MyAcademicID
Technical Workshop Report

Table of Contents

Overview

The focus of Activity 1 “Blueprinting a European eID for Higher Education” is to gather the requirements and draw the technical blueprint for a European eID for higher education, creating the digital environment for the ‘once only principle’ and taking into account existing deployed services and solutions such eduGAIN, eduroam, InAcademia and the European Student Identifier along with eIDAS.

As part of Task 1 “Requirements and technical outline for a European eID for higher education”, three technical workshops were organized in the first 5 months of the project with the goal to update the requirements for a European eID for higher education. As a starting point, we used the work that has taken place in the past in the GN4 and AARC projects as well as previous CEF projects such as the "GEANT eduGAIN - eIDAS comparison study" and the "Feasibility study on cross-border use of eID and Authentication Services (eIDAS compliant) to support student mobility and access to student services in Europe".

The first workshop was held on February 21st and 22nd. It focused on discussing the current status of the three pillars of MyAcademicID, eduGAIN, eIDAS and the European Student Card and understand the use cases of the e-services that constitute our initial target for enabling e-IDs in the project. In the second workshop, which took place on April 1st and 2nd, we re-centered the discussion around the requirements of the MyAcademicID use cases. During the first two workshops the complementarity between eduGAIN and the European Student Card project was identified and highlighted inthe discussion. The third workshop took place on May 6th and 7th where we discussed the implementation plan.

From Identification to Authentication

During the workshops, the complementarity between eduGAIN and the European Student Card was identified and highlighted as the European Student Card focuses more on the student identification to various services, while eduGAIN and eIDAS place their focus on the authenticated, electronic access to services. Indeed, there is a multitude of services that rely on the user being able to identify herself by presenting by presentation of a recognisable badge for example in the form of a card. Examples of services in this category include access to libraries and campus restaurants, but can also be extended to public transportation, cinema and theatre and many other services provided in the physical world which require the identification of the individual. The mere fact that the person possesses the card is enough for the service provider to give access or a discount to that person. Whether the person is indeed who she or he claims to be is of no importance for the delivery of the service.

On the other side, there are many services that should only be provided to the right individual. For example, access to my student records should be allowed only to myself and other authorised entities. Authenticated access for electronic service entails the verification of the user's identity through an implicit or explicit challenge to verify that the person in front of computer is indeed the person that she or he claims to be. The strength of this verification processes determines the level of assurance one can have on the authentication process.

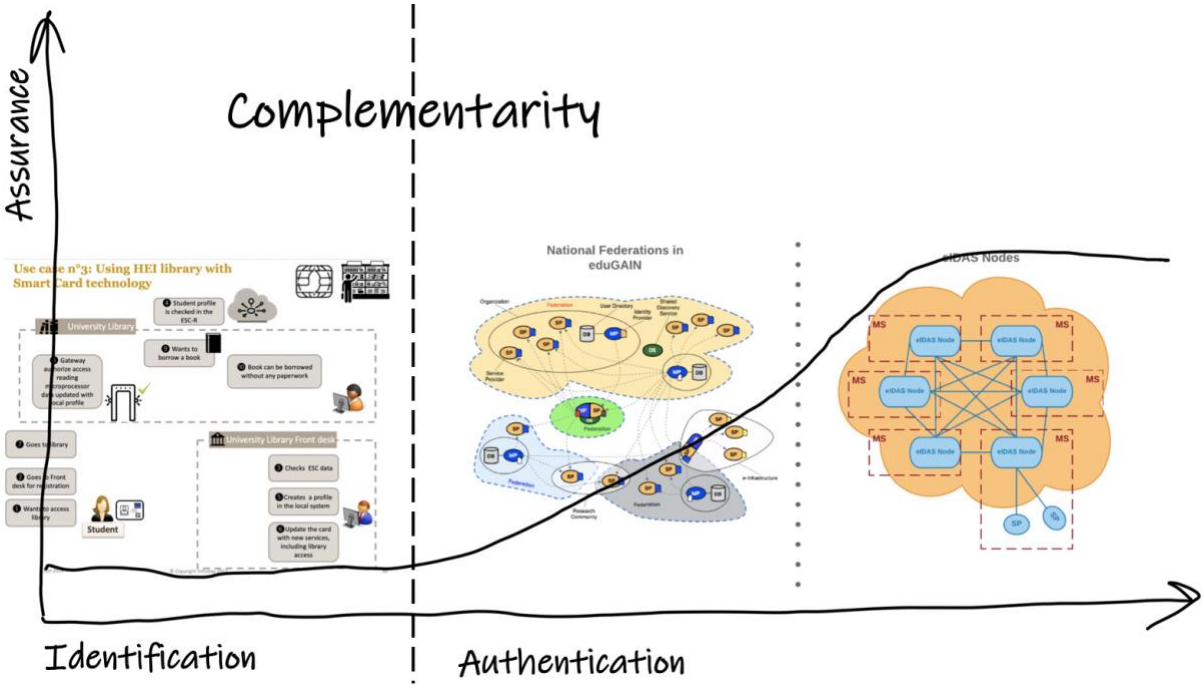


Figure 1. Complementarity among networks

In the figure above, on the left side we have all those use cases that can rely on student identification and which do not require use authentication. These use cases are ideal for the European Student Card, which can provide the means for the student to identify themselves to such services.

Moving rightwards on the diagram, we have the thousands of electronic services that are available through the National Federations in eduGAIN and which do require the authentication of the user at their home institution. For these cases, users/students have to verify themselves by means of authentication, which, today, primarily happens using username and passwords provided by the home institutions of the users. As part of the authentication process, services may request and receive extra information about the enacting users, such their name, e-mail address, institutional affiliation etc.

On the right side of the diagram, we have a number of use cases that require the use of citizen e-IDs that can provide higher level of assurance for the authentication process. As the use of

citizen e-IDs becomes more widespread, we expect to see more and more services relying on the use of e-IDS for the user authentication.

Enabling federated access for the MyAcademicID Services

In the MyAcademicID project we have identified a set of representative services from the European University Foundation that are used for enabling student mobility. These services include:

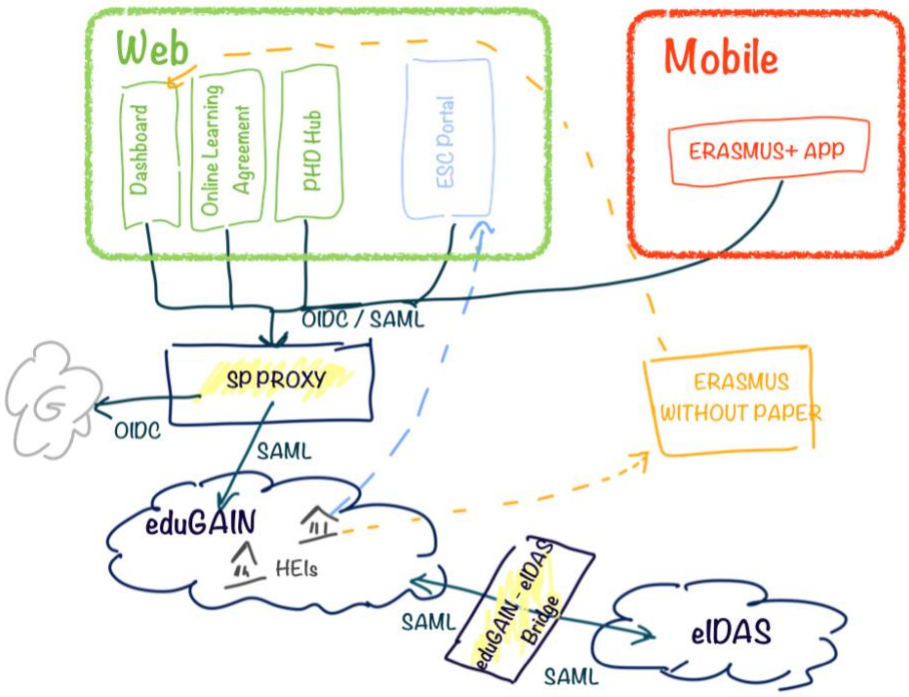
- the Erasmus Dashboard;
- the Erasmus Mobile App;
- Erasmus Without Paper;
- the Online Learning Agreement Tool;
- and the PhD Hub

Although it was not included in the list of services in the MyAcademicID project, during the workshops we recognised the additional value that the integration of the European Student Card would bring. Thus, we have included also the ESC portal in the list of project services.

These services have some common characteristics, but also important differences. The Erasmus Dashboard, the Online Learning Agreement, the PHD Hub and the ESC Portal, are web-based applications, which offer personalized services to users. The users need to authenticate themselves in order to access those services and at the same time the services need to know from which institution the user is coming from. The Erasmus Mobile App is very similar in requirements with the previous set of services, but it is a mobile application.

Erasmus Without Paper (EWP) is another service involved in the enablement of student mobility. The main difference between EWP and the previous services, is that EWP is not a user-facing service. EWP is a service that connects directly to the IT backends of institutions and can be used to transfer student records to other Erasmus services. As this service is not user-facing, it does not require student authentication.

In order to enable access to the mobile app and the web-based services, we are going to make them available to the National Federation through eduGAIN. This will allow (a) the users to authenticate at their home institutions and (b) the services to receive information about the users' affiliation and contact information from the home institutions. The services are going to be connected through multi-protocol SP Proxy (Service Provider Proxy) provided by GÉANT, which will allow the services to use the OpenID Connect protocol in order to authenticate users in eduGAIN, which uses the SAML protocol.



The use of the GEANT SP proxy, provides also a number of additional benefits:

- It allows to use the existing authentication stack of the services, which is based on OpenID Connect.
- Instead of having to connect multiple services in eduGAIN, only one service has to be connected, the GEANT SP Proxy.
- Handling the discovery of thousands Identity Providers and their metadata can be complex. This capability will be provided by the GEANT SP Proxy, so the services do not have to adopt the multi-federation, multi-identity-provider environment of eduGAIN.
- Currently, these service use Google authentication to authenticate users. The GEANT SP Proxy can be configured as an Identity Hub with Account Linking capabilities. Again, these are capabilities that will be provided by the GEANT SP proxy to all the connected Erasmus services, without having to modify anything on the service side.

User Identifiers

The student mobility processes require the use of a number of services, all of which are involved in different stages of the pipeline and which will need to be able to exchange data about the students who are in mobility. In order to enable these processes, we require a student identifier that can be made available by the institution and which can be used by all the services to uniquely identify the user.

Identifiers from the Identity Federation and eduGAIN

In the Identity Federation in eduGAIN there are a number of identifiers that are being used:

- **eduPersonTargetedID** is a persistent, non-reassigned, opaque identifier for a principal. In abstract terms, an eduPersonTargetedID value is a tuple consisting of an opaque identifier for the principal, a name for the source of the identifier, and a name for the intended audience of the identifier. The source of the identifier is termed an identity provider and the name of the source takes the form of a SAML V2.0 entityID, which is an absolute URI. The name of the intended audience also takes the form of an absolute URI, and may refer to a single service provider or a collection of service providers, although the latter is not commonly used.
- **eduPersonPrincipalName** is a scoped identifier for a person. It should be represented in the form "user@scope" where 'user' is a name-based identifier for the person and where the "scope" portion MUST be the administrative domain of the identity system where the identifier was created and assigned. Each value of 'scope' defines a namespace within which the assigned identifiers MUST be unique. Given this rule, if two eduPersonPrincipalName (ePPN) values are the same at a given point in time, they refer to the same person. There must be one and only one "@" sign in valid values of eduPersonPrincipalName. Syntactically, ePPN looks like an email address but is not intended to be a person's published email address or be used as an email address. In general, name-based identifiers tend to be subject to some degree of expected change and/or reassignment. Values of eduPersonPrincipalName are often, but not required to be, human-friendly, and may change as a result of various business processes. They may also be reassigned after a locally defined period of dormancy. Applications that require a guarantee of non-reassignment and more stability, but can tolerate values unfriendly (and unknown) to humans should refer to the eduPersonTargetedID attribute.
- **schacPersonalUniqueCode** specifies a “unique code” for the subject it is associated with. Its value does not necessarily correspond to any identifier outside the scope of the directories using this schema. The <country-code> must be a valid two-letter ISO 3166 country code identifier or the string “int”, and assigned by the SCHAC URN Registry for this attribute; <iNSS> is a Namespace Specific String as defined in RFC 2141 but case insensitive, from a nationally controlled vocabulary, published through the URI identified at the above mentioned SCHAC URN registry.
- **schacPersonUniqueID** specifies a "legal unique identifier" for the subject it is associated with. The <country-code> must be a valid two-letter ISO 3166 country code identifier or the string “int”, and assigned by the SCHAC URN Registry for this attribute; <idType>: Acceptable values must be declared per each country code through the URI identified at the above mentioned SCHAC URN registry.
- **subject-id:** This is a new identifier that is long-lived, non-reassignable, omni-directional identifier suitable for use as a globally-unique external key. Its value for a given subject is independent of the relying party to whom it is given. A value (the unique ID and scope together) MUST be bound to one and only one subject, but the same unique ID given a different scope may refer to the same or (far more likely) a different subject. The relationship between an asserting party and a scope is an arbitrary one and does not reflect any assumed relationship between a scope in the

form of a domain name and a domain found in a given SAML entity identifier. A value MUST NOT be assigned to more than a single subject over its lifetime of use under any circumstances. The unique ID should therefore be constructed in a fashion that reduces the probability of non-technical or political considerations leading to a violation of this requirement, and any such violation should be treated as a potential security risk to the relying parties to which the value may have been given. Relying parties should not treat this identifier as an email address for the subject as it is unlikely (though not precluded) for it to be valid for that purpose. Most organizations will find that existing email address values will not serve well as values for this Attribute.

Out of these identifiers, the first two (ePTID and ePPN) are commonly used in eduGAIN, but they do not meet our requirements. ePTID is targeted and thus will change from one service to the other, while for ePPN there is no guarantee that it will not be recycled, unless the Identity Provider supports the Research & Scholarship Entity Category. Both identifiers, usually, represent the user in the user directory and they do not reflect the student identifier that follows the student record.

Regarding the schacPersonalUniqueCode and schacPersonalUniqueID, these are identifiers that are not commonly found at the inter-federation, eduGAIN level, but might be found more commonly within the boundaries of a Federation or an institution.

The subject-id identifier, as it was mentioned, is a rather new identifier and the adoption is rather low.

Another identifier worth mentioned is the Orcid identifier. ORCID iDs are persistent digital identifiers for individual researchers. Their primary purpose is to unambiguously and definitively link them with their scholarly work products. ORCID iDs are assigned, managed and maintained by the ORCID organization.

The European Student Identifier

The European Student Card has introduced a new identifier: the European Student Identifier.

The European Student Identifier (ESI) is made up of the following data, separated by a hyphen:

1. **Country code** of the HEI that issued the card, on two upper case characters, according to the ISO 3166-1 norm, e.g.:

- DE for Germany;
- FR for France;
- IE for Ireland;
- IT for Italy;
- Etc.

2. **Region code** of the HEI that issued the card. This code is optional. The format is free. It is proposed to use the nomenclature of statistics territorial units (NUTS). NUTS codes already contain the country code, which should be omitted.
3. **HEI Participation Identification Code (PIC)**. This code identifies the HEI as part of the ERASMUS network. Every higher education institution in Europe has such a code, or can obtain one with a simple administrative action. This code includes 9 digits. It can be found by searching on the web site http://ec.europa.eu/education/participants/portal/desktop/en/home.html. In the rare case of a lack of PIC code, a pseudo code will be generated
4. **Student unique code** in the HEI where the student is enrolled. From case to case, this code may be national wide, regional wide or the home institution own identifier.

At the moment, the adoption of the European Student Identifier is rather low, limited to the institutions that have been piloting the European Student Card. One important characteristic of the European Student Identifier as defined the ESC project, is that it links directly to the student identifier that is in the student record. On the other hand, one limitation of the current specification is that it requires the use of a PIC number. The PIC number as a way to identify institutions has been reported to have a number of issues, the most important of which is that it is planned to be replaced by another identifier in the near future.

Requirements for a European Student Identifier

The ESI should be:

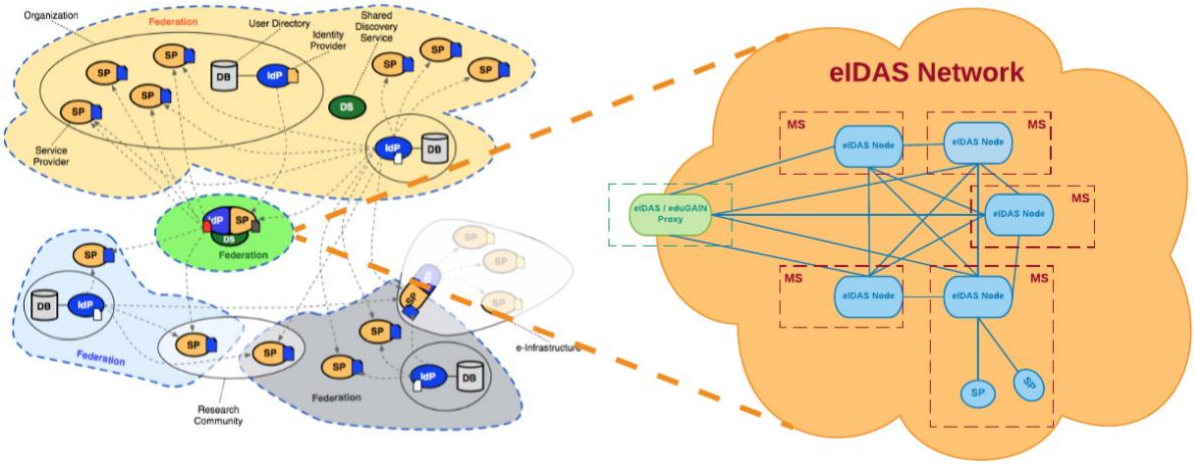
- Globally Unique: Each student should be uniquely identified across organizational and national boundaries
- Persistent: The identifier should follow the student during her/his time of studies
- Non-targeted: The identifier should be the same for all services involved in the student mobility processes
- Protocol neutral: The identifier should not change value depending on the propocal used. For example, it should be the same regardless is SAML or OpenID Connect is used
- Data transport neural: The identifier should not change value depending on how it is transported. For example, the students should be identified by the same identifier regardless of whether it happens through a federated authentication flow or a back-channel transfer of records.

In the next period, the project will engage with key stakeholders from the EC, the Identity Federations and the European Student Card initiative in order to decide on a solution that meets all requirements.

Connection with eIDAS

As the citizen e-IDs become wide-spread, users should be able to authenticate to services through eIDAS. As depicted in Figure 2, users should be able to authenticate using their national e-IDs.

In this case, MyAcademicID can be used by public services as a data provider. Service providers would be able to have information about the student's status or the university. For example:



A SAML-to-SAML protocol proxy acts as a bridge between the eIDAS Network and the Identity Federations in eduGAIN. In the identity federations in eduGAIN the proxy appears as an Identity Provider, while in the eIDAS Network, the service participates as an eIDAS Service node.

In the first phase, we are going to implement such a proxy and connect it to the test environment Swedish eIDAS. In parallel, we will start the discussion within the Cooperation Network regarding integration of eduGAIN as a virtual country in eIDAS.