

International authentication with eduGAIN ou comment fédérer les identités au niveau international

Olivier Saläun

RENATER

c/o CRI Campus de Beaulieu, Bat 12 D

263, Avenue du Gal Leclerc CS 74205

35042 RENNES Cedex

France

Lukas Hämmerle

SWITCH

Werdstrasse 2

8021 Zurich

Switzerland

Résumé

An identity federation consists of multiple organisations (e.g. universities and research institutes) that agree to use a common infrastructure for authentication and authorisation. eduGAIN is a global interfederation service that interconnects multiple identity federations, both technically and legally. It allows a user from one identity federation to access services in another identity federation. eduGAIN aims at connecting all SAML-based academic identity federations world wide. More than half of all known academic identity federations are already connected to eduGAIN. However, many institutions (e.g. universities) of the participating federations have yet to make the necessary adaptations to become part of eduGAIN. The adoption on institution level currently is slow due to complicated policy and data privacy issues. In particular the release of user information from one jurisdiction to another one.

RENATER has already signed the eduGAIN constitution and implemented the necessary technical changes on the federation level. Therefore, institutions that are members of the Fédération Education-Recherche can now benefit from eduGAIN as well. By joining eduGAIN, Service Providers can offer their services to a wider audience in the research and education community; this of course requires a few technical changes e.g. regarding the discovery service and the user attribute requirements. The same applies for Identity Providers that want to offer their users access to services operated in eduGAIN; they for example will have to fine-tune the attribute release policies and add a few internationally supported attributes. The article also outlines and explains the necessary steps which service operators and institutions need to take when getting eduGAIN-enabled.

Mots-clefs

SAML, Federation, Interfederation, Identity, Authentication, Authorisation, Shibboleth, Single Sign-On

1 Introduction

In a first part this article describes the history, goals and challenges of the global federated authentication and authorisation service eduGAIN¹. This is followed by a description of how the French identity federation (Fédération Education-Recherche) was integrated into eduGAIN and how web resources can make use of interfederated authentication and authorisation in France.

¹ <http://www.edugain.org>

1.1 Background

Around the millennium, the usage of web based e-learning applications and access to digital content increased considerably at higher education organisations in Europe. More and more user accounts had to be managed and users needed to remember an increasing number of passwords to access their e-learning applications. While these authentication problems could be easily solved within a single organisation by using the organisation's main user directory as authentication source for all applications, this approach is not applicable in case users from multiple organisations need to access an application. In order to solve this, identity management has to become federated identity management.

Organisations participating in so-called (identity) federations agree on a common set of policies as well as technical and legal standards. This creates the necessary trust that allows their users to access content and services within a federated identity management system. The main benefit for users: a single password is sufficient to authenticate at all services participating in the federation. Thanks to Single Sign-On, they also have to enter their password only once per web browser session to authenticate at multiple services. In the case of academic federations, these services not only include e-learning applications and library resources but also services offered by commercial companies. For example online shops that allow students to buy discounted products or download free software.

Thanks to federated identities, service operators also get additional information about their users. Name, email address and affiliations and other data can be released by a user's organisation in form of attributes. These attributes then also allow performing fine-grained access control. Subsequently, administrators of federated services can dispose of authentication and identity management issues and focus on the operation of the service itself.

2 About eduGAIN

Most identity federations were created within a single jurisdiction for a single country. However, shortly after the first federations were created, their operators envisaged to interconnect individual national identity federations, thus creating a so-called "Interfederation" infrastructure. This eventually led to the global interfederation service eduGAIN.

The primary goal of eduGAIN is to facilitate international collaboration between students, researchers and lecturers by interconnecting national identity federations to exchange trusted identity information. To achieve this goal, eduGAIN provides the required infrastructure (i.e. metadata data service to exchange federation SAML metadata) and frameworks (i.e. constitution, profiles) that cover technical, legal and policy aspects to facilitate cross-border interoperability.

2.1 History of eduGAIN

Even though the production operation of the interfederation service eduGAIN officially started in April 2011, its history goes back to the year 2004. At that time, no academic federation had yet a production federation but there were already a few pilot federations in Europe and the US. Only in 2005, the first academic identity federations started their production service². Since then numerous federated identity management infrastructures around the world were successfully established. Most of them are operated by the organisations that run the national research and education network (NREN) for a country.

With the advent of academic identity federations and the increasing internationalization of research, it was becoming obvious that the benefit of these identity federations could be increased if they were all connected. This would allow using digital identities from one identity federation to access services in another. Therefore, the GÉANT2 project in 2004 proposed an architecture for a new confederation called "GÉANT Authorisation INfrastructure for the research and education community", which led to the name "eduGAIN". Back then the few existing identity federations still were using different protocols and software implementations to exchange identity information. Some of them were proprietary and incompatible to each other. Consequently, the initial eduGAIN architecture proposed so-called "bridging elements" whose purpose was to translate identity assertions to a common language. The Security Assertion Markup Language (SAML) protocol was then chosen as lingua franca.

GÉANT2 succeeded to demonstrate that this concept technically worked. However, it was already overtaken by reality by the end of the project when it became obvious that SAML2 would become the de-facto standard for academic identity federations. Therefore, eduGAIN's architecture was adapted during the successor project GÉANT3. It became

² The first two were the Swiss SWITCHaaI and the Finish HAKA federations

simpler and more transparent because the bridging elements were obsolete. eduGAIN was transformed from a confederation model to an interederation service. GÉANT3 also showed that the policy side of an interederation services is very challenging. Therefore, a lot of effort was put into creating different policies and profiles to create a robust framework.

2.2 Architecture

An identity federation consists of entities from different organisations. In the context of SAML-based identity federations these entities usually are Identity and Service Providers. An Identity Provider (IdP) authenticates users in order to issue identity assertions about them. A Service Provider (SP) consumes identity assertions to log in users to applications like an e-learning or library web application. Therefore, (SAML) identity assertions always flow from an IdP to an SP. A federation is technically described by the IdPs and SPs listed in a SAML metadata file of that federation.

Identity federations use different architectures as is shown in Figure 1. Federations A and B are so called “full-mesh” federations (80% of all federations³) where each SP and IdP are visible to each other. In this model each participating organisation (i.e. a university) operates an own IdP with an own login service. Federation C and D are so-called hub-and-spoke federations where a central “hub” is involved in each login. In the architecture of Federation C (5%), this hub has a single central login page, which means that the hub needs access to all user directories of the organisations participating in this federation. The architecture of federation D (15%) is similar but each participating organisation operates its own Identity Provider with an individual login service. Identity assertions flow always via the hub that proxies and often extends the assertions before they are passed on to the SP.

eduGAIN itself uses a full-mesh architecture. One key component is the so-called Metadata Distribution Service (MDS). The MDS aggregates all SAML metadata files from the participating entities and republishes the aggregated metadata for consumption by the participating federations. Some federations like Federation A add all their entities to eduGAIN metadata. However, most federations (B, C, D) will only add a subset of their entities to eduGAIN using an opt-in procedure where each SP and IdP decides for itself whether it wants to be part of eduGAIN. Only entities that are listed in the eduGAIN MDS metadata can directly communicate with each other.

A federation joining eduGAIN has to apply some changes to become an eduGAIN member. Technically, this mostly concerns the exchange of SAML metadata and the release of user attributes. Often also the policies of a federation have to be adapted. Implementing these changes requires time and efforts. Additionally, the service operators of a national federation need first to be carefully informed about interederation. About its new opportunities but also about the – mostly data protection - risks.

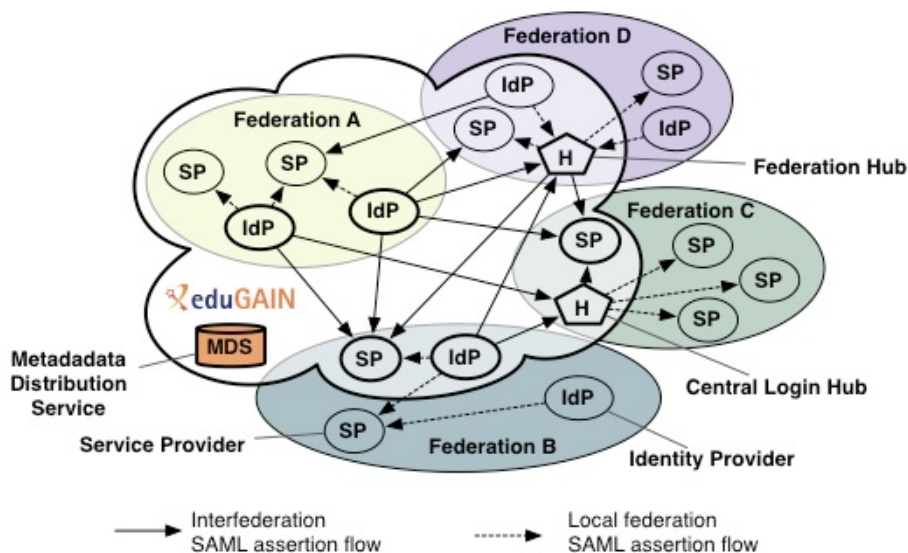


Figure 1 - eduGAIN's Technical Architecture

³ As of August 2013

On a technical level, eduGAIN requires the participating entities to publish and consume SAML metadata according to the eduGAIN Metadata Profile. The protocol used to exchange identity assertions is SAML2 using the SAML2int⁴ profile. However, eduGAIN's policy framework is flexible enough to also allow other protocols in the future.

2.3 Advantages of eduGAIN

In brief, eduGAIN allows to easier share resources and services cross-country. Similar like eduroam⁵ provides Internet access abroad, its sister eduGAIN provides access to services and content operated abroad. This allows for example collaborating in international research projects or cooperatively generating and using e-learning content with a partner university in another country.

Organisations that enable eduGAIN allow their researchers and students to access additional services operated by the higher education and research community. This provides them with additional opportunities for collaborations and cooperations.

As operator of a service, enabling eduGAIN allows making use of trusted federated identities across country borders. It basically widens the group of users that could be granted access with their federated identity. Instead of creating yet another account for users from abroad and basically do the identity management work yourself, using eduGAIN allows outsourcing the identity management.

2.4 Target Applications

There are basically three types of applications that will most benefit from eduGAIN.

The first group is the research applications. Large research projects today are often international projects with participants from all over the world. In order to collaborate and cooperate, the research community usually needs access to a set of common services. Most of these services need authentication and authorization. It is therefore obvious that these services benefit from eduGAIN's authentication and authorisation capabilities.

A second group consists of e-Learning applications. Operating a stable e-learning platform and providing high-quality e-learning content is very costly. Therefore, it can make sense for universities to cooperate in this area. There are already examples of such cross-country collaborations⁶. It is likely that more are to come.

A third group comes in form of the content providers that offer the academic community e-journals and papers. In the past few years content publishers like Springer, ScienceDirect, Thomson Reuters and many more have had to join multiple identity federations to offer their services to the academic community. This is a time-consuming process for all involved parties, the federation operators as well as the publishers. eduGAIN will allow them at some point to join only one federation, while still allowing users from different countries to access their services.

2.5 eduGAIN Adoption

As of August 2013, 34 national identity federations are known⁷ in the higher education and research worldwide. Since April 2011 already 18 national identity federations (52%) have signed the eduGAIN declaration and agreed to the eduGAIN constitution⁸. Two of them being non-European federations. 7 (20%) federations are in the process of joining and 5 (15%) are candidates to join. Even though GÉANT has been 50% funded by the European Commission, all SAML-based identity federations that "primarily serve the interests of the education and research sector" can apply for membership according to the eduGAIN constitution.

Approximately 10% of all entities in the participating federations already opted-in and applied the necessary changes to participate in eduGAIN. It must be noted that it is not reasonable for all SPs and IdPs to become part of eduGAIN. Some SPs are for example used only within a single organisation. Therefore, the adoption rate will not reach 100%.

⁴ SAML2Int profile: <http://saml2int.org/>

⁵ eduroam allows users from the education and research community to obtain Internet connectivity world-wide: <http://www.eduroam.org>

⁶ Example of an international e-learning collaboration "Virtual Campus Hub", <http://www.virtualcampushub.eu/>

⁷ Academic Identity Federations known: <https://refeds.terena.org/index.php/Federations>

⁸ eduGAIN Member Federations: http://www.edugain.org/federation_status.php

2.6 Obstacles and Difficulties

Federated identity management per se is non-trivial. Connecting multiple national federations in different jurisdictions via an interfederation service results in some complex legal and policy problems.

2.6.1 Chicken and Egg Problem

Each individual federation initially faces a chicken and egg problem. As long as there are not many interesting services (SP) yet to use, only few organisations invest time and money to federate their user accounts. On the other hand, if there are only few federated users, the motivation for services to get federated is low. Usually it takes rather long (years) for an identity federation to reach a tipping point after which a critical mass of SPs and IdPs automatically accelerates the adoption rate. The same is happening with eduGAIN. As is described in a next chapter, additional efforts have to be invested by SP and IdP admins to become eduGAIN-enabled. Spending these efforts is initially not very attractive with only few entities being part of eduGAIN. Unfortunately, there is also no global killer application that could motivate IdP administrators to invest the upfront effort sooner. Even though there are quite some research communities that operate services which could benefit from eduGAIN, there often are only few participants of a particular community at one organisation (e.g. a university). Therefore, it is difficult for these few people to convince their organisation to participate in eduGAIN.

2.6.2 Opt-In Process

The process of becoming eduGAIN-enabled varies from federation to federation. This often has legal and technical reasons. In some federations it is necessary that organisation first sign a document to enable eduGAIN, while in others this is not necessary because their normal federation policy already allows this.

Regardless if a legal opt-in is needed there still are technical steps required to become eduGAIN-enabled in most federations. Some identity federations using a “hub-and-spok” architecture (Federations C and D in Figure 1) have a central component that is involved in every authentication process. With such an architecture, ticking a checkbox might be sufficient to enable eduGAIN support for a service or an organisation. However, the majority of federations - including the Fédération Education-Recherche - use a full-mesh architecture (Federations A and B in Figure 1), which is completely distributed. This federation architecture requires a bit more efforts from the participating organisations and services because there is no central component involved where interfederation support can centrally be enabled.

The German DFN-AAI federation decided in 2011 to opt-in all their participating organisations into eduGAIN directly. They added all IdPs to the eduGAIN metadata. However, not all organisations were aware that they were listed as eduGAIN-enabled. And there also still were additional technical adaptations required for each IdP to be fully interoperable with eduGAIN. Many organisations however had not performed these adaptations yet. This then resulted in problems when users of such organisations tried to access an eduGAIN-enabled service. These problems backfired on the federation operator, which then made them also choose an opt-in procedure that ensured that only those organisations are exposed to eduGAIN, which also completed the technical steps.

Overall, even though the legal and technical efforts are often rather small for organisations and service operators, they slow down the eduGAIN adoption. Additionally, the people who decide if and when these efforts are invested first have to be properly informed about eduGAIN and its advantages.. Therefore, enabling organisations by default for eduGAIN is not an option for the time being for most federations.

2.6.3 Attribute Release and Data Privacy

Depending on the jurisdiction of a federation, different data privacy issues arise when an IdP releases user information. This is especially true if information is released to non-EU countries. The effect is that many administrators of organisations participating in eduGAIN are hesitant to release information in form of attributes about their users. The attributes however are what make federated identity management valuable. Not receiving all the required attributes of users is problematic for applications protected by a SP. This is also a problem because the user himself often has no influence over what information is released about him.

In order to handle the data privacy concerns, a document called the GEANT Data Protection Code of Conduct⁹ was created. This document serves as a brief declaration stating that the SP operator will follow certain processes and rules when processing user data received via eduGAIN. The goal is to increase the confidence that an SP handles user information properly. That a SP complies with the Code of Conduct is expressed in the eduGAIN metadata and therefore visible for everybody. Administrators can then configure their IdPs to release user attributes based on the existence of the Code of Conduct.

2.7 eduGAIN and Fédération Education-Recherche

As a national federation operator that is member of eduGAIN, RENATER is generating the list of French IdPs and SPs that wish to be part of eduGAIN. This list of SAML entities is a SAML metadata file, provided to eduGAIN operation team for inclusion in the global eduGAIN metadata file. From an SP/IdP administrator's point of view, joining eduGAIN is just checking a box on the same federation registry that is provided by RENATER to join the Fédération Education-Recherche¹⁰.

One may wonder why RENATER does not automatically include all its federation SAML entities in eduGAIN? As was illustrated by the example of the DFN-AAI federation above, including a SAML entity in the eduGAIN metadata is just one part of the job; the IdP/SP configuration also needs to be adapted to work in eduGAIN (attributes provisioning, release policy, attributes consumption, discovery service) as will be discussed in chapter 2.8.

2.7.1 The Federation Registry

The registration in Fédération Education-Recherche is handled by the so-called federation registry (le guichet de la fédération)¹¹. The registry is a web application that allows SP/IdP administrators to provide technical details about their SAML entities, join the test federation and eventually join the Fédération Education-Recherche. As a result, the federation registry updates the federation metadata file with the SAML entities description. The federation registry also builds attribute filter rules to be used by Shibboleth IdPs, thus ensuring relevant release of user attributes without requiring human intervention at each IdP. RENATER has recently extended its federation registry to also support registration of entities in eduGAIN.

Type	Intitulé et identifiant	Organisme de rattachement	Fédération de Test	Fédération Education-Recherche	eduGAIN
idp	GIP RENATER https://idp.renater.fr/idp/shibboleth	RENATER	✗	✓	✓
sp	CRU - Service de support RT https://support.cru.fr/shibboleth	RENATER	✓	✓	✗
sp	CRU - le site web https://www.cru.fr/shibboleth	RENATER	✗	⚠	⚠
sp	INHA - Serveur de listes https://listes.inha.fr/sympa	RENATER	✗	✓	✗
sp	JRES - Site inscription 2011 https://inscription.jres.org	RENATER	✗	⌚	✗

Figure 2 - RENATER's federation registry

The new federation registry allows declaring the technical details of each SAML entities once and then join one or more federations (see figure 2). Available federations are: Fédération de Test, Fédération Education-Recherche, eduGAIN and local federations. The notion of a local federation is still to be implemented by RENATER; it might be useful for group of institutions that share a set of SPs that should be known/accessible only by these institutions' IdPs.

⁹ Code of Conduct: <http://www.geant.net/uri/dataprotection-code-of-conduct/v1/>

¹⁰ Fédération Education-Recherche: <https://services.renater.fr/federation/index>

¹¹ RENATER federation registry: <https://services-federation.renater.fr/gestion>

The federation registry is accessible by any authenticated user to add a new SAML entity. However, registering this SAML entity in a production federation (including eduGAIN) requires prior approval by an institution's federation contact.

2.7.2 eduGAIN metadata: Workflow and format

As mentioned earlier, the eduGAIN Metadata Distribution Service (MDS) acts as a metadata aggregator. The workflow starts with RENATER's federation registry that generates a metadata file including only French SAML entities that opted in for eduGAIN. This metadata file is then aggregated by the eduGAIN MDS to build a global eduGAIN metadata file. However, this eduGAIN metadata file is not supposed to be published on the eduGAIN web site directly; instead it first is processed again by RENATER to filter out entities and resigning the metadata with RENATER's own x.509 certificate. That resulting eduGAIN metadata file¹² is then ready for consumption by French IdPs/SPs that are eduGAIN-enabled.

The eduGAIN metadata file has been split into two separate files by RENATER: `sps-edugain-metadata.xml` and `idps-edugain-metadata.xml`, each including a type of SAML entities. This reduces the size of metadata files SPs and IdPs need to load, which accelerates download and processing times and mitigates the risk of scalability issues in the long term.

Several extensions for the SAML2 metadata format were developed in recent years. EduGAIN mandates the use of two of these extensions: the SAML2 metadata extensions for registration and publication information¹³ and SAML2 metadata extensions for login and discovery user interface¹⁴. The first extensions allows keeping track of which federation operator registered a SAML entity at what time and it provides the URL to the federation registration policy.

The second extensions concerning the login and discovery user interface has not yet been implemented by RENATER. It aims at improving the user experience of the login page and the discovery service (also known as WAYF service) by including extra information in eduGAIN metadata. In particular the extension allows including new information like:

- the logo of an institution,
- the URL to a privacy statement URL,
- geographic coordinates associated to the SAML entity.

EduGAIN metadata also carries information about the user attributes that an SP requests for a user; this allows IdPs to automatically release the right set of attributes to SPs according to their release policy and the SP's expectations.

2.8 Joining eduGAIN

2.8.1 How a French IdP Can Join eduGAIN

IdP administrators should consider joining eduGAIN to make the user identities useable to access services operated in federations of other countries. These services may first concern only a small fraction of the user population (mainly researchers), but as eduGAIN grows, more services will require that an IdP participates in the inter-federation.

To make an IdP interfederate with eduGAIN, it needs to load the eduGAIN metadata file which is provided by RENATER. This metadata file is signed by RENATER; so one has to download RENATER's CA file and make sure the IdP software verifies the signature of the downloaded metadata file.

eduGAIN defines a set of user attributes that any IdP in eduGAIN should be able to provide given a standard syntax and semantics (see chapter 2.9 for a list of these attributes). IdP administrators should ensure that they are able to provide the recommended set of attributes for their users. Note that because not all values of `eduPersonAffiliation` (and `eduPersonScopedAffiliation`) are supported at an international level, certain values of these attributes should not be released to foreign SPs. This can be achieved with a Shibboleth IdP through attribute filtering.

Within the Fédération Education-Recherche, RENATER provides automatically built Shibboleth attribute filters for categories of services; this allows automatic configuration of attribute filters on the IdP side¹⁵. Equivalent attribute

¹² Downstream eduGAIN metadata file: <https://federation.renater.fr/edugain/>

¹³ SAML2 DRI extensions: <https://wiki.oasis-open.org/security/SAML2MetadataDRI>

¹⁴ SAML2 MDUI extensions: <https://wiki.oasis-open.org/security/SAML2MetadataUI>

¹⁵ Gestion automatique des filtres d'attributs: <https://services.renater.fr/federation/docs/fiches/attribute-filters>

filters are not provided to eduGAIN participants but an alternative is provided to prevent the need to manually configure attribute filters. The EduGAIN metadata includes a structured description of the SPs' attribute requirements; a recent extension of the Shibboleth IdP 2.4 filter rules¹⁶ allows configuring the automatic release of attributes to subsets of SPs.

2.8.2 How a French SP Can Benefit from eduGAIN

As an SP administrator, one should consider joining eduGAIN to provide access to federated services for foreign users whose IdP is (or might be) eduGAIN-enabled. This is especially interesting for international research projects whose users often need to access a common set of services. First, it is required that an SP redirects users to an appropriate Discovery Service, aware of all eduGAIN IdPs, and probably also IdPs of the Fédération Education-Recherche. One can operate an own instance of such a Discovery Service, or the one provided by RENATER can be used, which also supports fault tolerance capabilities¹⁷. To achieve a better integration of the DS service within an application, the embedded DS feature is also an option; for more information about the embedded DS setup, refer to RENATER's documentation¹⁸.

The SP software of course also needs to load the eduGAIN metadata file provided by RENATER. This metadata file is signed by RENATER; as already described above, one has to download RENATER's CA file and then ensure that the SP is configured to verify the signature of the metadata file.

Enabling an application for eduGAIN might require additional changes regarding the identity management of that application, in case it was not already adapted to consume foreign identities. There will be available different attributes for eduGAIN users: supAnn attributes will not be available but instead SCHAC¹⁹ attributes might be used, some values of eduPersonAffiliation won't be available either. These changes might have impacts on the access control management and also on the type of user identifiers that are used in an application.

2.9 User attributes recommended in eduGAIN

The eduGAIN attribute profile document contains the set of user attributes that should be available for each user of an eduGAIN-enabled IdP. The attributes are defined in the eduPerson²⁰ and SCHAC schemas, Other attributes may be used as well on a bilateral basis. Ideally both, the eduPerson and the SCHAC schema will be included in the French SupAnn directory schema.

The list of attributes recommended to provide for eduGAIN users is:

- displayName, common name (cn), mail,
- eduPersonPrincipalName, eduPersonTargetedId,
- eduPersonAffiliation and eduPersonScopedAffiliation,
- schacHomeOrganization, schacHomeOrganizationType.

Because of divergent interpretations of eduPersonAffiliation attribute nomenclatures in different countries, values "staff" and "employee" should not be used unless their semantics have been verified bilaterally. Note also that the values "researcher", "retired", "emeritus", "teacher", and "registered-reader" have been defined in SupAnn 2009²¹ only, but not in eduPerson.

¹⁶ Example of Shibboleth 2.4 IdP complex filter rules:

<https://wiki.shibboleth.net/confluence/display/SHIB2/IdPAddAttributeFilterExamples#IdPAddAttributeFilterExamples-AttributeInMetadata>

¹⁷ RENATER Discovery Service: <https://discovery.renater.fr/edugain/WAYF>

¹⁸ Using the embedded DS: <https://services.renater.fr/federation/docs/fiches/filtreidp>

¹⁹ SCHAC: <http://www.terena.org/activities/tf-emc2/schac.html>

²⁰ eduPerson: <http://middleware.internet2.edu/eduperson/>

²¹ SupAnn: <https://services.renater.fr/documentation/supann/index>

2.10 Conclusion

The goal of eduGAIN is to allow service operators to authenticate and authorise users from the world-wide education and research community. Users can use their own institutional account for accessing eduGAIN services. Once widely supported, especially international research projects and international content publishers will benefit from eduGAIN. National identity federations like the Fédération Education-Recherche can become eduGAIN member federations by agreeing on common policies and technical standards. This also requires some deployment efforts, not only by the federation operators but also by the individual institutions (e.g. universities), which have to adapt their configurations. The Fédération Education-Recherche is ready for eduGAIN. The next task includes supporting the French institutions to take the necessary steps too. The uptake of organisations joining eduGAIN has yet to accelerate and some mostly non-technical problems have yet to be solved. But in the end it can be said: The more organisations participate in eduGAIN, the greater are the potential benefits for the research and education community.