observation System (EPOS), which incorporates different earth science communities.

EIDA consists of 10 data centres across Europe (in the Netherlands, Germany, Switzerland, France, Italy, Turkey and Romania) currently storing more than 300TB of data from approximately 140 networks worldwide. In each network, the data is collected by broadband sensors, short period sensors and accelerometers at over 5000 stations. Seismologists can search and download existing datasets via a single web portal. On average, a single user requests around 130 GB of data per day.” (source: [SFH2])

3.2.2 Use Case

“Most of the data is, after registration, freely available via a web portal. Individual users can request data sets and download them via a desktop-client: the data is encrypted and password protected, with a new password required for each data centre. However, most user requests include data sets from multiple nodes, so multiple passwords need to be submitted every time to obtain the data.

The geoscience community is working to enhance their international collaboration and sharing of data. As a subgroup of EPOS, EIDA is looking to expand the seismological data archive through further sites joining. It also considers the interoperability with other Geoscience communities which will be combined under the umbrella of EPOS. The underlying infrastructure and services provided by EIDA therefore need to be scalable, allowing the addition of new services in the future.

The objective is to develop a scalable AAI solution, simplifying access and data downloads while complying with seismological and security standards.” (source [SFH2])

As described in greater detail in [SFH2], the EIDA use-case was initiated and collaboratively worked on together with GN4-1 NA4 Task 2 (“User Project Liaison”).

3.2.3 Collaboration Objectives

The goals of the collaboration with EIDA were to:

- Gather federated identity management requirements.
- Identify and evaluate different approaches that would allow distributing non-public research data using federated login via eduGAIN.

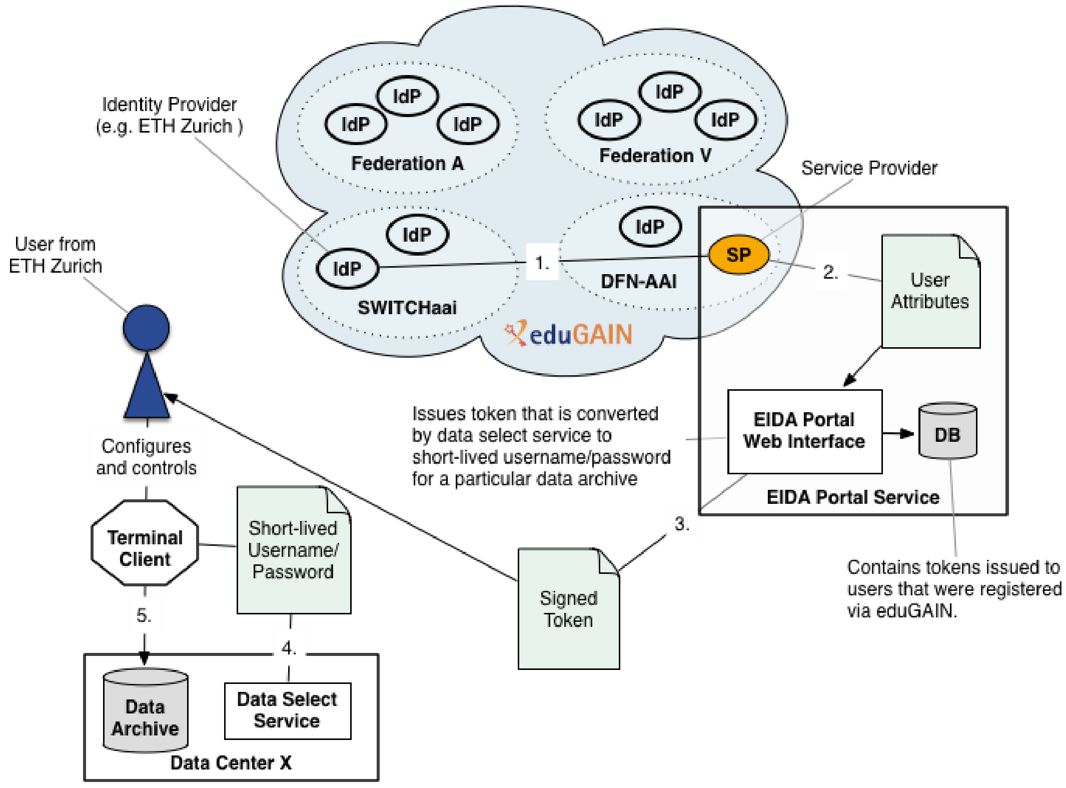
- Assist EIDA in implementing the most promising concept to make use of eduGAIN.
- Identify important organisations from EIDA’s perspective (e.g. organisations with most EIDA users) and ensure that these organisations are enabled to provide federated login via eduGAIN.

3.2.4 Expected Benefits

- EIDA users around the world will be able to access non-public data via eduGAIN.
- Less user management support needed for EIDA and EPOS.
- Less passwords needed for users to access data from one of the EIDA data centres.
- Increased federated identity management know-how and potential reusability for future EPOS/EIDA projects such as the [AlpArray] project.

3.2.5 Achieved Results

Two technical approaches were evaluated. The first of these involved use of a tool [webisoget] that enables downloading of data protected with federated authentication using a terminal client. The tool was successfully tested, however it presents a few drawbacks and disadvantages in terms of security (e.g. user’s organisation credentials stored on disk in plain text). Another approach was therefore used by setting up a portal that allows users to login via eduGAIN, which after successful login remembers the user’s identity including their email address. The user is then provided with a short-lived signed token that is configured in their terminal client to allow it to download public as well as non-public data from EIDA’s data centres. An overview of this process is given in Figure 3.1 below.



1. User from ETH authenticates at ETHZ Zurich IdP, SAML assertion with user attributes is then sent to EIDA SP in DFN-AAI via eduGAIN
2. SP verifies assertion and passes on user's attributes (email, name, organization) to EIDA Portal
3. EIDA Portal issues signed token to user via email or via web (https). User configures token in terminal client.
4. Terminal client asks Data Select service to check and convert signed token into a data center specific short-lived username/password
5. Terminal client uses username/password to query and download data from data center

Figure 3.1: Federated authentication using a terminal client in EIDA

3.2.6 Remarks

The paperwork for EIDA to operate a service in eduGAIN was completed and a beta service registered in the DFN-AAI federation in March 2016. A presentation to EIDA's technical committee is planned for 13 May 2016 to obtain its permission for the service to be rolled out to eduGAIN as well as to all EIDA data centres.

4 Basic eduGAIN Consultancy

The Enabling Users task worked in collaboration with GN4-1 NA4 and GÉANT partners to engage with research communities and service and e-infrastructure operators through:

- Dissemination (e.g. bilateral meetings to introduce the concept of federated login and eduGAIN)
- Consultancy (e.g. to learn their plans regarding eduGAIN and understand in which areas they could need assistance with eduGAIN or joining their local federation).
- Technical knowledge transfer (e.g. how to deploy Shibboleth and adapt web applications).
- Suggestions for improvements (e.g. how to best implement federated login).
- Feedback (e.g. testing federated login from different federations and providing inputs to further improve the login)
- Support (e.g. to debug login problems and propose solutions/workarounds)

The following research communities and operators were contacted:

- OpenAIRE
- eduOpen
- EUDAT
- SeaDataNet
- CESSDA
- Ariadne/Pathenos
- Human Brain Project (HBP)
- European Space Agency (ESA)
- CERN
- EPOS
- ELIXIR
- Square Kilometre Array (SKA)
- Cherenkov Telescope Array (CTA)

5 Additional Results

In addition to supporting and consulting for research communities, the Enabling Users task also developed enhanced tools for eduGAIN which are being included in the eduGAIN technical site [eduGAIN_Tech] as they mature. These developments are described below.

5.1 eduGAIN Access Check Service (EACS)

During GN3plus, the Enabling Users task had already heard from several Providers of eduGAIN services that they encountered difficulties in testing federated login on their own service. Not all SPs have a user account with an eduGAIN Identity Provider, and even when they do, they sometimes prefer to have multiple accounts with different attributes for testing.

To provide service operators with an easy solution to this problem, the Enabling Users task created a Test Identity Provider in eduGAIN that allows operators to access their own service (exclusively) with custom-tailored test identities that have different attributes. At the beginning of GN4-1, this Test Identity Provider was officially announced to the eduGAIN community as the eduGAIN Access Check Service [EACS]. The service is available to all Service Providers in eduGAIN.

5.2 eduGAIN Connectivity Check Service (ECCS)

As a distributed infrastructure with components operated by hundreds of different organisations, eduGAIN's service quality directly depends on the service quality of its components. From an authentication point of view, the most important components involved in federated login are Identity Providers. Some eduGAIN Identity Providers are not configured properly which causes login problems resulting in frustrated end users and increasing helpdesk requests for services operated by research communities and cloud providers. The eduGAIN Connectivity Check Service was created specifically to proactively identify eduGAIN Identity Providers that are not properly configured. It is a simple service that basically imitates a user performing a federated login to a set of well-known test eduGAIN services, using every eduGAIN Identity Provider. The check then analyses what web pages or error messages the eduGAIN Identity Providers return. If something other than a login page is returned, the check fails and the Identity Provider is flagged based on the type of error. The results of the check are published online [ECCS] for every eduGAIN Identity Provider.

The Enabling Users task implemented and piloted the service and brought it into production. Based on the results of the checks, it then informed the operators of those eduGAIN member federations that had at least one Identity Provider that failed the test. This information campaign had some immediate effects in the sense that most federation operators were made aware of any of their Identity Providers that were not configured properly. About one third of all federations replied within the first week that they had fixed the problem or promised to fix the problems. The information campaign will be repeated at the end of GN4-1 in the hope of further reducing the number of misconfigured eduGAIN Identity Providers. The effects of the campaign are shown in the following graphic in Figure 5.1.

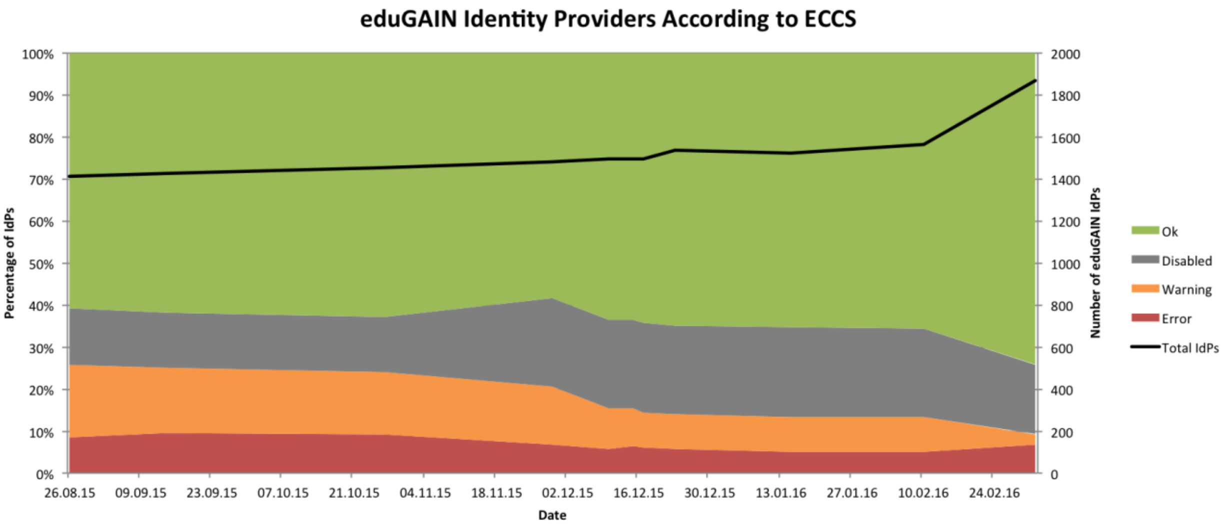


Figure 5.1: Status of eduGAIN Identity Providers in ECCS

Disabled IdPs are those excluded from ECCS checks because they cannot be reliably tested for technical reasons (e.g. due to the hub-and-spoke architecture of the federation) or because their federation operator requested to exclude them from checks.

The ECCS was first presented at the eduGAIN Town Hall meeting on 29 November 2015 in Vienna. It was then announced to all federation operators in eduGAIN (i.e. the eduGAIN Steering Group delegates and deputies) on 15 December 2015. The InCommon federation (US) added about 400 IdPs on 15 February 2016, which substantially increased the number of IdPs in eduGAIN. As shown for this period in Figure 5.1, most of their IdPs were properly configured from the start.

5.3 eduGAIN Attribute Release Check Service (EARCS)

A major problem for Service Providers in eduGAIN is that they sometimes do not receive the requested attributes from eduGAIN Identity Providers. The effect of this is that users (e.g. CLARIN researchers) are not able to access the services they would like to use via eduGAIN. Instead they are often presented with a cryptic error message saying that some of their account information is missing due to insufficient attribute release.

Insufficient attribute release may be due to different reasons. At the campus level, for IdP administrators these range from general data privacy fears linked to releasing information about their users, to lack of know-how on how to technically manage attribute release policies, as well as to lack of awareness of the problem. On a national scale, insufficient advice, training or tools on the part of the federation operator may also have an impact.

To mitigate this issue, some federations as well as research communities such as CLARIN have created attribute release checking services (the best known of which is probably the Interfederation Attribute Test [Interfed_Test] for their Identity Providers to verify which attributes an Identity Provider can release. As these are not official eduGAIN attribute release checking services, but are often federation-specific and lacking some useful features, the Enabling Users task (with inputs and expertise from CLARIN) designed and started the implementation of the eduGAIN Attribute Release Check Service.

The service consists of three Service Providers in different federations, each exporting the SP to eduGAIN. Each of these three Service Providers performs a specific check, the first two for attribute release based on the GÉANT Data Protection Code of Conduct and the REFEDS Research & Scholarship entities respectively, while the third is a general attribute release check. The service allows any user of an Identity Provider to take the test, which can be performed in one minute. The test results are published on the service's web page, which is public [EARC]. This seeks to increase transparency, awareness of the problem and the motivation to correct it within the eduGAIN IdP community.

As of March 2016, a test version of EARCS is available [EARC] and it is expected that pilot operation will start before the end of GN4-1.

5.4 SP Registration Process

Adding a service protected by a SAML Service Provider to eduGAIN means registering it with an eduGAIN member federation [eduGAIN_Status]. This means that services can only join eduGAIN via an existing eduGAIN member federation. Some services (e.g. from research communities, cloud providers, or e-journal providers) have no "natural" relationship to an eduGAIN member federation. In those cases, SP operators often have difficulties in finding out how to register their services with eduGAIN. One of the objectives of the Enabling Users task was therefore to create a "Simple SP Registration Process" (SSPRP) for registering a service with eduGAIN. This process requires a specified eduGAIN member federation to act as federation of last resort for the above-mentioned cases where there is no obvious member federation to register certain SPs.

Following evaluation of several candidate federations and discussions with the most promising candidates, the UK Federation, as the largest eduGAIN member federation, agreed to adopt the role of the federation-of-last-resort in the context of a limited pilot.

The Simple SP Registration Process consists of a web page providing straightforward step-by-step instructions and guidance on how to deploy a SAML Service Provider and how to register it via an existing eduGAIN member federation. The instructions clearly state that SPs should be registered with eduGAIN member federations that are “close” to them. Efforts will be made to work with the SPs to determine which federations these are, so as to minimise the impact on the federation of last resort.

The operators of SPs that do not have such a federation are invited to deploy and register their SP according to the instructions and policies of the UK Federation.

If an SP Operator has to register with the UK Federation (as a last resort), the web page will provide them with the instructions shown in the graphic below (Figure 5.2). These also include completing a web form to provide the service's technical data (SAML2 metadata) and non-technical data (contacts, names and descriptions).

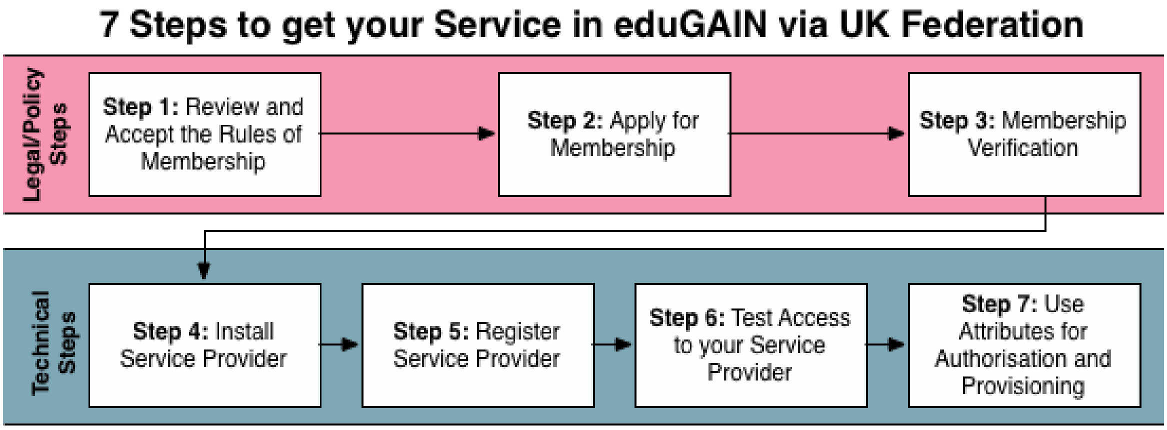

7 Steps to get your Service in eduGAIN via UK Federation
Legal/Policy Steps
Step 1: Review and Accept the Rules of Membership
Step 2: Apply for Membership
Step 3: Membership Verification
Technical Steps
Step 4: Install Service Provider
Step 5: Register Service Provider
Step 6: Test Access to your Service Provider
Step 7: Use Attributes for Authorisation and Provisioning

Figure 5.2: 7 steps to join eduGAIN via the UK Access and Management Federation

When the form is submitted, its content is sent to members of the Enabling Users task. Their role then is to pre-check and, if needed, complete this data before redirecting it to the UK Federation helpdesk on behalf of the SP Operator. By acting as an intermediary at this stage, the Enabling Users task is enabled to improve and optimise the Simple SP Registration Process during the pilot phase, with the aim of doing away with this step once the pilot is completed.

Once the data (including metadata and other information required by the UK Federation) has been submitted to its helpdesk, the UK Federation's normal registration procedures are applied. The Enabling Users task and equivalent functions in future projects will be available as third-level support contact for both the SP Operator and the UK Federation Helpdesk.

As of March 2016, the agreement with the UK Federation for this pilot was finalised and the user guide and active registration form made available online [Join eduGAIN]. In April 2016, the SSPRP was announced and feedback from federation operators was retrieved and incorporated. In the context of two SA7 Cloud Academy workshops, the SSPRP was presented to around a dozen representatives of cloud providers. Some of these it is expected will soon start using the registration guide, providing valuable feedback and inputs for the SSPRP pilot phase.

5.5 AAI Comparison

Federated Identity Management (FIM) is a complex topic, and given that there are multiple Authentication and Authorisation Infrastructures (AAI) using FIM technologies to provide FIM services, it is often difficult for research communities to know what to expect from these infrastructures and in what areas they differ, as well as how they could best make use of them. To shed some light on some of the better known AAIs, FIM services and FIM technologies, and to support GÉANT User Liaison teams, the Enabling Users task drew up its findings in this area in a document entitled "Comparison of Authentication and Authorisation Infrastructures for Research" [AAI White Paper].

The AAI comparison document sets out the characteristics and provides a simple overview and comparison of the general-purpose infrastructures, services and technologies used in research and education which are often misconstrued. The document was written with the help of key experts from the e-Infrastructures, services and technologies it describes. It provides interested research communities with information about some of the most basic aspects of the AAIs, service and technologies as well as practical answers to questions such as how to make use of a particular service or technology and where to get support.

The document has also been published in the eduGAIN wiki [eduGAIN-Wiki] and will be distributed to relevant community mailing lists, such as the FIM4R list.

5.6 Training Events

Members of the Enabling Users task were also involved as contributors and presenters in a number of training events aimed at federation operators and Service Provider operators of research communities, as described in the sections below.

5.6.1 TNC 2015 Attribute Release Training

Insufficient attribute release by Identity Providers is one of the main problems for operators of Service Providers and one which is not limited to eduGAIN. A few identity federations also have an attribute release problem within their own national federations. In the context of eduGAIN, this issue is exacerbated by the higher number of Identity Providers affected by it and the wider heterogeneity of Identity Providers and their policies.

For this reason, the Enabling Users task, in collaboration with SA5 Task 1, Harmonisation, also helped organise an attribute release training session at the TNC15 Networking Conference in Porto. The goal of the training was to make participants aware of the issues, show how attribute release is managed in three federations with different architectures and generally encourage the implementation of attribute release based on entity categories such as the GÉANT Data Protection Code of Conduct [CoCo] and REFEDS Research & Scholarship [RANDS].

Even though the training was held before the start of the conference, it was attended by 50 participants representing 26 federations. In February 2016, 15 (58%) of the 26 federations whose members had attended the training had one or more entities in their metadata which declared support for the GÉANT Data Protection Code of Conduct or the REFEDS Research & Scholarship entity categories.

5.6.2 AAI Workshop for Service and Resource Providers

In collaboration with ELIXIR, DARIAH and the AARC project, the Enabling Users task organised a two-day training on 15-16 March 2016 in Manchester (UK) providing a general introduction on how to deploy and use Shibboleth Service Providers. The training was targeted at administrators who were new to federated login.

On day one, a basic overview of the key principles of federated identity management, SAML, Shibboleth (the most popular SAML implementation in eduGAIN) was provided. Participants later took part in a hands-on session in which they installed and configured the latest Shibboleth Service Provider themselves on a virtual machine image that they were provided with.

On the second day, the training participants were split in two groups, dealing with topics that were relevant to operating federated services either in ELIXIR or DARIAH respectively. Based on the positive feedback received and on demand, another training with the same format is planned to take place in June 2016 in Germany, specifically for the DARIAH research community.

5.7 Knowledge Database

The eduGAIN Wiki [eduGAIN_Wiki] was set up during the GN3plus project to serve as a database of eduGAIN knowledge. During GN4-1, its structure was improved to better serve the different target groups. It was also extended to include several topics that were found to be relevant for operators of eduGAIN services. These topics include:

- How to use an SAML Proxy (SimpleSAML PHP) to enable federated login from potentially hundreds of different eduGAIN Identity Providers to applications that support SAML login only, with one or very few Identity Providers (e.g. because they cannot process SAML2 metadata files or because they don’t include Identity Provider discovery).
- A recommendation on the best SAML attributes to use to identify a user in an eduGAIN context.
- Extended technical Best Current Practices for joining eduGAIN as a federation.
- Processing SAML2 metadata using the increasingly popular tool pyFF.

5.8 Dissemination of eduGAIN to Scientific Communities

To increase the deployment of eduGAIN and federated identity technology in general as well as awareness of certain issues (support for interfederation/eduGAIN and insufficient attribute release of Identity Providers) among Campus administrators, presentations were given at the following events:

- TNC 2015, 15-18 June 2015, Porto, Portugal.

- EGI Community Forum 2015, 10-13 November 2015, Bari, Italy.
- JRES, 8-11 December 2015, Montpellier, France.
- AAI Workshop for Service and Resource Providers, 15-16 March 2016, Manchester UK (in collaboration with AARC, DARIAH and ELIXIR).
- Scheduled: EUNIS, 8-10 June 2016, Thessaloniki, Greece

The task was also represented at:

- 6th RDA Plenary, 23 - 25 September 2015, Paris, France.
- FIM4R/REFEDS/eduGAIN Town Hall meeting, 30. November-1. December 2015, Vienna, Austria.

6 Conclusion and Recommendations

Separation between AARC and GÉANT is complex to manage but delivers some advantages. AARC also includes research communities and libraries alongside NRENs, making it easier for it to pilot solutions over multiple infrastructures compared to the experience in GN3plus. However, it is important for GÉANT to have close contact with user communities after the end of a pilot to ensure that recommendations from AARC transition to service and become operationally sustainable. Without in-depth subject matter expertise, it is also challenging to determine on a case-by-case basis whether a community should be served as a support case by GÉANT or adopted in AARC as a pilot.

Experience has shown that AARC and GÉANT should work very closely together, supporting each other with training material and results, to avoid the need for duplications. This maximises the number of research communities who can be supported, while keeping a clear separation in scope between working with existing solutions and platform-driven development, and being able to improve eduGAIN based on user-driven input from AARC.

In GN4-1 the Enabling Users task focused on two new collaborations with the CLARIN and EIDA research communities, while still maintaining contact with the five research communities that had previously been supported through close collaborations in the GN3plus project [EUE].

In addition to working closely with CLARIN and EIDA, the Enabling Users task provided other research communities with consultancy, expertise and eduGAIN know-how on a smaller scale. The task was also involved in organising and providing training specifically targeted at the technical personnel of two large research communities, ELIXIR and DARIAH. In this area, to avoid duplication of effort and materials as well as the need for AARC to engage separate expertise covering the areas of GÉANT services, it was considered more efficient for GÉANT to directly provide training to AARC’s target communities.

Work was also started on a number of new tools aimed at mitigating the main problems encountered by eduGAIN users, i.e. configuration and attribute release issues. Both of these do not require extensive user-driven innovation to be addressed, but they have a high impact, as where they are not resolved they endanger the service’s quality as perceived by both users and service operators. Direct contact with user groups enables GÉANT to identify, monitor and address such issues pragmatically while more long-term approaches are developed in projects such as AARC. As the team works in close contact with eduGAIN operations, these tools can be integrated in the eduGAIN technical suite quickly, simply and efficiently rather than needing a complex handover between separate projects.

References

[AAA] AAA Study, Licia Florio (TERENA) et al, 2012 http://www.terena.org/publications/files/2012-AAA-Study-report-final.pdf

[AAI White Paper] http://www.geant.org/Resources/Documents/Comparison-of-AAIs-for-Research_White-Paper_v1.0.pdf

[AARC] Authentication and Authorisation for Research and Collaboration project https://aarc-project.eu/

[AlpArray] http://www.alparray.ethz.ch/home/

[Attribute_Checker] https://wiki.edugain.org/How_to_configure_Shibboleth_SP_attribute_checker

[CERN] http://www.cern.ch/

[CLARIN] http://www.clarin.eu/content/about-clarin

[CLARIN-SPF] https://www.clarin.eu/content/service-provider-federation

[CoCo] GÉANT Data Protection Code of Conduct http://www.geant.net/uri/dataprotection-code-of-conduct/v1

[DARIAH] http://www.dariah.eu/

[EACS] eduGAIN Access Check Service, https://access-check.edugain.org

[EARC] http://earc.eduid.hu/

[ECCS] eduGAIN Connectivity Check Service, https://technical.edugain.org/eccs/

[eduGAIN] http://services.geant.net/edugain/Pages/Home.aspx

[eduGAIN_Status] https://technical.edugain.org/status

[eduGAIN_Tech] http://technical.edugain.org/

[eduGAIN_Wiki] https://wiki.edugain.org/

[EIDA] http://www.orfeus-eu.org/eida/eida.html

[E-INFRA-7] http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/2141-einfra-7-2014.html

[ELIXIR] http://www.elixir-europe.org/

[ESA] http://www.esa.int/ESA

[EUE] "Towards Horizon 2020 - The Enabling Users Experience" (GN3plus deliverable D9.4, DS5.5.1), Lukas Hämmerle et al. December 2014

[FIM4R] "Federated Identity Management for Research Collaborations", 2012, CERN-OPEN-2012-006, Broeder, Daan (MPI) ; Jones, Bob (CERN) ; Kelsey, David (STFC) ; Kershaw, Philip (STFC) ; Lüders, Stefan (CERN) ; Lyall, Andrew (EBI) ;

Nyrönen, Tommi (CSC) ; Wartel, Romain (CERN) ; Weyer, Heinz J (PSI, Villigen) http://cds.cern.ch/record/1442597/

[GÉANT] GÉANT, pan-European research and education network http://www.geant.net/

[GN3plus] "Multi-Gigabit European Research and Education Network and Associated Services (GN3plus)", DANTE (Matthew Scott, Niels Hersoug), 2012

[GN4-1_D9.4] http://www.geant.org/Projects/GEANT_Project_GN4-1/Documents/D9-4_Report-on-Harmonisation-Development-and-Pilots.pdf

[Interfed_Test] https://attribute-viewer.aai.switch.ch/interfederation-test/

[Join_eduGAIN] https://wiki.edugain.org/How_to_Join_eduGAIN_as_Service_Provider

[Policy] eduGAIN Policy

http://www.geant.net/service/edugain/resources/Pages/home.aspx

[RANDS] REFEDS Research & Scholarship entity category

https://refeds.org/category/research-and-scholarship

[RDA-FIM] https://rd-alliance.org/groups/federated-identity-management.html

[REMS] Resource Entitlement Management System (REMS)

http://rems.csc.fi/

[SFH2] D4.1 "Support for Horizon 2020 – Handbook on how GÉANT will Support Horizon 2020 Projects", V. Capone (GÉANT), J. Dyer (GÉANT), R. Sabatino (GÉANT), B. Weber (GÉANT), February 2016

[webisoget] http://staff.washington.edu/fox/webisoget/

[Wiki_eduGAIN] http://wiki.edugain.org/

Glossary

AA	Attribute Authority
AARC	Authentication and Authorisation for Research and Collaboration
AAI	Authentication and Authorisation Infrastructure
eduGAIN	educational Global Authentication Infrastructure
CLARIN	Common Language Resources and Technology Infrastructure
COCO	GÉANT Data Protection Code of Conduct
DAC	Data Access Committee
DARIAH	Digital Research Infrastructure for the Arts and Humanities
EC	European Commission
EGA	European Genome-phenome Archive
EIDA	European Integration Waveform Data Archive
ESA	European Space Agency
EO	Earth Observation
EPOS	Earth Plate Observation System
ESFRI	European Strategy Forum on Research Infrastructures
FaaS	Federation as a Service
FIM	Federated Identity Management
FIM4R	Federated Identity Management for Research – a forum for AAI providers, e-Infrastructures and users
FIMig	Federated Identity Management Interest Group of the RDA
FP7	Seventh Framework Programme for Research and Technological Development
GN3plus	GÉANT 3 Plus Project
IdM	Identity Management
IdP	Identity Provider
JRA	Joint Research Activity
LoA	Level of Assurance
MDS	Metadata Distribution Service
NA	Networking Activity
NREN	National Research and Education Network
PKI	Public Key Infrastructure
RANDS	REFEDS Research & Scholarship Entity Category
REFEDS	Research and Education Identity Federations
RDA	Research Data Alliance
SA	Service Activity
SA5	Service Activity 5 "Application Services"
SAML	Security Assertion Markup Language
SCI	Security for Collaborating Infrastructures SSH Social Science and Humanities.

SP	Service Provider
SSO	Single Sign-On
VC	Video Conference
VO	Virtual Organization