12-03-2018

Deliverable D9.3
Best Practice for User-Centric Federated Identity

Deliverable D9.3

Contractual Date:	31-10-2017
Actual Date:	12-03-2018
Grant Agreement No.:	731122
Work Package/Activity:	9/JRA3
Task Item:	Task 3
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	SURFnet
Document ID:	GN4-2-17-10451F
Authors:	D. Vaghetti (GARR) ed.;H. Bourgault (RENATER);R. Brugger (SWITCH);C. van Genuchten (SURFnet);M. Héder (MTA-SZTAKI);M. Kremers (SURFnet);K. Meyer (GÉANT);C. Sclifos (RENAM)

© GEANT Limited on behalf of the GN4-2 project.
The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 731122 (GN4-2).

Abstract

This deliverable outlines the basic principles of user-centric federated identity, focusing on those use cases where this approach provides an improvement over current practices. It provides assessments of the demand and readiness level of federations to adopt user-centric federated identity and outlines example architectures and best practices for its delivery as well as the related policy and governance.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

The current practice for identity federation in research and education is to manage identity with an organisation-centric approach, where users are assigned an identity as part of their enrolment process in organisations. This identity attached to the organisation can then be used to access services in a federation or inter-federation. This deliverable looks at identity in a wider context and outlines how to combine user-managed or user-centric identities with organisational identity. In this scenario organisations are no longer the sole providers of identity in a research and education context, but the advantages of classic identity federation are preserved, especially in terms of vetting, trusted sources of attributes, and privacy.

This change in architectural approach aims to address challenges for lifelong R&E identity and stemming from an increased mobility of users between institutions that have been observed over more than 10 years of R&E identity federation operations in production. Focusing on those use cases where a user-managed rather than organisation-managed R&E identity is an improvement over current practices, assessments are provided of the demand for this sort of approach and the readiness level of federations to adopt it, and example architectures and best practices outlined for delivering user-managed, lifelong R&E federated identity and the related policy and governance.

A user-managed R&E identity model can especially improve the current user experience and user management practices over the organisation-centric approach for use cases related to mobility between institutions, multiple affiliation and lifelong learning. Account linking enables the user to have access to all relevant affiliations at a given time, and the existence of an educational identity independent of individual organisations enables lifelong learning, while providing a closer trust relationship with the sector than fully commercial or social identity providers.

Currently two NRENs, SWITCH (Switzerland) and SUNET (Sweden), are taking the lead and already implementing this new approach, although based on different architectures. GARR (Italy) is planning to roll out its own version of this model, incorporating national eGOV-ID identities. These three architectures are compared and analysed in the document.

Finally, a set of practices are identified to guide potential adopters who are considering moving from an organisation-centric approach to user-centric R&E identity management. Policy, governance, data protection, technology, security and interoperability principles are covered.

The work presented in this deliverable, including practical examples of NRENs that are adopting this approach and the additional knowledge gathered, aims to encourage more federations to consider the shift from purely organisational identities to lifelong individual R&E identities which enable users to manage their relevant affiliations more flexibly.

1 Introduction

Current Federated Identity Management (FIM) is largely based on an organisation-centric approach, where users are only enabled to use federated authentication once they have been enrolled in an R&E institution. This kind of approach has a major benefit in that it guarantees that the user belongs to a recognised institution which is a member of the identity federation. However, its downside is that while this system works very well for stable and unique relationships, it is not so efficient in managing looser associations, e.g. when a user has a relationship with several institutions, either simultaneously or in sequence.

In the current R&E landscape, as students and/or researchers progress in their academic careers, intermittent and concurrent relationships between users and institutions are now common. The organisation-centric approach reveals many shortcomings when dealing with identity provisioning and de-provisioning for these dynamic and loose relationships. The same problems arise when managing multiple concurrent affiliations or leveraging (long-lasting) external attribute authorities such as ORCID. These issues have a negative impact on the mobility of both students and researchers, as well as on lifelong learning activities and linking industry and government users to the R&E ecosystem.

For this reason, a new approach should be considered that decouples user identity from user role(s) and affiliation(s). The cornerstones of this type of approach to digital identity are:

- It should be focussed on the user rather than the organisation – the R&E identity is created and maintained by and for its owner.
- It should be persistent – the identity is associated to the owner via a lifelong identifier.
- It should preserve the user's privacy – each user controls the propagation of personal attributes.
- It should be institutionally backed within R&E – Universities and Research Centres provide trusted attributes for affiliated users.

In Section 3 of the document, the architectures of the SWITCH and SWAMID edu-ID projects and a of a third proposed architecture based on an IdP/SP proxy that shares some of the targets of the first two are described and compared. In Section 4, best current practices and recommendations for implementing a user-managed approach to lifelong R&E identity are presented to support those NRENs and Identity Federations that are currently evaluating changing their federation operational models in this direction.

2 Evolution of the Federated Identity Management Approach

2.1 Background

Identity federations have been implemented at a national level for over 10 years, providing a trusted technical and policy framework for users to access a wide variety of services. Many of the initial use cases involved access management to resources owned or contracted by an institution, such as e-learning platforms or journals, but without the university IT department or service having to shoulder the support burden of providing students and staff with multiple credentials. Use cases were then expanded to include a wider variety of services including, with the advent of eduGAIN, those offered across national boundaries. In many federations today, the driver is still institutional needs, with services requiring explicit endorsement by an institution to participate.

While this approach works well for organisations and services, there is still scope to improve the experience for the user, as well as a need to address changes in the external environment. The growth of social and governmental identities over this period, along with changes to academic career paths which make multiple organisational relationships more likely, mean that users end up with a significantly more complex set of identities and relationships with the R&E environment than previously envisioned. To determine the best response to these challenges, federations introduced a concept by the working title of 'eduKEEP' to investigate a possible transition from an organisation-centric identity management architecture to one where the user is positioned as the owner of their R&E identity.

2.1.1 Prior Work in GN4-1

The eduKEEP work item was started during GN4-1 as part of the Joint Research Activity 3 Trust and Identity Research, Task 1 Attributes and Authorisations (JRA3 T1). The eduKEEP team looked at a user-managed identity model for identity federations and created a solution concept for user-centric identity management based on lifelong or long-term identity, a central IdP, and the splitting of the login process in three distinct phases: authentication, identity enrichment, and service access. A summary of this work was published as part of the GN4-1 JRA3 T1 Deliverable [GN4-1_D15.1].

In GN4-2, the eduKEEP team (as part the of Trust & Identity Technology Development task - GN4-2 JRA3 T3) used this solution as the initial reference model. While this solution was eventually refined and further improved according to specific use cases, it retains a set of core features that can still be found, at least partially, in the currently implemented edu-ID architectures exposed in Section 3.

2.1.1.1 *Solution Concept*

To achieve its goals, the eduKEEP team defined a solution concept that leverages existing identity federations, and at the same time proposes a significant change in approach. The main reasons for making changes to the current architecture can be summarised as follows:

- Better division of responsibilities for authentication and attribute release:
  - eduKEEP makes a clear distinction between the two and distributes responsibilities within these two processes to the entities and organisations that can better commit to them.
- Trusted interaction of multiple parties for authentication and attribute release:
  - eduKEEP assumes information will be coming from different authoritative organisations or entities, which will be queried as Attribute Authorities.

2.1.1.2 *Architecture*

The GN4-1 reference model implementation is shown in Figure 2.1 below and includes a centrally managed IdP in which the user can manage his/her own data, which may also be enriched with data from other sources, such as entitlements and affiliations provided by various institutions.

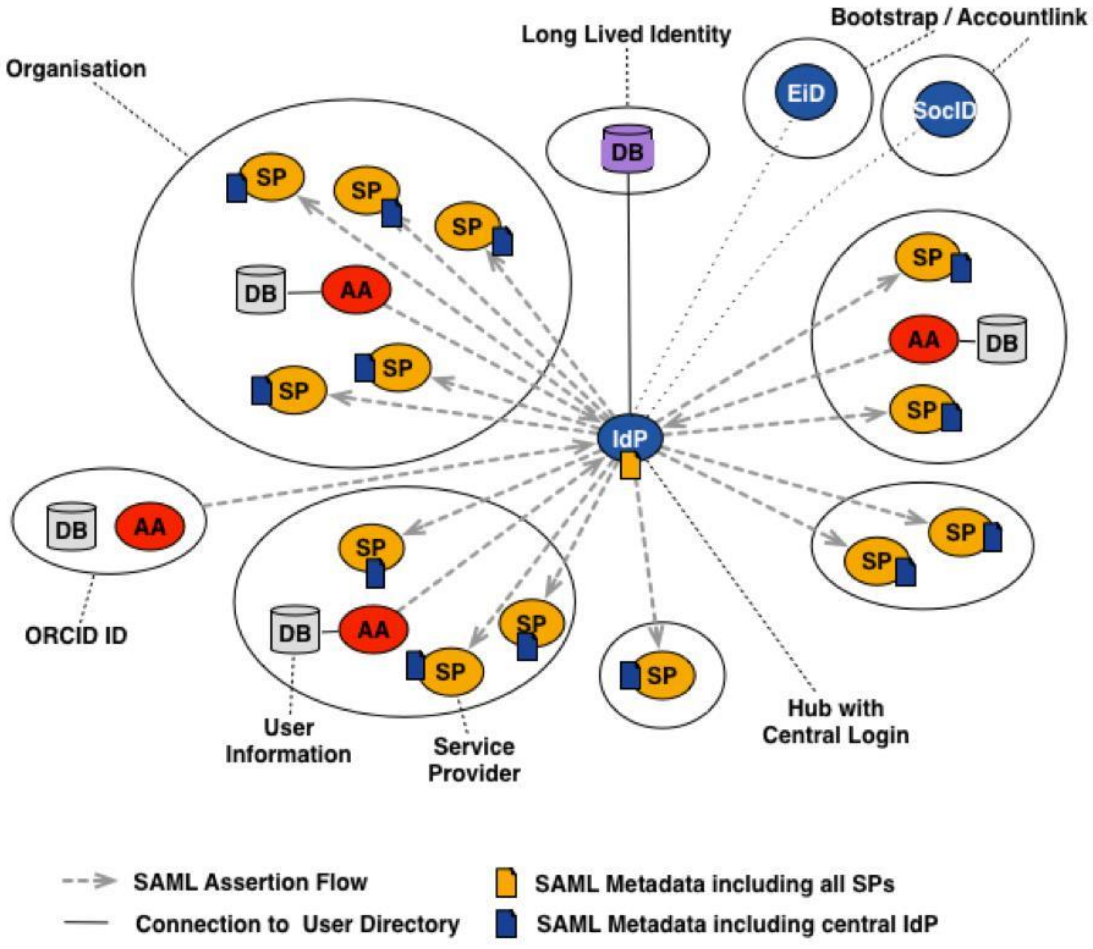

Figure 2.1: An edu-ID architecture with a central IdP design in GN4-1 D15.1

The enrichment phase can be managed by the central IdP itself, which will be thus responsible for querying the Attribute Authorities containing additional user information. Otherwise, the enrichment phase can be managed by the Service Provider, which in this case will act as the querying entity.

Examples of attributes that could be part of the enrichment process are shown in Figure 2.2 below and include, but are not limited to, user affiliation(s), external identifiers (such as ORCID), groups, and role information.

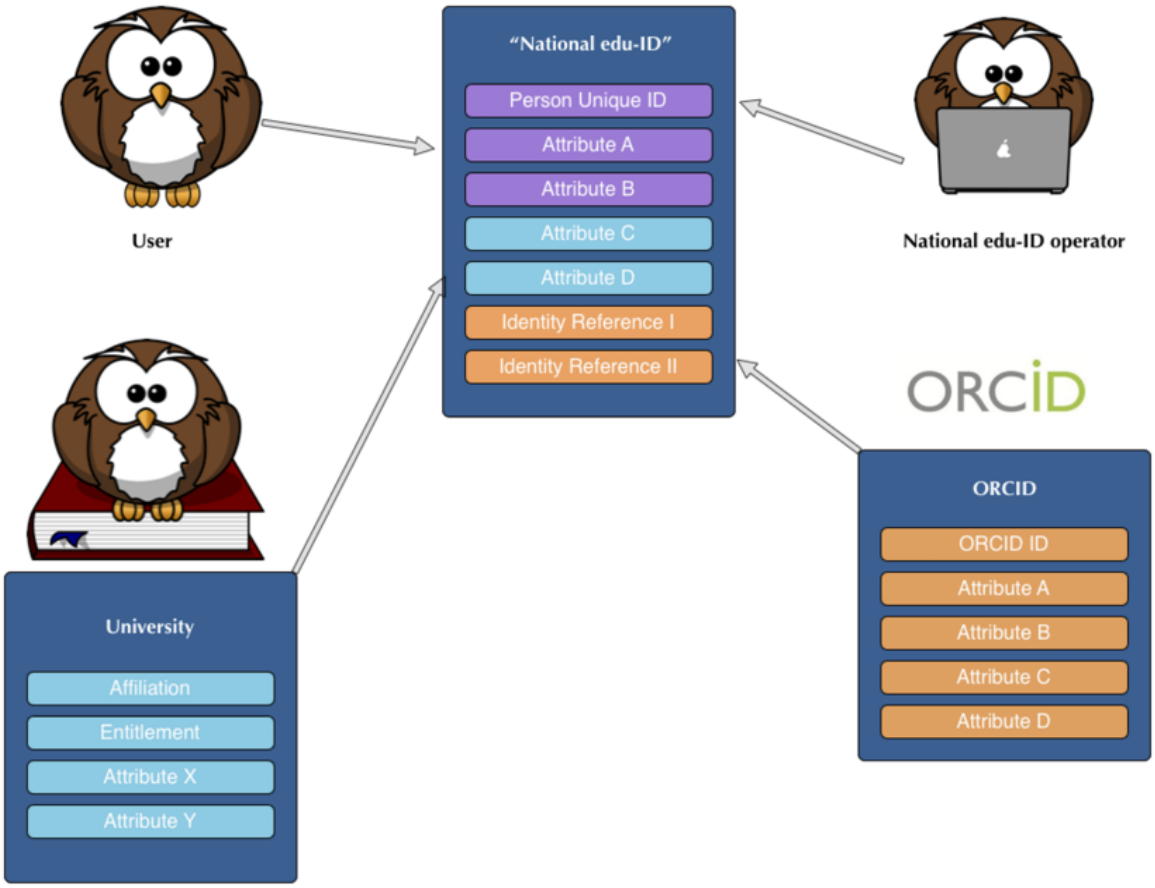

Figure 2.2: After authentication the user identity is enriched with attributes from multiple sources (GN4-1 design in D15.1)

While this model forms the basis for the best practice recommendations, the examination of the case studies carried out in GN4-2 and detailed in Section 3 showed that there were some differences in actual deployments compared with the reference model. This shows that the user-centric identity model cannot be reduced to a single architecture, let alone implementation. Nonetheless, a set of shared key features are:

- Long-lived identities – or at least long-lived identifiers.
- User-managed basic attributes.
- Centralised Identity Provider at the NREN level.
- Attribute enrichment processes based on external Attribute Authorities.

Some practices, as described in Section 4, were therefore identified, which may apply more broadly to these aspects, rather than to the strict implementation of the GN4-1 reference model.

2.2 Common Issues of the Organisation-Centric Identity Model

The user-managed R&E identity approach aims to provide solutions to common issues that arise with the organisation-centric identity model employed by the great majority of identity federations participating in eduGAIN. In the organisation-centric model, the digital identities of individuals are issued by an organisation within a federation which the individual is affiliated to or a member of. Digital identities are usually created when such an affiliation link is established and terminated when the link is dissolved. The following issues may be encountered with this type of identity:

- In the case of an individual changing university or employer, the identity is lost and a new one is created independently.
- Digital identities are mostly restricted to individuals who are members of an organisation of the federation. Therefore, these are not ideal to support trusted interactions with external parties, e.g. in the context of project collaborations, where some users may not have formal membership in an organisation of the federation.
- In the context of lifelong learning associated with concurrent, overlapping, intermittent relationships with educational organisations, digital identities are created and terminated many times, which is inefficient and a burden on the user who may lose access to resources and to a coherent e-portfolio associated with an identity.
- Creating digital identities from scratch has the additional disadvantage that multiple identities that do not necessarily relate to each other are created for a same individual.
- Multiple concurrent affiliations (quite common for researchers and lecturers) produce multiple, concurrent, unlinked and often disliked or unwanted identities.
- The organisation is responsible for all available attributes, including about the individual. This imposes technical, legal and administrative limitations and burdens on what attribute information can be managed consistently.
- The individual is dependent on the baseline norms and processes of the organisation for identity vetting which may not meet the needs of services that have strong requirements in this space. The organisation may not find it cost-effective to improve the baseline for the sake of a small number of cases.

2.3 User-Managed, Persistent and Privacy-Preserving Institutional Identity for R&E

To address the shortcomings and issues of the organisation-centric identity model a new, user-managed type of identity should be considered. In the context of R&E, such a model should preserve the advantages of the current approach, specifically in terms of privacy preserving and ensuring that institutional attributes are provided.

2.3.1 User-Managed

The identity is created and maintained by its owner, usually leveraging a self-sign-up service and providing only basic information such as name and email. The user can then use that identity to bootstrap the accounts for all the Institutions to which they belong, provided that they meet their vetting criteria, or after passing assurance elevation as established by the edu-ID system itself (for an example, see the

SUNET eduID.se architecture). This enables multiple organisational relationships to be handled.

2.3.2 Persistent

To support lifelong learning, the identity is associated to the owner via a lifelong identifier. The system should not require that the user provision a new identity to change university, or to start a parallel collaboration in a research group. The identity should also be persistent during life changes, e.g. marriage, divorce, moving house, gender transitioning and other such events.

2.3.3 Privacy Preserving

Each user controls the propagation of personal attributes. These attributes are transmitted securely (encrypted) and shared based on the principle of data minimisation. Currently, in the organisation-centric approach the institution is responsible for all attribute release, and for seeking consent from the user or using other mechanisms such as legitimate interest to ensure this minimisation (see Section 2.1 of the milestone document *Assessment of DP Legislation Implications* [M9.2 DP Assessment]).

2.3.4 Institutionally Backed

In order to provide a clear advantage for use of this model in R&E, although the identity is managed by the user, institutions should still enhance it by providing trusted attributes for their affiliated users. Currently, institutions provide both the identity and the affiliation information. In the user-managed identity model, identity is decoupled from affiliation, and the Institutions to which the user belongs do not act as Identity Providers, but rather as Attribute Providers.

2.4 Stakeholders

2.4.1 Users

In this context, a User can be anybody who needs some form of identity within research and education e.g. to access a service, to participate in a public-private research collaboration, to use an academic library etc. Typically, the trust framework to enable is provided through R&E identity federation and eduGAIN. Students and affiliates of research institutes are the most common types of users considered. However, services could theoretically be offered from within R&E to the public at large.

2.4.2 Service Providers

Service providers provide services (e.g. a digital library) to end users. They rely on the identity federation to provide the trust framework which is used to authenticate their users, and to make authorisation decisions.

2.4.3 Institutions

Users are affiliated with institutions or, in other words, institutions, e.g. Universities, Research Institutions, *have* users. In the user-managed federated identity model institutions do not provide authentication themselves but may provide user attributes via an attribute authority. Academic institutions usually run a couple of service providers and, in most cases, operate an attribute authority.

2.4.4 Federation Operator

The federation operator is responsible for maintaining the core services of a federation, e.g. a metadata service, as well as for maintaining contracts with federation members.

2.5 Use Cases

Some examples of the most important types of use cases for the described approach are given below.

User-focussed use cases:

- Alice has finished school and wants to start studies in biology. In order to sign up to university she first creates her own digital, academic identity. If she is admitted at the university, Alice then uses all relevant services using the same identity that she created to sign up.
- Adam is undecided as to whether he wants to study history at Trinity College or geology at King's College. He signs up for admission exams at both universities using the same digital identity.
- A year ago, Bob received his master's degree in psychology. Using his digital identity, he continues to access eligible learning material on the file sharing service provided by the university.
- Clare is an experienced and renowned manager in a pharma cooperation. She teaches business administration seminars at various universities. For the duration of a seminar, she is granted access to the learning management system to upload training material and communicate with the seminar participants.
- Gerard is a researcher based in Grenoble. Next week he will move to Manchester for a 3-month research activity. Access to the experiment setup in Manchester is granted through his academic identity.
- Sophia is a part-time worker. Next month she will apply for a 6-month university course in physiology. She logs in to the online application using the digital identity she had during her former studies.

- A professor holds one or several courses at a university. As part of a collaborative project (or academic mobility under Erasmus), he will also hold a course (or several) at another university (or universities) in another country. In this case he will require access to basic university resources (library, post), as well as to resources at the host university to check students' results. The professor is assigned an affiliation to the foreign university that is eventually linked to his identity.

Federation Operator use cases:

While resolving limitations for the issues already identified, by combining different sources of attributes and enriching them, the approach also allows a federation to more scalably improve the features available within the identity federation that may be currently challenging to implement. Some examples are:

- A Federation Operator wishes to provide a multi-factor authentication solution for his community, in order to increase login security for sensitive services. The edu-ID central IdP can be used as single point to offer a second-factor authentication service.
- A Federation Operator wants to deploy a new standard/protocol for user authentication (such as OpenID Connect) for its entire federation. It can easily achieve this by adding support for this new standard/protocol directly to a central edu-ID IdP.

3 Federation Perspectives

3.1 Federation Survey

In order to identify and quantify the actual level of interest in the user-centric identity federation concept, an online survey was conducted at the beginning of 2017.

The target of the survey were the operators of identity federations, primarily those that are members of eduGAIN. The survey was circulated to members of the REFEDS Federation Operator Group (FOG) and the wider REFEDS community. 16 of the 40 eduGAIN member federation operators completed the survey (40% response rate). A single consolidated response per participating federation was received.

The survey questions were grouped under four categories:

- Interest in and readiness for the user-centric identity concept.
- Specific issues: how to deal with mobility, identity bootstrap, and multiple affiliation.
- Available identity infrastructure facilities and features: account linking, attribute authorities, assurance profiles, and attribute quality.
- Targeted questions: different sets of questions for the 'not interested' and 'interested' groups respectively.

The questions were aimed at assessing the potential for adoption of this approach within R&E, gathering information about specific aspects of implementation to support with Best Practice Recommendations and determining the overall readiness level for its implementation within the community.

3.1.1 Results

The term *user centric* is used in the survey results shown below to highlight the disruptive difference between the current organisation-based model and the proposed approach whereby it is the user who manages his/her, identity which is then enriched by a range of sources, including affiliation information from organisations.

3.1.1.1 *Interest in Concept*

Responses in this category indicate that the degree of interest in the subject among those who answered is high, with two federations already deploying this type of solution, and 10 expressing an interest in the concept. Adjusting the data to take into account the federations that gave no answer

(N/A), the picture changes considerably, but the level of interest is still far from being negligible at 30 percent. When also taking into consideration other data sources such as the REFEDS survey on federation funding and staffing, it is reasonable to conclude that not every federation would have the resources to adopt this approach. 30% is therefore a significant level of interest for what is a highly disruptive development.

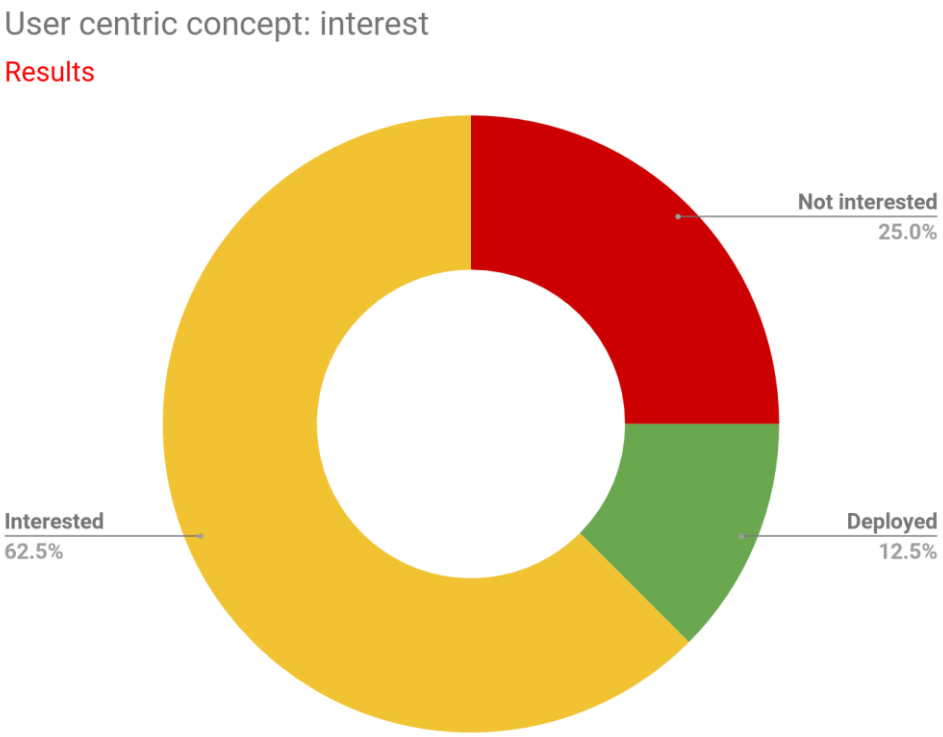

Figure 3.1: User-centric survey results. User-centric concept: interest - Results.

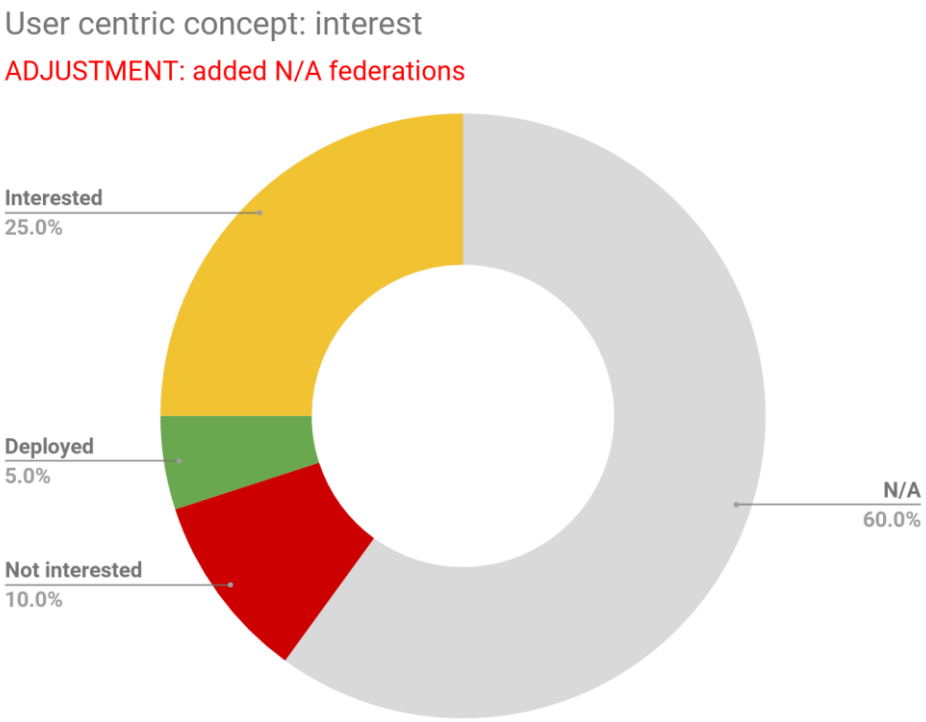

Figure 3.2: User-centric survey results. User-centric concept: interest - including N/A federations.

3.1.1.2 *Specific Issues*

Survey respondents were asked to give their views on specific issues relating to the implementation of user-centric identity. Other questions examined the shortcomings of the organisation-centric identity approach and possible solutions to these which could be provided by a user-centric approach. Notably, the proposed user-centric solutions were chosen only by the 'interested' (and 'deployed') group, while the 'not interested' group expressed preference for the other solutions.

Questions	Answers	User-centric 'interested' + 'deployed' groups		User-centric 'not interested' group	
		Currently	Desired	Currently	Desired
Lifelong learning How do you manage lifelong learning?	Creating new accounts for new and returning users	91%	0%	50%	0%
	(user-centric solution) Linking local attributes to external identities	0%	50%	0%	0%
Mobility How do you manage faculty, staff and students from exchange programs?	Creating temporary accounts	75%	0%	75%	0%
	Most services are federated, no need to issue new accounts	0%	66.6%	0%	50%

Questions	Answers	User-centric 'interested' + 'deployed' groups		User-centric 'not interested' group	
		Currently	Desired	Currently	Desired
	(user-centric solution) Linking local attributes to external identities	0%	50%	0%	0%
Multiple affiliation How do you manage multiple concurrent affiliations?	No need to, the user have to select the proper IdP each time	93%	0%	75%	0%
	(user-centric solution) For selected services account linking and attributes aggregation is performed through an IdP/SP Proxy	9%	50%	0%	0%
	(user-centric solution) We avoid issuing new accounts leveraging lifelong educational identities that are linked to the affiliated organisations.	0%	41%	0%	0%

Table 3.1: User-centric survey results: questions on specific issues

Based on the answers to the *specific issues* section, some general considerations can be drawn. First of all, as mentioned, the 'not interested' group never chose the proposed user-centric solutions, confirming their initial response.

On the other hand, in most cases only half of the 'interested' group chose the selected user-centric solutions to particular challenges posed by organisation-centric approaches, showing that even among the federation operators interested in the concept, the level of confidence in the user-centric approach is not yet solid enough to gain the support of the absolute majority [1] of interested federations.

3.1.1.3 *Available Identity Infrastructure Facilities and Features*

Answers in this category asses the availability of Attribute Authorities and Account Linking facilities in the respondent Identity Federations, as well as assurance profiles and attribute quality metadata as features of the federated identities.

[1] Given that the audience of the survey is a very highly technical one, their partial lack of confidence can also be attributed to the limited amount of technical details in the proposed answers.

Questions	Answers	User-centric 'interested' + 'deployed' groups	User-centric 'not interested' group	Total %
Account linking Which of the following account linking mechanisms is predominant in your identity federation?	(negative answer) No Account Linking	25%	50%	31.25%
	(negative answer) Service Providers responsible for Account Linking	41.6%	25%	37.5%
	Account Linking accomplished through an IdP/SP Proxy service	16.6%	0%	12.5%
	Account Linking provided by the "home" IdP	8.3%	0%	6.25%
	A Virtual Organisation platform is used to link the user accounts	8.3%	0%	6.25%
	Account Linking undertaken by a trusted external IdP	0%	25%	6.25%
Attribute Authorities Which of the following best describe the current use of Attribute Authorities in you Federation?	(negative answer) The Federation does not manage any Attribute Authorities or VO platforms services	50%	75%	56.25%
	The Federation does manage Attribute Authorities and/or VO platforms services	33.3%	25%	31.25%
	The Federation IdPs, at large, can also act as AA when needed	16.6%	0%	12.5%
Level of Assurance What LoAs do you offer?	(negative answer) LoAs not provided but minimal identity vetting is undertaken	33.3%	100%	50%
	Users starts with a basic Level of Assurance that is upgraded to a higher one on a case by case basis.	16.6%	0%	12.5%
	LoAs are used based on other means (e.g. eIDAS)	8.33%	0%	6.25%
	A LoA is assigned by the originating IdP	33.3%	0%	25%
Attribute Quality How do you manage attributes quality in your Identity Federation?	(negative answer) We do not assess or manage the quality of attributes	50%	75%	56.25%
	(negative answer) We use the LoA for the whole identity rather than for individual attributes	41.6%	25%	37.5%
	Each attribute has a separate quality statement	8.33%	0%	12.5%

Table 3.2: User-centric survey results: availability of Attribute Authorities and Account Linking facilities in respondent Identity Federations

The overall majority of the Identity Federations answered negatively to this set of questions, but with different rates depending on the group. The 'not interested' group had the highest rate of negative answers, while the 'interested' plus 'deployed' group comparatively had the lowest.

Availability of the above facilities and features can both directly and indirectly provide a solid foundation for the development of a user-centric system. Thus, with reference to infrastructure-readiness the 'interested' and, unsurprisingly, the 'deployed' groups are the best positioned.

3.1.1.4 *Targeted Questions*

The final part of the survey included two different sets of questions for the 'not interested' and 'interested' groups respectively.

The set of questions for the 'not interested' group was based on the use cases where the user-centric approach is generally considered an improvement, asking the federation operators to express their judgement on the relevance of the issues.

Questions	Answers		
	This topic is relevant but a solution is not known or implemented	This topic is relevant and it is already solved	This topic is not relevant to our identity federation
Individuals have one academic identity for life	50%	0%	50%
Individuals who are affiliated with more than one university at the same time have a single identity	50%	0%	50%
Alumni and other former university members have a digital identity	0%	50%	50%
Researchers can participate in international projects with a single identity	50%	50%	0%
Continuing education students have a digital identity	50%	0%	50%

Table 3.3: User-centric survey results: questions on use cases for the 'not interested' group

As the 'not interested' group only comprises four respondents, it is rather difficult to draw general conclusions from its responses. Nonetheless, the majority of chosen answers polarise between 'not relevant' and 'relevant but without solutions'. It would therefore seem that among the 'not interested' group, the main reasons for disregarding the user-centric approach are rather varied, at least as regards the selected use cases.

The set of questions for the 'interested' and 'deployed' groups aimed to assess their readiness level for adopting a user-centric architecture for their federation.

Questions	Answers <TOTAL %> (<INTERESTED NUMBER>, <DEPLOYED NUMBER>)		
	Yes	Uncertain	No
Legal Questions are Clarified	0% (0, 0)	42% (3, 2)	58% (7, 0)
The Technical Architecture is Known	25% (1, 2)	50% (6, 0)	25% (3, 0)
There are a Critical Mass of Users	17% (1, 1)	50% (6, 0)	33% (3, 1)
There is a Governance Model	17% (0, 2)	17% (2, 0)	66% (8, 0)
There is a Business Plan	17% (0, 2)	0% (0, 0)	83% (10, 0)
All stakeholders agree to (partially) centralise academic identity management	0% (0, 0)	58% (6, 1)	42% (4, 1)

Table 3.4: User-centric survey results: questions on planning and deployment for the 'interested' group and 'deployed' groups.

The answers to this final set of questions clearly show that the two federations that are deploying a user-centric solution are confident in terms of technical architecture, governance model and business plan, while on legal questions, critical mass of users, and agreement from stakeholders they have a similar level of confidence as that expressed by the 'interested' group.

The overall picture is that a clear majority are uncertain, or simply do not have an answer, so their readiness level is correspondingly low. However, considering that at the time of the survey two identity federations already had a user-centric system in production, solutions beyond proof of concept are available. This leads to conclude that the uncertainty might be the result of a lack of information rather than of solutions or practices. This best practice document is intended to provide meaningful help for those who are uncertain to make a more informed decision on whether and how to evolve their identity federation to support account linking, lifelong identifiers, attribute enrichment and user-managed identities.

3.2 Study of Existing and Planned Implementation

Information about the current and planned new user-managed or lifelong identity-based architectures was collected thanks to the assistance of the developers of these solutions in each NREN including through direct contact and interviews.

3.2.1 SWITCH edu-ID Architecture

The SWITCH edu-ID implements a hub-and-spoke identity federation architecture with a central identity provider and is specifically made for Swiss universities [SWITCH edu-ID]. An important design goal is the provision of identities not just to students, staff and teachers, but also to university guests, further education students, private library users, event participants and other non-typical users.

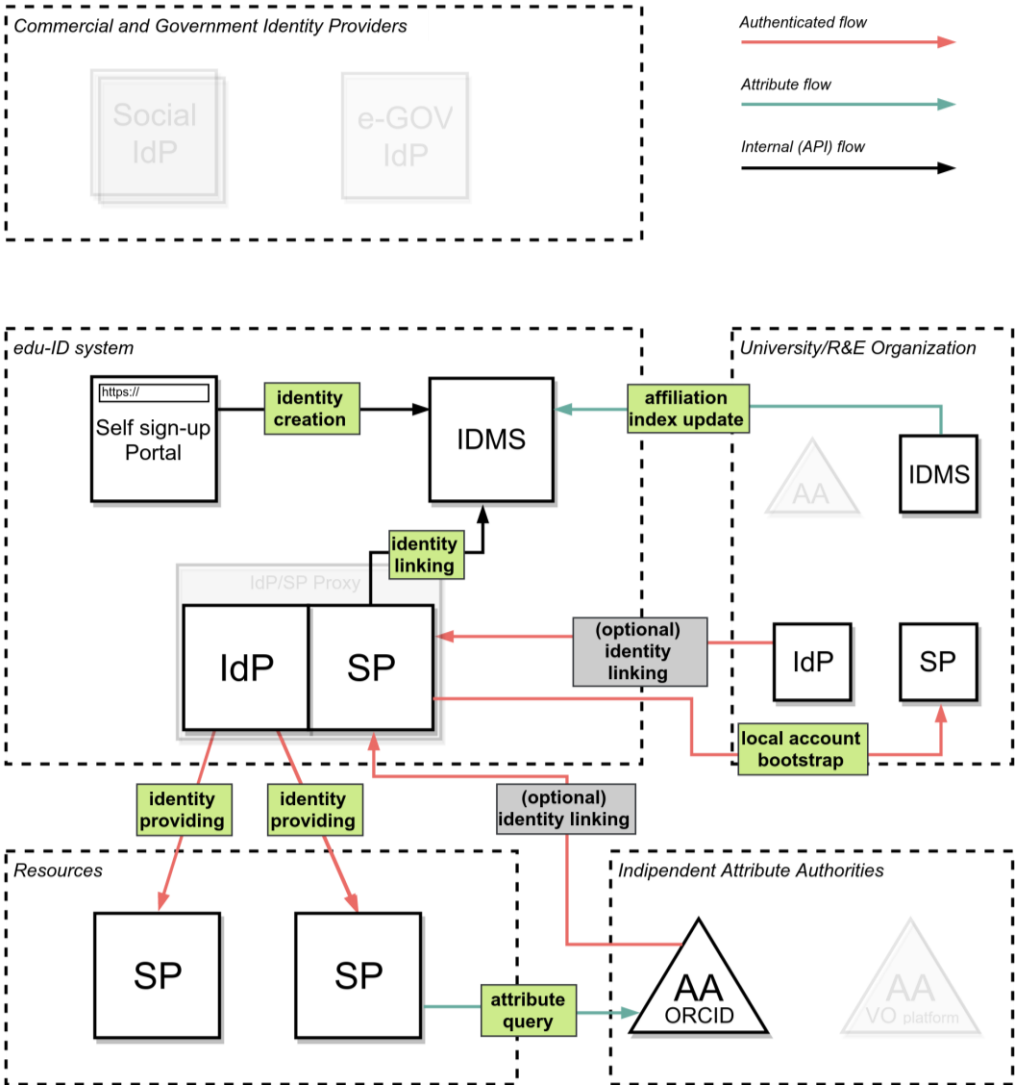

Figure 3.3: SWITCH edu-ID architecture.

The SWITCH edu-ID directory contains user-managed attributes and an affiliation index that indicates which universities the user is currently affiliated with. The edu-ID identity creation is performed through a self-sign-up portal. When a new user is registered at a university (either bootstrapping the local account leveraging the edu-ID user-managed account or through other processes) an organisational account is created and linked to the edu-ID. The organisational identity management system then sends an account linking message to the IdM provisioning service, which updates the affiliation index. Likewise, the affiliation index is updated when a user leaves a university which then sends an account unlinking message.

To access a service, a user authenticates at the central SWITCH edu-ID IdP. The attribute aggregator collects the user-managed attributes and the attributes from all the organisations which the user is currently affiliated with. The attributes are then filtered and reduced to the needs of the service and according to the set of permitted attributes as defined by the university. Finally, with the user's consent, the identity is provided to the service.

A user who has left a university and has no affiliation with any other university keeps the private, user-managed part of a SWITCH edu-ID identity. Although many services will still require users to have a current affiliation with a university, an increasing number of services will be open to users other than students or university staff.

3.2.2 SUNET eduID.se Architecture

The primary goal of SUNET eduID.se is to simplify the Swedish University admission process for future and current students, leveraging a centralised identity provider [SUNET eduID.se].

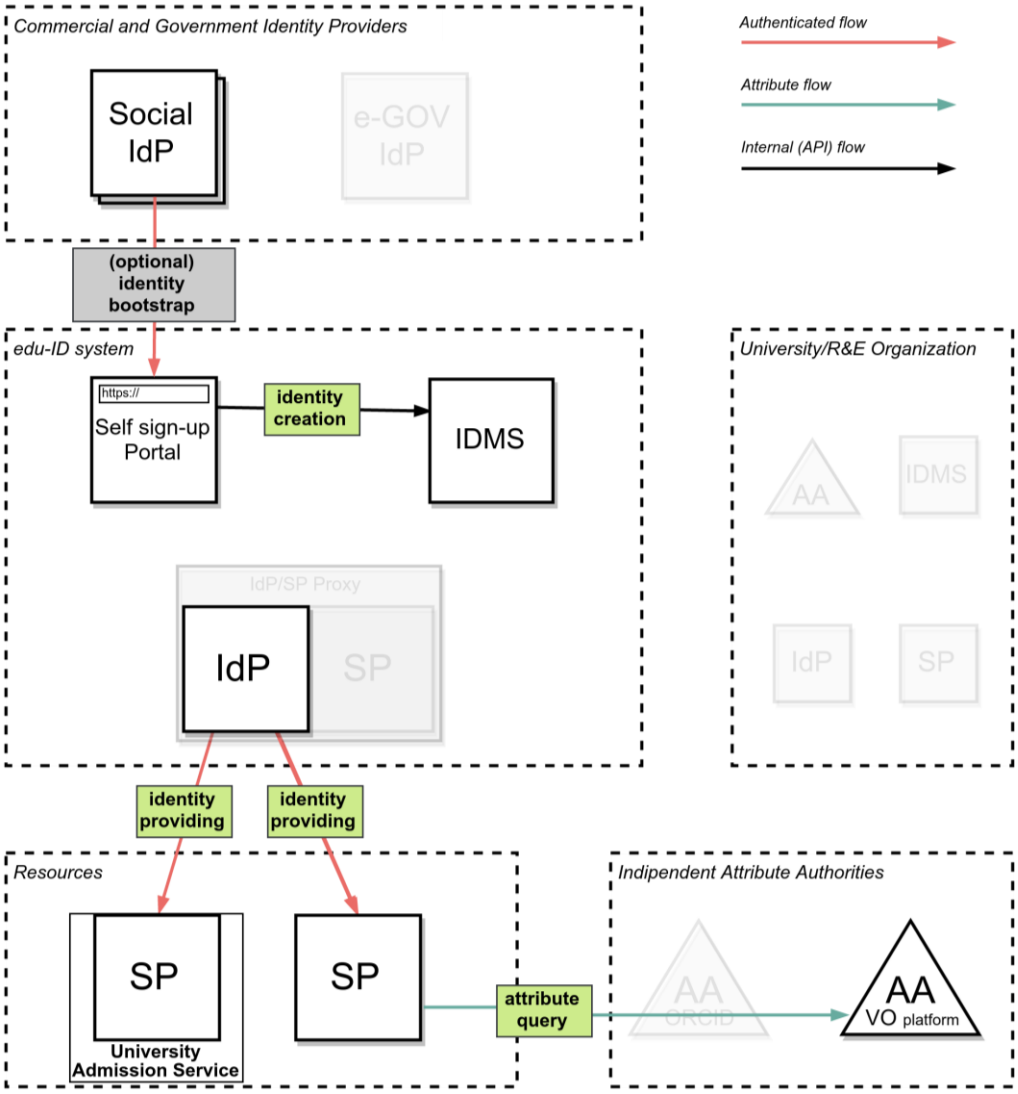

Figure 3.4: SUNET edu-ID architecture.

The self-sign-up portal is the component involved in the self-registration of users. User-managed attributes can be completely self-asserted, or a social IdP can be used to bootstrap the identity. User registration is completed through e-mail confirmation. Self-registered accounts are assigned a basic Assurance Level [SWAMID AL1].

Registered users can access a dashboard to manage their accounts, and confirm their identity providing a National Identity Number (Swedish 'personnummer'). Confirmed identities are assigned a higher Assurance Level [SWAMID AL2].

eduID.se acts primarily as an authentication endpoint for the Swedish online university admission service providing a lifelong identity that is not bound to the admission service itself, nor to other academic institutions.

Other uses of the service are also planned:

- As a centralised solution to access diploma results.
- As the primary IdP for Institutions that ask for the service (in combination with the SWAMID VO platform).
- As an IdP of last resort in combination with the SWAMID VO platform.

3.2.3 GARR Proxy-Based edu-ID Architecture

A proxy-based edu-ID architecture could be set up to leverage trusted external identity services that provide lifelong identities. GARR is currently planning to adopt this solution and to leverage the Italian national eGOV-ID system as a source of lifelong identities.

In a proxy-based edu-ID architecture, the existing Institutional Identity Providers will also act as Attribute Providers. Specifically, they must support the SAML Attribute Query outside an authentication session.

The proxy implements an account linking service to bind a user account originating from the national eGOV-ID systems with an institutional federated one. In order to establish this link, different strategies can be envisioned, the simplest and most effective of these being to ask the user to log in at both ends thus collecting the two identifiers. Only these identifiers are then stored persistently, while all other attributes are collected in real time, leveraging the authentication session and a SAML Attribute Query (see below). Users must be also given the chance to modify the link between their two identities through a proper user interface. Moreover, to prevent stale linked identities, the system should implement an account linking update procedure to renew the Institutional federated account identifier.

The proxy will also assign a new identifier to the user, which will be provided to the services down the line.

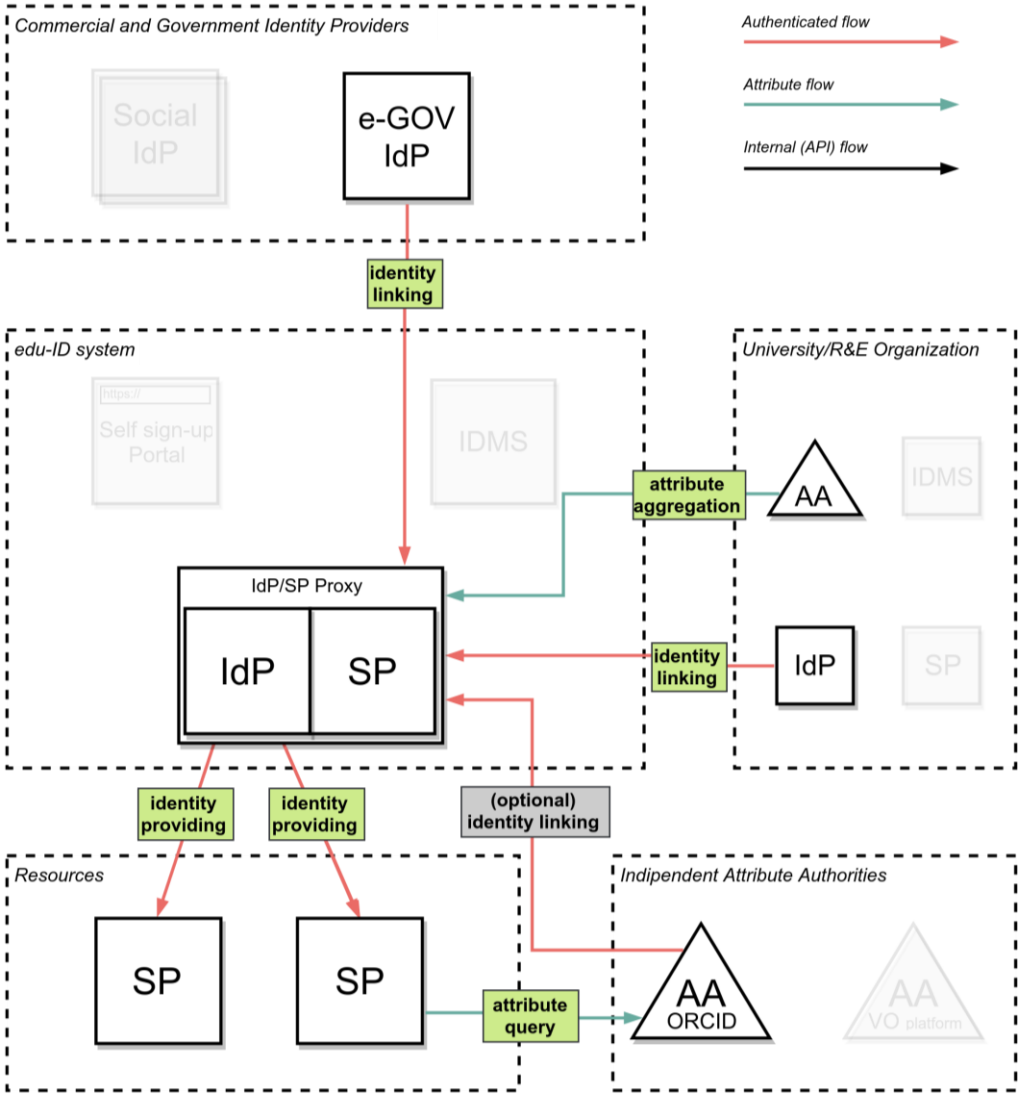

Figure 3.5: GARR edu-ID proxy proposed architecture.

Once the accounts are linked, to access a service:

1. The user selects the edu-ID proxy as a source of authentication in the service provider discovery interface.
2. On the edu-ID proxy, the user then chooses the eGOV-ID system as the IdP against which to authenticate.
3. Upon successful authentication, the proxy issues a SAML Attribute Query directed to the user's Institutional Identity Provider to collect additional attributes, and more specifically the 'affiliation' attribute(s).
4. The proxy then aggregates the attributes collected from the eGOV-ID IdP with the user's HO attributes.
5. Finally, a single coherent SAML Assertion paired with the user's proxy identifier is sent to the Service Provider.

3.3 Case Study Comparison and Analysis

3.3.1 Grounds for Comparison

Two of the three edu-ID solutions included in the comparison are real case scenarios where the concept of user-managed architecture is already being applied by NRENs and identity federations.

3.3.2 Comparison Criteria

The criteria used to compare these solutions are fully detailed in Appendix B.

3.3.3 Comparison Results Overview

Feature	SWITCH edu-ID	SUNET eduID.se	GARR edu-ID proxy
A - Target audience	R&E Community, Persons with loose relationship to universities	R&E Community, Persons with loose relationship to universities	R&E Community
B - Long-term identity	YES	YES	YES *
C- Identity suitable for AuthN	YES	YES	NO
D - A new identifier is provided	YES	YES	YES *
D1- Persistent (stable over time) identifier	YES	YES	YES *
D2 - Globally Unique Identifier	YES **	YES **	YES *
E- Central external IdP acting on behalf of Home Organisation IdPs	YES	NO	YES
F- Account Linking	YES	YES	YES
F1- Linked Account AuthN	NO	NO	YES
G- Self-asserted Identity	YES	YES	PLANNED
G1- Identity Assurance Elevation	YES	YES	YES
G2- VO-based vetting	NO	NO	PLANNED
H- Central Attribute Aggregation	YES	NO	YES
I- Attribute Release Policy - Delegate Management to Home Organisations	YES	NO	NO

* external dependency ** confirmed identities

Table 3.5: edu-ID architectures comparison overview

3.3.4 Discussion of Results

As already summarised in the comparison overview above, the three edu-ID solutions analysed have both shared and specific features and implementation levels (of which a more detailed analysis follows below). Furthermore, compared to the abstract solution described in Section 2 (defined in the GN4-1 JRA3 T1 Deliverable [GN4-1_D15.1]), it is interesting to note that almost all the key features are totally or partially implemented.

GN4-1 D15.1 Feature	SWITCH edu-ID	SUNET eduID.se	GARR edu-ID proxy
Long-lived identities	Implemented	Implemented	Implemented leveraging eGOV long-lived identities
User-managed basic attributes	Implemented	Implemented	Not implemented
Attribute enrichment processes based on external Attribute Authorities.	Implemented	Partially implemented (only through the SWAMID VO platform)	Implemented
Centralised Identity Provider at the NREN level.	Implemented	Implemented	Implemented

Table 3.6: edu-ID architectures and GN4-1 D15 eduKEEP solution comparison overview.

3.3.4.1 *Similarities / Common Concepts*

Target audience: R&E community and beyond

The three edu-ID solutions are mainly designed for students, staff and researchers of universities.

The SWITCH and SUNET solutions are also open to other users who have a loose relationship to the university or home organisation (guests, applicants, alumni, event participants, etc.), which means that in theory anyone can create an edu-ID identity on these two platforms.

User-managed identities

The features and workflows related to the creation and management of true user-managed identities only apply to the SWITCH and SUNET solutions. In these two solutions, users create their identity at a central AAI platform through a lightweight self-registration process. Users are the owners of their identity and can easily manage it (i.e. by setting or updating personal attributes). As identity owners, their consent is systematically required before their underlying attributes are passed to services.

In the case of the edu-ID proxy-based solution, the user-managed principle does not really apply, since the user is not in control of the creation of the identity.

A long-term, unique and stable digital edu-ID identity

The edu-ID identities managed and delivered to the end-users by the SWITCH and SUNET solutions share the same characteristics. They are intended to be long term, which means that users can keep them for life once created.

They also include a unique and stable (over time) identifier, which ensures identification of a user in a unique way across different domains and constitutes the foundation on which the account linking process rests (see below).

Note that, in the case of the edu-ID proxy-based solution, the digital identities delivered by the eGOV-ID platform also respect these requirements.

Authentication to a single central AAI platform

In the case of the SWITCH edu-ID and the edu-ID proxy-based solutions, the users of the federation authenticate to a single central AAI platform operated by the NREN (i.e. respectively a central IdP and a central IdP/SP proxy), regardless of home organisation. In other words, this central IdP (or IdP/SP proxy) is working as a substitute for Home Organisation IdPs that are found in a classic identity federation.

The SUNET edu-ID solution is also running a central IdP but this is currently only used to bootstrap the admission process. However, there are plans to make the solution capable of working as a substitute for Home Organisation IdPs in the future.

Account linking

Each edu-ID solution implements some kind of account linking mechanism to connect the user's edu-ID account (or eGOV-ID account) with one or more additional accounts the user owns (e.g. local account at the University, external social account, etc.).

This type of mechanism is specifically used by the SWITCH edu-ID solution as part of the data provisioning and attribute aggregation processes, which involve both the edu-ID account stored at the central AAI platform level and a local account the user owns, stored at the university level. As a second use case, account linking also enables linking an edu-ID account with one or more external account(s) (social accounts, ORCID account, etc.) belonging to the same user. In this case, linked identities are aimed at importing an external identifier (from an external identity provider) into the Switch edu-ID profile where it can be further processed and, if necessary, made available to services in the form of attributes.

In a similar way, the SUNET edu-ID solution uses account linking to link a SUNET edu-ID account with an external social account (Google, Facebook, etc.), but this time for the purpose of simplifying the edu-ID account creation process for the user.

Finally, the edu-ID proxy-based solution uses account linking as part of the attribute aggregation process performed by the proxy (see below) as well as to link identities for further authentication purposes. The link established here between the eGOV-ID account and the institutional federated account at the proxy level allows the user to perform authentication using any of the linked accounts.

In each of those cases, the account linking mechanisms leverage the unique identifier related to each account the user owns.

Identity enrichment or attribute aggregation

The SWITCH and proxy-based edu-ID solutions both implement an identity enrichment process to complete the basic digital identity of the user obtained from authentication with additional attributes coming from relevant sources (i.e. universities), and then to pass on the resulting combined set to the service providers.

This enrichment process, called attribute aggregation, is in both cases operated centrally (at the central AAI platform level) by a single entity, comprised by the attribute aggregator component (combined with the affiliation index) for the first solution, and the IdP/SP proxy for the second.

Self-asserted identity with identity assurance elevation

In the user-managed model, each edu-ID solution has to deal with self-asserted identities (as a result of self-registration for an identity). All the examined solutions have put in place procedures or strategies to verify the quality of the identity in general and/or the associated attributes.

In case of the SUNET solution, an identity vetting procedure is part of the edu-ID identity creation process. It requires the Swedish “personnummer” (the national tax identification number), combined with a postal address or a mobile phone number, and allows users to elevate their identity from a basic (SWAMID AL1 / self-registered accounts) to a higher assurance profile (SWAMID AL2 / confirmed accounts).

The SWITCH edu-ID solution implements the same kind of verification process but instead of considering the quality of the whole identity, it rather focuses on the quality of each selected individual attribute of the identity (which means that assurance levels are expressed here for each attribute).

In the case of the edu-ID proxy-based solution, the main identity vetting service provided by the proxy is based on account linking with high assurance level eGOV-ID accounts as a way to increase the assurance profile of current institutional federated accounts. Note that, in a longer-term perspective, there are also plans to leverage a VO platform centrally together with the proxy. In this specific case, the identity of the user is then vetted basing upon the validation mechanisms associated to the VO enrolment process (invitation, user pre-registration or admin approval).

Naming

Federations that have adopted this approach tend to prefer the term ‘edu-ID’ to describe it, though it should be noted this term is not exclusive and is in use in other contexts.

3.3.4.2 *Specific Aspects of Individual Solutions*

In contrast to the above concepts, some features, while interesting, remain very specific to each edu-ID solution. These specific features are detailed below.

SWITCH edu-ID: attribute quality validation process

The SWITCH edu-ID solution provides vetting processes to increase the quality level not only of an identity but also of selected individual attributes of an identity, whenever it is required by a service. For this purpose, it implements an attribute model which explicitly supports varying quality by associating quality statements to individual attributes. These quality statements are in particular expressed in forms of attribute assurance levels (from low level for self-asserted attributes with minimal trust to higher levels for confirmed attributes).

Each service defines its own individual quality requirements (as it is up to the service to verify each quality aspect of a user's identity) and has various ways to react to an identity with insufficient quality. The user may be denied access to the service or he may be guided through processes to increase the level of assurance of his identity and/or related attributes to match the requirements.

SWITCH edu-ID: management of attribute release policy delegated to Home Organisations

Although a user-managed approach, the Switch edu-ID solution let the Home Organisations maintain full control of how their organisation specific user attributes are to be released to target services. The management of Attribute Release Policies is in this way performed by the HO AA Administrators at the Central AAI platform level through the access to a dedicated Web UI.

SUNET eduID.se: a central user-managed IdP for identity bootstrap

The SUNET eduID.se platform relies on a centralised user-managed identity provider that aims to allow anyone who is eligible to obtain a lifelong edu-ID identity with a high assurance level (i.e. confirmed account).

Students, for example, can then directly use this confirmed account to log into the Swedish online university admission service (among other institutional services) to confirm their place at the university, as part of the admission process. Moreover, they can also use their edu-ID account to activate local accounts at home organisations.

SUNET eduID.se is currently officially validated against the Kantara Identity Assurance Framework.

GARR edu-ID proxy: leveraging external trusted eGOV identities

The GARR edu-ID proxy-based solution allows a federation's users to authenticate to services using an external trusted governmental identity. As described earlier, it mainly relies on a central proxy – implementing both account linking and attribute aggregation services – used in combination with Home Organisation Identity Providers acting also as Attribute Providers.

In this model, the existing identity federations do not manage the authentication process, which is externalised to a trusted platform put in place at a governmental level for all citizens of the respective countries, and only perform the enrichment of digital identity (attribute aggregation) before granting access to the services.

One of the main advantages here is the possibility of leveraging account linking with high assurance level eGov accounts to increase the Level of Assurance (LoA) of current institutional accounts used within the federation.

4 Best Current Practices and Recommendations

Combining the needs of survey respondents with an analysis of the available architectures, this section introduces a series of best current practices for a user-managed identity solution in R&E. The term *edu-ID* is used throughout this section as this is the term adopted by all current deployers.

4.1 Stakeholders

In the edu-ID service concept, two new stakeholders, who may share some overlapping characteristics, emerge:

edu-ID System Operator

The envisioned central identity system operator is responsible for maintaining core identity-related services, e.g. operating the central IdP, and maintaining contracts and relations with attribute providers and the federation.

edu-ID Governing Body

The user-managed identity governing body oversees policy and legal relations for the service. It should ensure that its membership is carefully composed to competently represent key stakeholders.

Existing stakeholders may undergo a change in role. e.g., although campus organisations are no longer the providers of identity, by providing affiliation they are still fundamentally contributing to the effective value of the system and are therefore responsible for a critical feature.

If close relationships are built with particular identity issuers such as government eID providers, they too become stakeholders.

4.2 Policy

This section highlights some of the most important aspects of policy linked to edu-ID management. The complete policy recommendation can be found in Appendix A.

The first group of recommendations (Section A.3) deal with the ownership and governance model of the operator.

The legal status and governance of the organisation that operates the long-lasting identifier is important, since the user must place long-term trust in it. The obligations of the edu-ID towards the user should be safeguarded from changes in the governance or ownership of the provider organisation by including these explicitly in the policy document.

For example, a public entity involved in the IT background of an edu-ID Federated Identity System might be privatised because of budget constraints or as a result of an outsourcing strategy. This can negatively affect personal data sharing policies or service levels from the user's perspective, by making it either harder for their information to be shared to access services, or their information being more freely shared than they might originally have reasonably expected.

User data (see A.4) is released to third parties to enable them to access services. Users should at all times be able to review the full list of SPs as well as the eligibility criteria for adding them. Also, users should be aware of the model of the data release: is it happening in-session only or in the background (i.e. Attribute Query) as well?

Users should also be able to review a data release history. The policy should specify how long the account history is retained. This includes not only data release history but user data modifications (old values) as well.

The policy should define the assurance levels used by the Operator or endorse those from other specifications, as applicable. This can help in finding a common denominator among the very different sources of data gathered. If data for different assurance levels are treated differently, the Policy document should express this.

If the system foresees user consent (lawfulness of processing is not based on other factors, see [GDPR]), per-SP revocable consent (not all-or-nothing) should be provided and the revocation method should be explicitly stated in the policy. Because of the nature of long-lasting identifiers, the common causes of termination are expected to be demise of the user and user-initiated termination.

The policy should cover the process of user-initiated account termination, including the timeframe for the required steps. Account termination should mean the retrospective erasure of history of user activity and any other personal data. It is recommended that the technical solution of this step replace the identifiers with pseudonyms that make it impossible to trace back to the original identifier, but maintain the consistency of the underlying databases, e.g. all forgotten users can be distinguished from each other. Non-identifier fields should read 'deleted' or similar suitable value. The policy should also cover the process of operator-initiated account suspension and termination. It should contain contact information and steps for terminating the account of a deceased user and grant access to digital assets to inheritors if applicable. It is highly recommended that the Operator has in place an account expiration policy – e.g. account deletion after years of non-usage or receiving no reply to repeated contact attempts.

4.3 Governance

The governance model for edu-ID identity management should be explicitly explained in the policy document of the provider. Should this approach become successfully established in a given country, this could also mean that a large number of services might become available only through this system.

This can create a situation where participation in this system is formally optional but mandatory in practice, e.g. opting out means serious disadvantage in higher education, etc.

For this reason, it is best to approach the governance of such a system as if it wielded such power at the outset. In consideration of this, as a basic principle it could be required that the users have a meaningful representation in the decisions taken, such as on data sources to be included and data release targets in the federation. In current models, typically only the organisations are represented.

Users should be notified in advance about the upcoming addition of Service Providers and be provided with right of objection via an official channel if the change affects them. They should be represented in decisions about individual users, possibly by a delegated member forming part of an ethical committee dealing with misuse.

edu-ID operators should also strive for transparency and instruments for self-reflection and self-investigation should be available. That is, there should be an internal organisational role for auditing and the proactive, internally initiated investigation of incidents, and the results of these activities should be published. The Operator should release a compliance record regularly. This should include aggregates about authority requests and similar issues.

Governance models largely depend on structures established in each country or NREN, so it is recommended that each NREN check their own model and make sure that all stakeholders are part of the process.

4.4 Data Protection

Data protection within both the R&E sector and the Identity management context is vitally important not only in terms of legislation, but also to mitigate the impact of data breaches and misuse on a particular person or organisation.

The status of Federated and Interfederated Identity provisioning needs to be understood in the light of the upcoming implementation of the new EU General Data Protection Regulation in May 2018 [GDPR]. The GN4-2 project has produced a milestone document on the impact of this new regulation on identity federations [M9.2 DP Assessment]. The findings in this deliverable rely on the outcome of that work. Successfully understanding the provisions of the GDPR can help ensure that work done between IdPs across countries is stable and compliant. In this respect, a lifelong IdP service acts in the same way as an existing IdP and is affected in principally the same ways.

The structure of Identity Federations and the separation of user data and service data between the IdP and SP elements potentially offers great advantages within the field of GDPR as it provides a modular approach to data management, control and security. Within this distributed structure, the concept of ‘personal data’ as shared between attribute providers and the edu-ID operator as well as the reporting requirements and the ‘right to be forgotten’ need to be clearly defined.

In the case of lifelong learning and/or research, different types of data are used (and potentially shared), starting with personal data and ending (in the case of research) with commercial data, public data, and data that may refer to other individuals (for example in medical research) and may even include potentially censored information (e.g. state secrets). This data (and the history of the data) is

distributed across multiple systems in multiple organisations and potentially across multiple jurisdictions.

As implementation guidance for GDPR is still evolving, precise practices in this area are not yet defined. The requirements of the GDPR in respect of federated identity will need to be addressed with expert legal advice. The scope of this investigation is not limited to lifelong identity services but will affect and be influenced by the wider AAI environment and actual policy formation should be undertaken by the wider AAI community rather than within the scope of a specific project or activity [M9.2 DP Assessment].

4.5 Technology

In terms of architecture, the concept of user-centric federated identity does not require new technology or protocols, nor on the other hand does it have to be restricted to the technologies or protocols already commonly in use in classic identity federation. All the proposed use cases can be covered with existing technologies, protocols and tools, such as currently deployed SAML stacks, SCIM and OIDC. SAML implementations are very mature; OIDC implementations are also in production. User access to applications SAML (proven, heavy-weight) or OpenID Connect (light-weight in complexity and ease of integration) are recommended, and multiple technologies can be simultaneously supported.

To transmit user information in the context of identity management processes (e.g. de-provisioning user accounts when a user initiates self-termination), authentication protocols such as SAML or OIDC are generally not appropriate as they are usually tied to user sessions. To implement IdM processes, more flexibility and less standardisation are required. The relatively new SCIM protocol is a promising candidate for this purpose.

A REST-based communication protocol permits to create, read, update and delete users, and to add or remove them from groups, therefore for provisioning of identities generic REST-based APIs might also be considered.

4.6 Security

General Single Sign-On Identity Providers' security recommendations also apply to user-managed SSO solutions. Nonetheless, the scale is different at least in two dimensions: lifetime and quantity of managed identities. In an organisation-centric identity federation there are usually tens or hundreds of Home Organisation IdPs each managing their own user base, usually counting thousands or tens of thousands of users. A user-managed edu-ID project can easily involve several hundred thousands or millions of users. A security breach in an HO IdP may pose a serious risk for the HO's community, and in the case of an edu-ID system could potentially affect the entire community i.e. every user with data in the system and every service connected to it.

In this section, specific terminology from the ISO 27000 family of standards [ISO/IEC 27000] is used, which is shown in italics. The best practices for user-managed federated identity are rather technology-agnostic, therefore a *risk analysis* based on principles and workflows is already possible.

The main *threats* associated with such a user-managed approach are the following:

- In centralised edu-ID models, a huge amount of user information, such as personal data and account history is accumulated in one place. This increases the *level of risk* of attacks targeting that database as the *consequences* of *an information security incident* are higher. If the likelihood of such an incident can be lowered by the fact that a single place of storage and processing is easier to defend then this risk is manageable.
- In the centralised edu-ID models, the consequences of a denial-of-service attack are significant since a wide range of services can be made unavailable for masses of users with an attack on one service. It also means that account hijacking opens up several services to the attacker. The overuse of an authentication and authorisation system (that is, usage of the system as intended so excessively that it overloads the system) has not been a characteristic source of failures in the past. Denial of service attacks however might pose a *high level of risk,* as a whole community might be rendered unable to access the resources.
- The *level of risk* of eavesdropping and decryption of single messages coming in and out of the system are fairly low, since the information attainable is minimised by design (e.g. profile information, targeted id) and is envisioned as less of a target for attackers. But this depends on the actual kind of information that is going to be transmitted. The risk of users elevating their own privileges depends on the services offered.

4.7 Practices for Interoperability with Classic Federated Identity

Many aspects of interoperability are well covered by specifications, directives and policies. These are sufficiently general that they can be reused in the User-Managed Identity model. SIRTIFI should be used in incident handling. Entity categories [R&S], data protection directives [CoCo] and the ongoing REFEDS Assurance Framework [RAF] work can also be leveraged. The user-managed approach might even help the compliance to such specifications as it provides a larger, centralised, entity with more resources on the identity provider side. In current federations it is often a problem that some entities have so few resources that they can barely achieve basic functionality and never go beyond that.

4.7.1 User-Managed Identities

As the edu-ID identity is initiated by the user, which can occur without automatic affiliation with an organisation, self-asserted (or unconfirmed) identities will exist for users within a federation. This raises the double question of:

- How service providers should deal with these kinds of identities within a federation.
- How federations should manage the exposition of identities with a minimal level of confidence to eduGAIN at the interfederation level.

Regarding the first point, it is worth noting that, within the Swiss federation and just for the services available in that federation, self-asserted identities can access SWITCH-AAI service providers with no affiliation value at all.

It is indeed up to the service to verify the quality of a user identity and in case of insufficient quality (no affiliation or insufficient LoA), service providers can deny access to users or invite them to increase the quality of their identity to match the requirements. Generally, it is assumed that in the future an ever greater share of the IdPs will signal assurance.

In the light of this it is recommended that Service Providers parse assurance profile information in order to recognise self-asserted identities (and the quality of the identity and/or the attributes in general), thus being enabled to take the right access control decision. While this is particularly applicable for the edu-ID use case, it must be noted that assurance validation should become a common practice for all federated identity use cases.

On the issue of exposing self-asserted identities to eduGAIN, two different approaches have been adopted among the current edu-ID initiatives presented herein:

- In the Swedish edu-ID model, self-asserted identities are exposed to eduGAIN with a basic assurance profile (SWAMID AL1) and presented with the value *affiliate* for the 'eduPersonScopedAffiliation' attribute (i.e. <affiliate@eduid.se>).
- In contrast, in the SWITCH edu-ID model, self-asserted identities are currently not exposed to eduGAIN at all, and it is planned that only identities which strictly comply with the eduGAIN policy will be exposed to it in future.

Based on these two instances, it is recommended that self-asserted identities should not be exposed to eduGAIN, and should a federation still make that choice, it must clearly signal this by assigning a proper assurance profile to the identity and implementing any other appropriate technical measures.

4.7.2 edu-ID Identity Providers Metadata and Scope Authority

In terms of the SAML entityID specification, edu-ID systems are Identity Providers, thus metadata representation can be straightforward. Nonetheless, there are different uses of edu-ID IdPs that should be taken into account which introduce complexity and have an influence on metadata representation and scope.

In the case of the SWITCH edu-ID IdP, the final goal is to substitute the user's Institutional IdP. While that institutional IdP will still exist as infrastructure, it will only provide attributes towards the SWITCH edu-ID. The implication is therefore that the edu-ID IdP is authoritative for multiple institution scopes. In terms of metadata representation, multiple scopes authority can be implemented either by listing all the scopes for which the IdP is authoritative under a single entityID, or by having one entityID per scope, but pointing to the edu-ID authentication endpoints.

On the other hand, in the case of the SUNET eduID.se, just one IdP is published to the SWAMID federation, and thus a single entityID is authoritative for a single scope.

Finally, in the case of an edu-ID proxy architecture, while the proxy is authoritative for its own scope as an IdP, it should also be authoritative for the scopes of the affiliation attributes collected from the Institutions' Attribute Authorities.

5 Conclusions

The investigation conducted by the eduKEEP team as part the of Trust & Identity Technology Development task (GN4-2 JRA3 T3) has made it clear that, while the current organisation-centric approach is successfully in use in many contexts, opportunities exist to improve approaches to identity management to make it easier for users to change between institutions or have multiple affiliations. The results of the survey carried out confirm that many federations are aware of a need to address these issues.

The key strengths of the user-managed identity concept are its strong safeguards and transparency toward the user and the fact that it allows the creation of user-managed, lifelong identifiers. Beyond R&E, similar lifelong identifiers exist in the form of social media accounts and eGOV-ID initiatives. However, some of these approaches, in particular those of social media, lack the privacy and transparency measures that are desirable in R&E.

With edu-ID, these accounts can be technically integrated with all the tools used by the research and education community, thus taking advantage of their convenience and ubiquity while providing additional value and protection. Conversely, if no action is taken in this area by R&E organisations, these identifiers may become de-facto dominant without their being consistently integrated within the sector.

On the other hand, the user-managed identity model also introduces new concerns about the increased responsibilities of federation operators as they hold more personal data. Strong safeguards are therefore recommended, as set out in the policy recommendation template for governance structure setup provided in the appendix, and legal consultation is strongly advised on data protection considerations.

Some countries (Switzerland, Sweden) are already developing their NREN identity systems in the direction of user-managed identity. These systems have served as the models for the architecture and technology best practices proposed here.

Use cases for a user-managed, persistent, privacy preserving, and institutionally backed identity solution providing enhanced user flexibility and improved support of lifelong learning, are presented in this document. The models and use cases together aim to provide a reference for federations in addressing the limitations of organisation-centric ID and the challenges posed by the changing identity management environment.

Acknowledgments

The comments and suggestions received by the members of the identity federation community greatly improved the current work. In particular, the eduKEEP team wishes to acknowledge the contributions of:

Marina Adomeit (AMRES)

Pål Axelsson (SUNET)

Christoph Graf (SWITCH)

Christos Kanellopoulos (GÉANT)

Hans Nordlöf (SUNET)

Francesca Pegazzano (GÉANT)

Wolfang Pempe (DFN)

Special thanks also to the respondents to the user-centric identity federation survey.

Appendix A Policy Recommendation

A.1 Purpose

This document presents the potential elements of a policy for a user-managed identity system to support lifelong learning in R&E as deployed at a national level.

A.2 Definitions

Term	Description	Abbreviation
Lifelong Account	A Lifelong account is an account that the user can acquire and use for life. Example: SWITCH edu-ID account. A Lifelong Account might be identified with a Lifelong Persistent Identifier. It can also be associated with several pseudonyms and targetedId-s.	LLA
Lifelong Persistent Identifier	An identifier that is permanently assigned to a certain person for life (default case). Examples: ORCID, ResearcherId, Scopus ID. Although all these can be revoked, and an ORCID can be superseded by another identifier registered by the same user, in the default case these are lifelong unchanging identifiers.	LLI
Provider Organisation	The organisation that is hosting the Lifelong Account. This is deliberately not referred to as 'home organisation' so as not to overlap with the Organisation-Centric concept.	PO
Auxiliary Identifiers	All identifiers associated to the account that are not lifelong. For instance both ORCID and SWITCH edu-ID rely on email addresses, but these may belong to institutions the user leaves (e.g. university) and this should be changed in the LLA. Targeted Ids and Pseudonyms may also be auxiliary identifiers, although these can also be lifelong.	AXI

A.3 I - On the Legal Entity of the Operator of Lifelong Persistent Identifiers

The contract between the PO and the user could last for 50+ years. Besides the obvious opportunities this presents, it also brings in new challenges.

The legal status and governance of the organisation that operates the long-lasting identifier are important, since users must be able to place their complete trust in it in the long term.

***I/1 - Nature of PO**: The ownership (public/private) and the governance of the organisation are especially important. Both of these might change several times while the brand remains unchanged.*

***I/2 - Ownership and governance** are crucial in the context of personal data protection. Therefore, it is justified to place safeguards in the policy so that personal data is not automatically carried over to the changed organisation without the user's consent.*

***I/3 -** The policy document should explain the **ownership model of the PO** and the governance **model of LLI.** The former should explain who is ultimately responsible for the personal data stored in the system. The latter should explain **who is authorised** to access the personal data, share it with authorities, and remove and suspend accounts.*

For example, a public entity involved in the IT background of a long-lived entity might be privatised because of budget constraints or as a result of an outsourcing strategy. An entity belonging to a fairly autonomous university might be transferred to a ministry, meaning that it is effectively placed under direct government control. The situation of the data can also change without ownership change. An academic institution might be mandated to have a government-appointed superintendent on its board.

This is why it would not be unreasonable to protect the users from not only obvious legal changes (PO is overtaken, dissolved), but also the less obvious but significant PO governance changes.

***I/4 -** Users should be **notified about any changes in Ownership or Governance** before they occur. Notification before an actual change should be given early enough to ensure that users have a **reasonable time frame within which to opt-out by terminating their accounts**.*

A.4 II - Data Handling

A.4.1 Basics

For a user, creating an LLA means entering into a very long relationship with the identity provider. To earn the user's trust for such a commitment, the presence of strong safeguards within the policy governing such relationships is justified. The policy elements below serve this purpose.

An LLA involves releasing user information when necessary to the right services. In this respect, a crucial point of a policy should be to define with precision both the **list of potential recipients** and the **timing of the data releases** to service providers.

II/1 - *The PO should maintain a* ***list of current SPs*** *including details about their data consumption and the timing of the data releases to them - e.g. only in user session, or also back-channel without the presence of the user.*

II/2 - *The list of SPs* ***should retain history****, meaning that the date of introduction of new SPs and removal of old SPs should be recorded and published.*

The policy must include a **long-term definition of what services belong to the federation** - not just the current list, but also the generic criteria set for which kind of organisations can become SPs in future.

II/3 - *The policy has to include* ***eligibility criteria*** *for adding new SPs.*

Another important aspect is to exactly define what is being released. This means giving the **attribute name and value** (because the released value might be different per service provider), and also the identifier, as it could be targeted to the SP.

II/4 - *A full, detailed, per-SP* ***attribute release configuration*** *should be available to the users.*

About the identifier, it is important to clarify what it might reveal about the identity of the user – preferably nothing.

II/5 *- Attribute Release should be based on* ***per-SP revocable consent.*** *This does not necessarily mean a pop-up or other intrusive technique, but should include a user-facing* ***list of previously given consents*** *and the possibility of removing them.*

II/6 - The format of the identifier *should be an explicit part of the policy.*

II/7 - Identifier reuse *(giving an old, used identifier to a new user) should be prevented by policy.*

The re-assignment of identifiers has no practical benefits, but is able to create problems in obvious ways, and therefore should be ruled out.

A.4.2 Transparency

II/8 - Data release history: users should be able to access their full, detailed data release history that contains both attribute names and values, along with the times and recipients.

The expiration and removal of such history might be constrained by the overall legal framework. For example, the PO might be required by law to retain certain elements of the history for a given period of time.

II/9 - The policy should specify **how long the account history is retained.** This includes not only data release history but user data modifications (old values) as well.

A.4.3 Compliance Report

Nowadays big end-user IT companies such as Facebook and Google release aggregate numbers of policy/secret service data requests fulfilled and denied. In this way they can showcase their relative autonomy and integrity. Also these numbers are usually so low (e.g. a couple of hundreds or thousands per billions of users) that they really communicate a message that authorities' access is not a big issue. Therefore, sharing this information is both ethically desirable and useful for the organisation's image.

II/10 - *The PO should release a* ***compliance record regularly.*** *This should include aggregates about authority requests and similar issues.*

A.4.4 Assurance Levels

II/11 - The policy should define the assurance levels used by the PO, if applicable. If data of different assurance levels are treated differently, the Policy document should express this.

A.5 III - End of Life

Users should be able to initiate the termination of their LLA. This termination should occur by a reasonable deadline.

III/1 - *The policy should contain the* ***process of user-initiated account termination****, including the time frame of the steps.*

Accounts might be suspended or terminated by the PO for several reasons, including abuse of the system, gross violations of the terms of use, etc.

III/2 - *The policy should contain the process of PO-initiated account suspension and termination. The policy should define the governance model for this, e.g. an ethical committee or a decision mechanism.*

A lifelong account means that the policy should address the issue of the demise of the service users.

III/3 - *The policy should contain contact information and process for* ***terminating the account of the deceased*** *and grant access to digital assets to inheritors if applicable.*

Relatives of the deceased might not get around to initiating the termination of the account. Therefore, it is important to have other means to remove such accounts.

III/4 - *It is* ***highly recommended*** *for the PO to operate an account expiration policy - e.g. after years of non-usage or receiving no reply to repeated contact attempts, the account will be deleted.*

A user might self-terminate an account and at a later time apply for an account again. If enough time has passed, the old record is already deleted, as per the data retention policy. If the user is allowed to create a new account then this can be seen as the implementation of a right to forget policy. Preventing this seems not to be feasible. For example, in order to ban re-registrations, records about past accounts would have to be kept indefinitely (otherwise there would be no way of knowing that a

registration is in fact a re-registration). In this respect it is also important whether the user will be allowed to get the same ID again (possible if the ID is generated from unchanged personal details).

III/5 - *The Policy should* ***handle re-registrations****. The most reasonable approach is to allow it, unless the system envisages the option of a lifelong ban.*

Appendix B Comparison Criteria for edu-ID Architectures

A - Target audience

The target audience is the end-user population for which the edu-ID solution is intended (i.e students, R&E community, etc.)

B - Long-term Identity

The identity is reusable and does not have a limited lifetime bound to an affiliation. For example, a long-term identity provides students with an account they can use from their first university application until they are become alumni, and possibly researchers or members of the university staff.

C - Identity suitable for AuthN

The created identity is paired with a credential, thus it can be used by the user for authentication purposes. In other words, users have their own personal account (as a result of the new edu-ID creation process) that they can use to authenticate themselves.

D - A new identifier is provided

When identity for a new individual is created on the central AAI platform (Central IdP or IdP/SP Proxy), this latter assigns him/her a permanent, opaque, non-re-assignable identifier.

D1 - Persistent identifier

A persistent identifier stands for an identifier which is immutable. The same identifier is used over time. In an authentication context, persistent means that the identifier is not a per-session identifier, but is a stable value, and is fundamental to link an account to resources hosted on a Service Provider.

D2 - Globally Unique Identifier

As part of a digital identity, a globally unique identifier is any identifier which ensures identifying a user in a unique way across different domains. Globally, Unique Identifiers are usually composed of two parts:

- a locally unique part, usually generated leveraging UUID.
- a scope that represents the issuer entity.

In a federation context, when required, a unique identifier can be shared among all the different entities of the federation, and enables collection of items of information from different sources and tying them to the same digital identity and thus to the same user.

E - Central external IdP acting on behalf of Home Organisation IdPs

An organisation externalises the responsibility of user authentication to an external IdP (i.e. outside of the organisation's direct control), instead of using its own home IdP. The IdP responsible for the authentication will also act on behalf of the Home Organisation, meaning that it is capable of asserting attributes scoped with the Home Organisation domain.

The external IdP can be a single central IdP or an IdP/SP proxy.

F - Account Linking

Account linking occurs in the case in which a single user has different accounts/credentials on different entities (IdPs or services).

It is a process that permits to create a persistent association between distinct accounts belonging to the same user. Account linking is accomplished coupling unique identifiers related to each user account.

F1 - Linked Account AuthN

Given that two or more user accounts have been linked at an entity (an SP, an IdP/SP proxy, or a Central IdP), 'linked account authentication' refers to the possibility of using any of the linked accounts to perform authentication.

G – Self-asserted Identity

The digital identity of the user is issued by the user him/herself without additional validation by a third-party or identity vetting procedures.

G1 - Identity Assurance Elevation

The concept of Identity Assurance assesses the strength and rigor of the identity proofing, the authentication, and the attributes' reliability. It is generally based on different assurance levels (from low to high) or profiles with different sets of characteristics.

"Identity Assurance Elevation" typically refers to the case where a user jumps from a non-validated ID (self-asserted identity) to one that is proofed by a trusted party.

G2 - VO-based vetting

The identity of the user is vetted by leveraging the different identity check mechanisms associated to a Virtual Organisation's (VO) enrolment process (invitation, user-pre-registration or admin approval).

H - Attribute Aggregation at IdP

After authentication, the digital identity of the user is enriched with further information (i.e. attributes) coming from other components of the architecture (Attribute Authorities).

This process of identity enrichment is taken care of at the IdP level (performed by the central IdP or the IDP/SP proxy), thus the Service Provider will receive all the aggregated information without having to query any other entity out of the IdP.

I - Attribute Release Policy – Delegate Management to Home Organisations

The delegation of ARPs management allows the Home Organisation to autonomously manage how its users' attributes (affiliations, group membership, etc.) are to be released to target services, instead of having this operation performed by a central third-party (typically the Federation Operator).

A dedicated tool with web UI (providing an abstract view of the underlying user repository) is generally needed to this end.

References

[eduGAIN]	https://www.geant.org/Services/Trust_identity_and_security/eduGAIN
[eIDAS]	https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eIDAS+Profile
[FOG]	https://wiki.refeds.org/display/GROUPS/FOG
[GDPR]	http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC
[GÉANT]	https://www.geant.org/
[GN4-1_D15.1]	https://www.geant.org/Projects/GEANT_Project_GN4-1/Documents/D15-1_Report-on-the-Achievements-of-JRA3-T1-and-Recommendations-on-Future-Work.pdf
[CoCo]	https://www.geant.org/uri/Pages/dataprotection-code-of-conduct.aspx
[ISO/IEC 27000]	https://www.iso.org/standard/66435.html
[M9.2 DP Assessment]	https://www.geant.org/Projects/GEANT_Project_GN4/deliverables/M9-2_Assessment-of-DP-Legislation-Implications.pdf
[OIDC]	http://openid.net/connect/
[RAF]	https://wiki.refeds.org/display/GROUPS/Assurance+Working+Group
[R&S]	https://refeds.org/category/research-and-scholarship
[REFEDS]	https://refeds.org/
[SAML]	https://www.oasis-open.org/committees/security
[SWITCH edu-ID]	https://www.switch.ch/edu-id/
[SUNET eduID.se]	https://www.eduid.se/en/
[SWAMID AL1]	https://www.sunet.se/swamid/policy/al1/
[SWAMID AL2]	https://www.sunet.se/swamid/policy/al2/

Glossary

Account Linking	An association between distinct accounts identifying the same user.
Attribute Aggregation	The process of enriching a digital identity with further attributes coming from other components of the architecture or other parties.
Auxiliary identifiers	All identifiers associated to the account that are not lifelong. For instance, both ORCID and SWITCH edu-ID rely on email addresses, but those may belong to institutions the user leaves (e.g. university) and this should be changed in the user-managed lifelong account. Targeted Ids and Pseudonyms may also be auxiliary identifiers, although those can be lifelong, too.
GDPR	European General Data Protection Regulation.
HO	Home Organisation
IdM	Identity Management
IdP	Identity Provider. The party responsible for providing the authentication service and to release the attributes and the identifier that represent the user's identity.
LoA	Level of Assurance. The amount of confidence that can be placed in an identity is known as assurance, and it has been traditionally defined by assigning a Level of Assurance (LoA). More recently, the concept of levels related to assurance has been questioned, and the latest assurance standards are instead employing a components or profile (or both) based assurance framework.
OIDC	OpenID Connect
Persistent identifier	A persistent identifier stands for an identifier which is immutable.
PO	Provider Organisation
REST	REpresentational State Transfer
SAML	Security Assertion Markup Language

SCIM	System for Cross-domain Identity Management
SIRTIFI	Security Incident Response Trust Framework for Federated Identity
SLA	Service Level Agreement.
SP	Service Provider. The party responsible for providing services to authenticated users.
SSO	Single Sign-On
TOU	Terms of Use
VO	Virtual Organisation