30-04-2019

Deliverable DJRA1.4: Evolution of the AARC Blueprint Architecture

Deliverable DJRA1.4

Contractual Date:	31-03-2019
Actual Date:	30-04-2019
Grant Agreement No.:	653965
Work Package:	JRA1
Task Item:	JRA1.3
Lead Partner:	GRNET
Document Code:	DJRA1.4

Authors: AARC Consortium Partners;AppInt members;Nicolas Liampotis (ed.)

Abstract

The AARC blueprint architecture provides a set of building blocks for software architects and technical decision makers who are designing and implementing access management solutions for international research collaborations. This document describes the evolution of the AARC Blueprint Architecture, starting with a summary of the changes since AARC-BPA-2017. It also describes the community-first approach which enables researchers to use their community identity for accessing services offered by different infrastructures. The deliverable then presents extracts from the set of guidelines and informational documents that accompany the current iteration of the Blueprint Architecture.

The research leading to these results has received funding from the European Community's Horizon2020 Programme under Grant Agreement No. 730941 (AARC2).

Document Revision History

<This section to be deleted or hidden before publication of document>

Version	Date	Description of change	Person
1	2019-04-15	First draft issued	N. Liampotis
2	2019-04-24	Second draft issued	N. Liampotis
3	2019-04-30	Final version issued	N. Liampotis
		Review	
		Approved	

Table of Contents

Table of Figures

Table of Tables

Executive Summary

The AARC Blueprint Architecture (BPA) builds on top of eduGAIN and adds the functionality required to support common use cases within research collaborations, such as access to resources based on community membership. The AARC BPA champions a proxy architecture in which services in a research collaboration can connect to a single point, the SP-IdP-Proxy (hereafter termed "proxy"), which itself takes the responsibility for providing the connection to the identity federations in eduGAIN, thus reducing the need for each service having to separately connect to an identity federation/eduGAIN.

This document describes the evolution of the AARC BPA, including a summary of the changes since [AARC-BPA-2017]. The current iteration of the BPA (AARC-BPA-2019) comprises five component layers grouped by their functional role:

- *User Identity* - groups services which provide electronic identities that can be used by users participating in International Research Collaborations.
- *Community Attribute Services* - components related to managing and providing information (attributes) about users, such as community group memberships and roles, on top of the information that might be provided directly by the identity providers from the User Identity Layer.
- *Access Protocol Translation* - addresses the requirement for supporting multiple authentication technologies and defines an administrative, policy and technical boundary between the internal/external services and resources.
- *Authorisation* - contains components for controlling access to services and resources.
- *End-services* - where the external services interact with the other elements of the AAI.

The current iteration of the BPA focuses on the interoperability aspects, to address an increasing number of use cases from research communities requiring access to federated resources offered by different infrastructure providers. Hence the "community-first" approach, which introduces the Community AAI. The purpose of the Community AAI is to streamline researchers' access to services, both those provided by their own infrastructure as well as services shared by other infrastructures. Specifically, in the community-first approach, we can distinguish among three types of services that can be connected to the Community AAI:

1. community services - provided only to members of a given community
2. generic services - provided to members of different communities
3. infrastructure services - provided by a given research infrastructure or e-Infrastructure to one or more Community AAI (typically through a dedicated infrastructure proxy)

AARC-BPA-2019 is accompanied by a set of guidelines and informational documents.

There are documents that provide guidance on the interoperable expression of information, including:

- community user identifiers [AARC-G026]
- group membership and role information [AARC-G002]
- resource-specific capabilities [AARC-G027]
- affiliation information [AARC-G025]

Furthermore, based on the analysis of the authorisation architectures from nine different use cases detailed in [AARC2-DJRA1.2], we have identified three main authorisation models in [AARC-I047] that make use of an SP-IdP-Proxy,

1. Centralised Policy Information Point: the proxy aggregates user attributes, such as group membership information and roles, and makes them available to the end-services
2. Centralised Policy Management and Decision Making: the proxy conveys the authorisation decision to the end-services in the form of capabilities

3. Centralised Policy Management and Decision Making and Enforcement: the proxy enforces the decision directly at the proxy

The problem of combining assurance information associated with one or more external identities linked to the community identity is addressed in [AARC-G031]. The provided guidelines also include compensatory controls for assessing assurance component values in the absence of assurance information from the external identity provider.

In [AARC-G049], a portable and technology-agnostic way is defined for allowing services to receive hints about which identity provider to use. This mechanism (termed “IdP hinting”) can greatly simplify the discovery process for the end-user, by either narrowing down the number of possible IdPs to choose from or by making the actual selection process fully transparent.

1 Introduction

The purpose of the AARC Blueprint Architecture (BPA) is to provide a set of interoperable architectural building blocks for software architects and technical decision makers, who are designing and implementing access management solutions for international research collaborations. During the last two years, the AARC project continued to work closely with e-infrastructures, research infrastructures, research communities, AAI architects, and implementers to evolve the AARC BPA through a better understanding of the experiences and needs regarding sharing and accessing resources within research collaborations.

In order to address an increasing number of use cases from research communities requiring access to federated resources offered by different infrastructure providers, the work during the second phase of the project focused on interoperability aspects, including the interoperable expression of attributes across BPA-compliant AAIs. Hence the “community-first” approach, which introduces the Community AAI. The Community AAI streamlines researchers’ access to services, both those provided by their own infrastructure as well as services shared by other infrastructures.

The remainder of this document is organised as follows: Chapter 2 presents the latest iteration of the AARC Blueprint Architecture (AARC-BPA-2019), starting with a summary of the changes since [AARC-BPA-2017]. Chapter 2 also describes the community-first approach for facilitating researchers' access to services and resources offered by different infrastructures. The deliverable then presents extracts from the guidelines and informational documents that accompany AARC-BPA-2019. Specifically, Chapter 3 groups together the sets of guidelines for expressing attributes, including community user identifiers, group membership, resource capabilities and affiliation information. Chapter 4 provides best practices for managing authorisation, specifically targeting models for community-based authorisation. Chapter 5 provides guidelines for combining assurance information and for assessing assurance component values in the absence of assurance information from the external identity providers. Chapter 6 describes a portable and technology-agnostic mechanism for simplifying the IdP discovery process for the end-user. Lastly, conclusions are drawn.

2 Evolution of the AARC Blueprint Architecture

As shown in Figure 2.1, the latest iteration of the AARC Blueprint Architecture (AARC-BPA-2019) defines five (5) component layers, namely, User Identity, Access Protocol Translation, Community Attribute Services, Authorisation and End Services. Each layer groups one or more components based on their functional role.

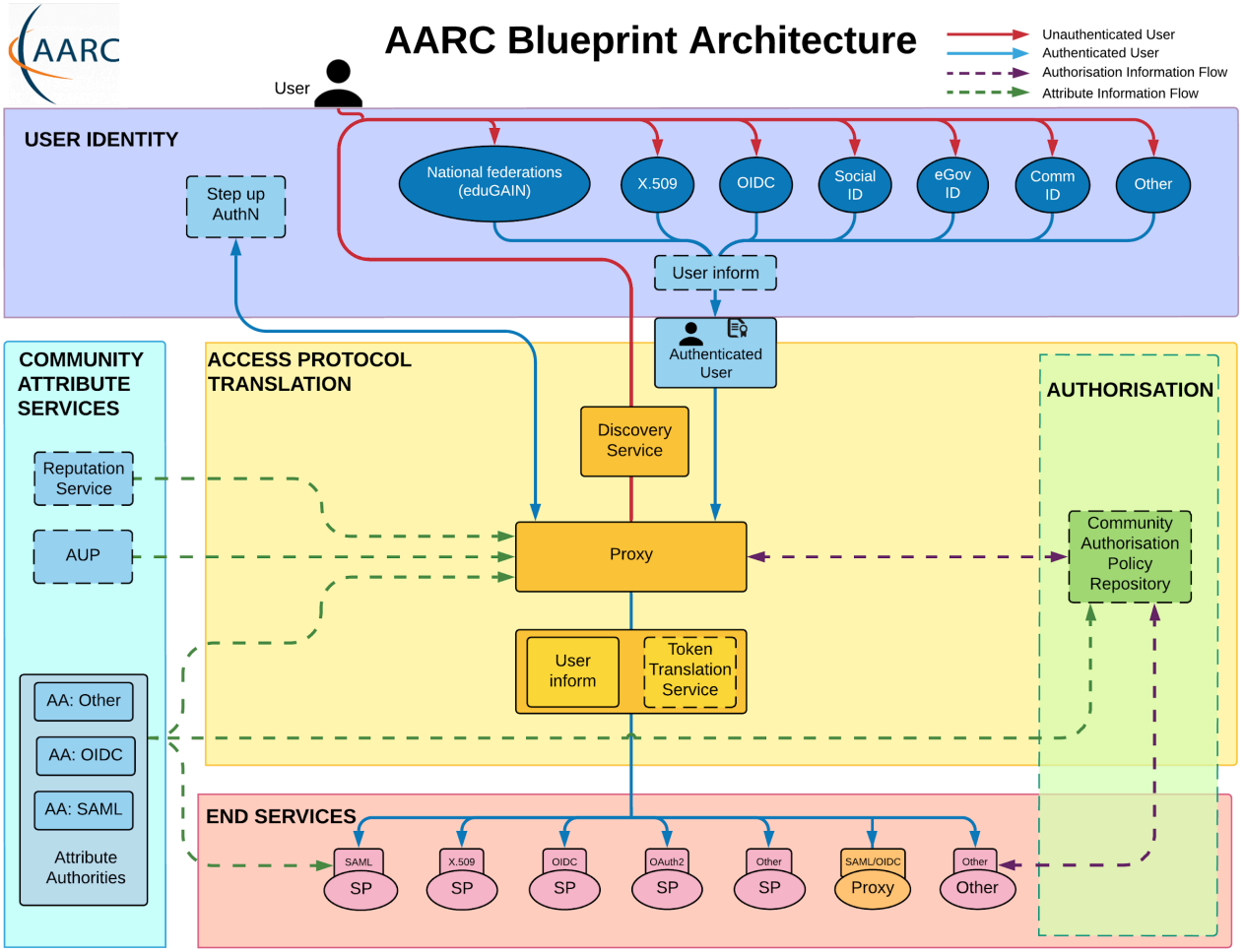

Figure 2.1: AARC Blueprint Architecture (AARC-BPA-2019)

The *User Identity Layer* contains services for the identification and authentication of users. In existing implementations in the research and education space, these services typically include Security Assertion Markup Language (SAML) identity providers, certification authorities and, more recently, OpenID Connect (OIDC) or OAuth2 Providers (OPs). Although the focus of the services in this layer is to provide user authentication, often some end-user profile information is released as part of the authentication process.

The *Community Attribute Services Layer* groups services related to managing and providing information (attributes) about users. Typically, they provide additional information about the users, such as community group membership and roles, on top of the information that might be provided by services from the User Identity Layer.

The *Access Protocol Translation Layer* addresses the requirement for supporting multiple authentication technologies. It includes the following services:

- SP-IdP-Proxy (proxy), which serves as a single integration point between the Identity Providers from the User Identity Layer and the Service Providers in the End Services Layer. Thus, the proxy acts as an SP towards the Identity Federations for which this proxy looks like any other SP, while towards the internal SPs it acts as an IdP.
- Token Translation Services, which translate identity tokens between different technologies.
- Discovery Service, which enables the selection of the user's authenticating IdP.
- User inform, which allows users to be informed regarding the processing of their personal data

The *Authorisation Layer* controls access to the End Services Layer. The AARC BPA allows the implementers to delegate many of the complex authorisation decisions to central components, which can significantly reduce the complexity of managing authorisation policies, and their evaluation for each service individually.

The *End Services Layer* contains the services users want to use. Access to these services is protected (using different technologies). These services can range from simple web-browser-based services, such as wikis or portals for accessing computing and storage resources, to non-web-browser-based resources such as APIs, login shells, or workload management systems.

2.1 Revisions since AARC-BPA-2017

The current version of the AARC blueprint architecture builds upon the previous one [AARC-BPA-2017] (depicted in Figure 2.2), while retaining full backwards compatibility. As shown in Figure 2.1, it retains the same five layers, each of which includes one or more functional components, grouped by their complementary functional roles. The User Identity Layer, the End Services Layer and the Authorisation Layer are still there, while the User Attribute Services Layer has been renamed Community Attribute Services Layer (see definition of Community Identity in the Glossary) and the Identity Access Management (IAM) Layer has been renamed Access Protocol Translation Layer and retains its prominent role in the architecture. Within the Access Protocol Translation Layer, the layout of the Token Translation Service (TTS) has been updated to better visualise the role of the TTS in the flow of attributes between the proxy and the connected services. It is worth noting that, in the new version of the architecture, "User consent" has been renamed "User inform" to indicate the points where users (data subjects) need to be informed regarding the processing of their personal data. This change is in line with the current consensus [CORMACK1, CORMACK2, AARC-G016, AARC-G042] which considers *legitimate interest*, rather than *consent*, the correct legal basis (Article 6.1(f) of [GDPR]) for the processing of personal data in the context of granting access to resources for collaborative and research communities, which is typically done for professional reasons.

The reader will note another proxy among the end services. A proxy is by definition a *service* for the IdPs facing it, and it is sometimes possible to daisy-chain proxies. This approach enablesaccess to resources offered by infrastructures through infrastructure proxies, as described in the following section.

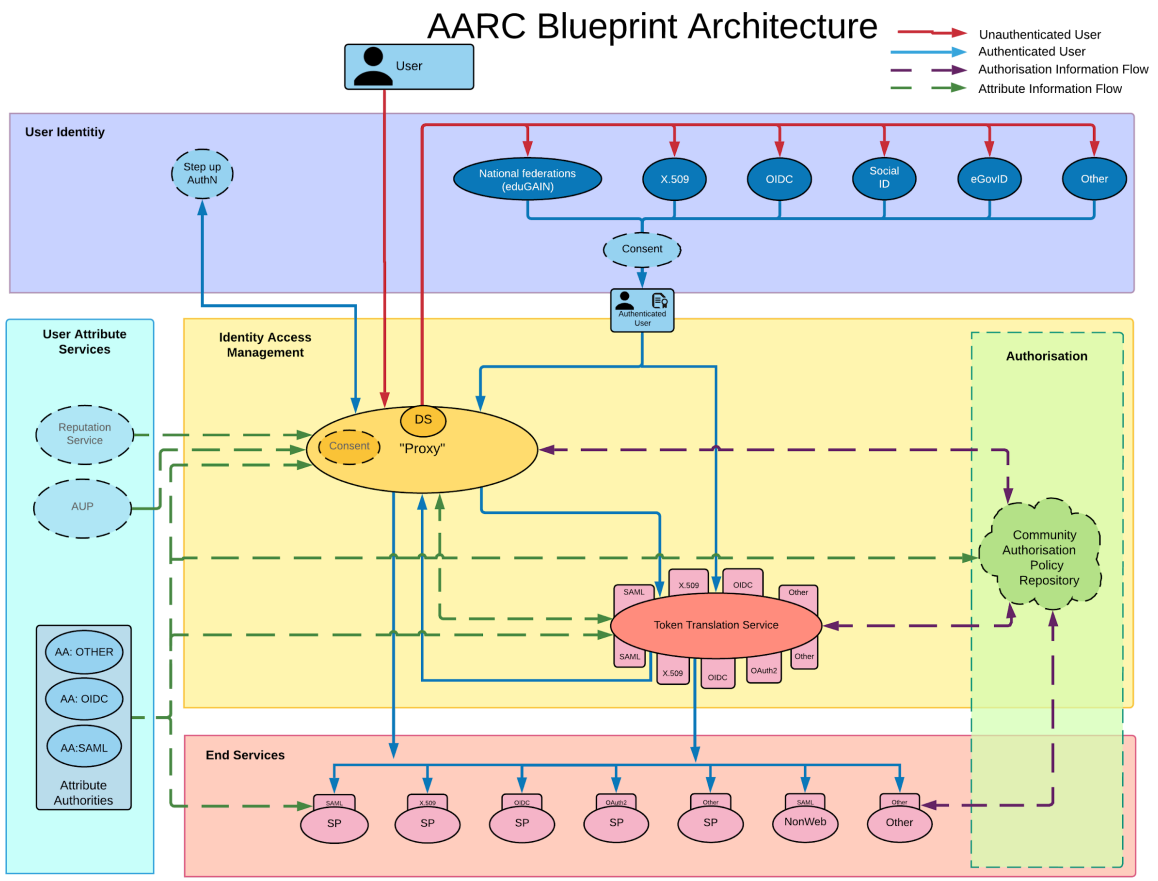

Figure 2.2: AARC Blueprint Architecture (AARC-BPA-2017)

2.2 Community-first approach to the AARC Blueprint Architecture

This section focuses on the interoperability among AARC BPA compliant AAIs that are operated by different research and e-Infrastructures. Interoperability is needed by research communities requiring access to federated resources offered by different infrastructure providers. Hence the “community-first” approach, which introduces the Community AAI. The purpose of the Community AAI is to streamline researchers’ access to services, both those provided by their own infrastructure (if they have one) as well as services provided by infrastructures shared with other communities. User authentication to the Community AAI uses primarily institutional credentials from national identity federations in eduGAIN, but, if permitted by the community, can also use other IdPs.

The Community AAI follows the proxy-based architecture shown in Figure 2.1. It can therefore add attributes to the federated identity that in turn can enable services to control access to their resources. Furthermore, the Community AAI is responsible for dealing with the complexity of using different identity providers with the services offered to the community. We can distinguish among three types of services:

1. community services - provided only to members of a given community
2. generic services - provided to members of different communities (e.g. the RCauth.eu Online CA service)
3. infrastructure services - provided by a given research infrastructure or e-Infrastructure to one or more Community AAI (typically through a dedicated infrastructure proxy)

As shown in Figure 2.3, community-specific services only need to connect to a single identity provider, i.e. their Community AAI. On the other hand, generic services need to connect to multiple Community AAIs in order to serve different communities. Being connected to multiple Community AAIs requires generic services to provide some form of IdP discovery, in order to be able to redirect the user to the relevant Community AAI[1]. Additionally, the generic services should support some means of doing "IdP hinting" (see Chapter 6), thereby allowing "community branding" of the service and automatically redirecting the user to the corresponding Community AAI.

Communities may also require access to various services which themselves are behind (another) proxy, as often is the case with resources offered by e-Infrastructures or Research Infrastructures (Infrastructures hereafter). These Infrastructure Proxies can be connected to different Community AAIs - see Figure 2.3. So, just as for the generic services, Infrastructure services should be able to hint to the Infrastructure Proxy which Community AAI to use (see Chapter 6).

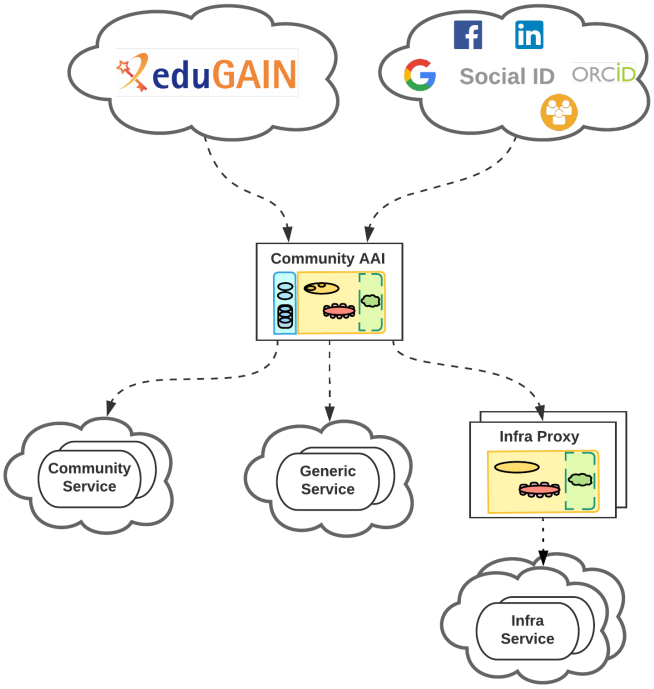

Figure 2.3: Community-first approach based on the AARC Blueprint Architecture. Researchers access services/resources using their institutional (eduGAIN), social or community-managed IdP via their Community AAI. Community services are connected to a single Community AAI, whereas generic services can be connected to more than one Community AAI. e-Infrastructure services are connected to different Community AAIs through a single infrastructure SP proxy.

It should be noted that the "community-first" approach does not impose a requirement on communities to deploy and operate a Community AAI on their own. Communities could make use of either dedicated or multi-tenant deployments of AAI services operated by a third-party, typically a generic e-Infrastructure. A multi-tenant AAI service deployment supports different communities, as depicted in Figure 2.4. It typically appears as a single entity to its connected IdPs and SPs. Such multi-tenant deployments are aimed at medium-to-small research communities/groups or individual researchers. Yet it should be emphasised that also in the multi-tenant AAI scenario, the community managers are responsible for managing their community members, groups and authorisation attributes.

[1] Primarily to get the user's identity via the community IdP, but also potentially to obtain attributes from community attribute authorities.

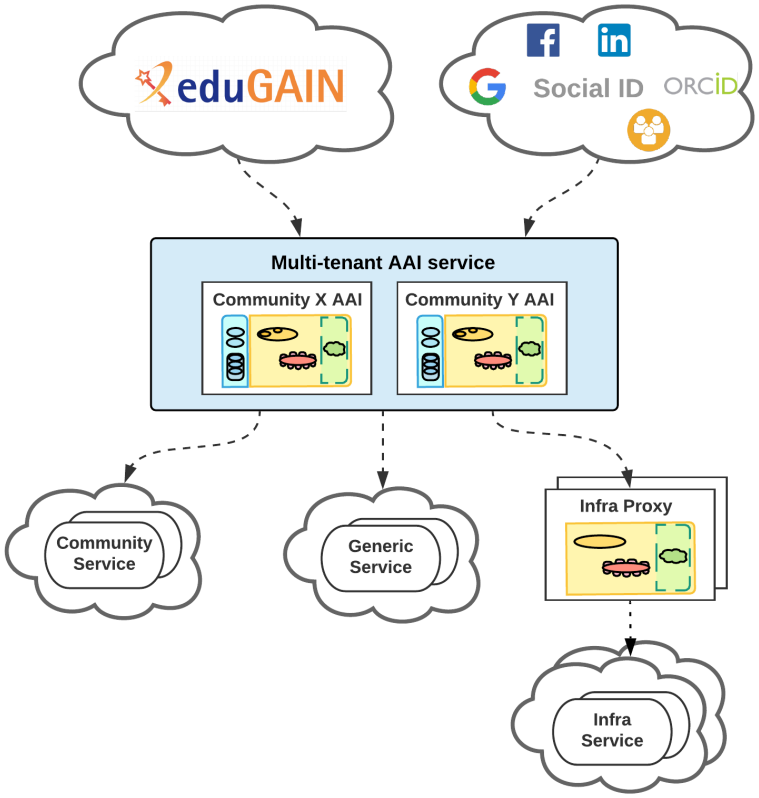

Figure 2.4: Multi-tenant deployment of AAI services in “community-first” approach to the AARC Blueprint Architecture.

3 Interoperable expression of information

This chapter provides guidelines for expressing attributes interoperably across AARC BPA compliant AAIs.

3.1 Community user identifiers

This section describes how to express community user identifiers such that the values can be transported in an interoperable way across AARC BPA compliant AAIs (see [AARC-G026]).

3.1.1 General guidelines

The community user identifier:

- MUST be assigned so that no two values created by distinct identity systems could collide.
- SHOULD be opaque
- once assigned, MUST NOT be reassigned to another principal
- SHOULD be permanent
- MUST be persistent
- MUST be expressed as a combination of two distinguishable components, namely the <uniqueID> and the <scope>, where:
    - The <uniqueId> component MUST follow the syntax of the unique id component of the [SAML-SubjectID-v1.0], as defined in Section 3.3.1 of the specification.
    - The <scope> component MUST be a domain controlled by the issuing entity or the community for which the identifier has been issued for.
    - The <scope> component MUST follow the syntax of the scope component of the [SAML-SubjectID-v1.0], as defined in Section 3.3.1 of the specification.
    - Value comparison of community user identifier components MUST be performed case-insensitively.

3.1.2 Considerations for different federated identity protocols

This section discusses protocol-specific considerations for expressing community user identifiers.

3.1.2.1 Security Assertion Markup Language 2.0 (SAML)

The community user identifier SHOULD be communicated using the General Purpose Subject Identifier (general purpose subject-id), formatted as `<uniqueID>@<scope>`, as defined in [SAML-SubjectID-v1.0].

3.1.2.1.1 NameID considerations

The use of the general purpose subject-id attribute is meant as a replacement of the `<saml:NameID>` element as a means for identifying users. However, some SAML profiles such as the Single Logout Profile, require the use of a `<saml:NameID>` element. Therefore, it is RECOMMENDED that the

urn:oasis:names:tc:SAML:2.0:nameid-format:transient NameID Format be used in conjunction with the release of the subject-id attribute.

3.1.2.1.2 REFEDS Research and Scholarship Entity Category compliance considerations for IdP Proxies

The REFEDS Research & Scholarship (R&S) specification [REFEDS-R&S] defines a bundle of attributes that Identity Providers are encouraged to release to R&S services. This bundle includes a shared user identifier, which is defined as a persistent, non-reassigned, non-targeted identifier. An Identity Provider that exhibits the R&S entity attribute in its metadata to indicate support for the R&S Category is REQUIRED to release the shared user identifier. According to [REFEDS-R&S], the shared user identifier is defined to be the eduPersonPrincipalName (ePPN) attribute (if non-reassigned). Therefore, to implement support for the R&S category, an IdP SHOULD also transport the subject-id value via the ePPN attribute. Note that the syntax of the ePPN value is compatible with that of the subject-id.

3.1.2.2 OpenID Connect (OIDC) and OAuth 2.0

The community user identifier SHOULD be communicated using the standard public sub claim, formatted as <uniqueID>@<scope>. The OIDC specification [OIDC-CORE-v1.0] requires the sub claim to be present in the UserInfo Response and in the ID Token, assuming the openid scope has been requested.

3.2 Group membership and role information

Information about the groups a user is a member of is commonly used by SPs in order to authorise user access to protected resources (see also Section 4.1). This section provides a URN namespace specification for expressing group membership and role information such that the values can be uniformly interpreted across infrastructures (see [AARC-G002]). The values should be communicated using the eduPersonEntitlement attribute (which is multivalued) [EPE].

3.2.1 Syntax

An eduPersonEntitlement attribute value expressing group membership and role information has the following syntax (components enclosed in square brackets are OPTIONAL):

<NAMESPACE>:group:<GROUP>[:<SUBGROUP>]...[:role=<ROLE>]#<GROUP-AUTHORITY>

where:

- <NAMESPACE> is in the form of
  urn:<NID>:<DELEGATED-NAMESPACE>[:<SUBNAMESPACE>]...
  where:
  - <NID> is the namespace identifier associated with a URN namespace registered[2] with IANA, as per [RFC8141], ensuring global uniqueness. Implementers can and should use one of the existing registered URN namespaces, such as urn:geant [URN-GEANT] or urn:mace [URN-MACE]
  - <DELEGATED-NAMESPACE> is a URN sub-namespace delegated from one of the IANA registered NIDs to an organisation representing the e-infrastructure, research infrastructure or research collaboration. It

[2] Generic top level namespaces require IANA approval as per Section 6.2 of [RFC8141]: https://www.iana.org/assignments/urn-namespaces/urn-namespaces.xhtml

is recommended that a publicly accessible URN value registry for each delegated namespace is provided.

A <NAMESPACE> can have a variable number of elements. For example `urn:geant:edugain`, `urn:geant:nikhef.nl` and `urn:geant:nikhef.nl:idm` are all valid <NAMESPACE> values.

- the literal string **"group"** indicates an eduPersonEntitlement value expressing group membership information;
- <GROUP> is the name of a Virtual Organisation (VO), research collaboration or a top level arbitrary group. Group names MUST be unique within a given namespace;
- an optional list of <SUBGROUP> components represents the hierarchy of subgroups in the <GROUP>;
- the optional <ROLE> component is scoped to the rightmost (sub)group; if no subgroup information is specified, the role applies to the top level group/VO;
- <GROUP-AUTHORITY> is a non-empty string that indicates the authoritative source for the entitlement value. For example, it can be the FQDN of the group management system that is responsible for the identified group membership information. The <GROUP-AUTHORITY> is specified in the f-component of the URN ([RFC8141], Section 2.3.3); thus, it is introduced by the number sign ("#") character and terminated by the end of the URN. Any characters outside the ASCII range that appear in the <GROUP-AUTHORITY> MUST be percent-encoded using the method defined in Section 2.1 of the generic URI specification [RFC3986]. As described in Section 3.2.2, the <GROUP-AUTHORITY> MUST NOT be taken into account when determining equivalence of URN-formatted eduPersonEntitlement values expressing group membership and role information.

3.2.2 Semantics

Each eduPersonEntitlement attribute value represents a particular position of the user within a VO, research collaboration or generally a top level arbitrary group. A user may be a member or hold more specific roles within the groups associated to this top level group. Groups are organised in a tree structure, meaning that a group may have subgroups, which in turn may have subgroups, etc.

This hierarchical structure implies that if someone is member of a subgroup, then they are also member of the parent group. For example:

<NAMESPACE>:group:*parent-group*:*child-group*#<GROUP-AUTHORITY>

implies membership in *parent-group*, i.e.:

<NAMESPACE>:group:*parent-group*#<GROUP-AUTHORITY>

Ownership of any role always implies membership of that particular (sub)group. However, holding a more specific role in a subgroup does not imply the same role in the parent group. For example:

<NAMESPACE>:group:*parent-group*:*child-group*:role=*manager*#<GROUP-AUTHORITY>

implies plain membership in both *child-group* and *parent-group*, but NOT:

<NAMESPACE>:group:*parent-group*:role=*manager*#<GROUP-AUTHORITY>

Conversely, asserting a role in *parent-group* does not imply that the person has the same role (or a role with the same name) in *child-group*; if the person is a member of *child-group* and has the same role in *child-group*, then an extra eduPersonEntitlement value is needed to communicate this.

Determining if two eduPersonEntitlement values refer to the same group membership (and role, if specified) requires testing for URN-equivalence as per Section 3 of [RFC8141]. Thus, the mandatory group authority information specified in the f-component of the URN MUST be ignored in this process. For example, the following two URNs are equivalent:

<NAMESPACE>:group:*parent-group*:role=manager#*group-authority1*

<NAMESPACE>:group:*parent-group*:role=manager#*group-authority2*

3.3 Resource capabilities

This section provides a specification for expressing resource-specific capabilities using entitlements (see [AARC-G027]). In the rest of this document, resource-specific capabilities will be referred to as just capabilities. A capability defines the resource or child-resource a user is allowed to access, optionally specifying certain actions the user is entitled to perform. Capabilities can be used to convey - in a compact form - authorisation information.

3.3.1 Syntax

Capabilities SHOULD be expressed according to the following syntax (components enclosed in square brackets are OPTIONAL, three dots ('...') indicate additional entries of the type after which they are placed, the backslash ('\') being the continuation character):

<NAMESPACE>:res:<RESOURCE>[:<CHILD-RESOURCE>]...[:act:<ACTION>[,<ACTION>]...]#<AUTHORITY>

where:

- <NAMESPACE>[3] is controlled by the e-infrastructure, research infrastructure or research collaboration that manages the capability. It is in the form of

  urn:<NID>:<DELEGATED-NAMESPACE>[:<SUBNAMESPACE>]...

  where

  - <NID> is the namespace identifier associated with a URN namespace registered with IANA[4], ensuring global uniqueness. Implementers SHOULD use one of the existing registered URN namespaces, such as urn:geant [URN-GEANT] or urn:mace [URN-MACE].
  - <DELEGATED-NAMESPACE> is a URN sub-namespace delegated from one of the IANA registered NIDs to an organisation representing the e-infrastructure, research infrastructure or research collaboration. It is RECOMMENDED that a publicly accessible URN value registry for each delegated namespace be provided.

  A <NAMESPACE> can have a variable number of elements. For example
  - **urn:geant:edugain**
  - **urn:geant:nikhef.nl**
  - **urn:geant:nikhef.nl:idm**

  are all valid **<NAMESPACE>** values.

[3] The <NAMESPACE> definition follows that in Section 3.2.1 (see also [AARC-G002]).
[4] Generic top level namespaces require IANA approval as per Section 6.2 of [RFC8141]

- The literal string **"res"** indicates that this is a *resource-specific* entitlement as opposed to, for example, an entitlement used for expressing group membership [AARC-G002].
- **<RESOURCE>** is the name of the resource. Whether the name should be unique is an implementation decision.
- An optional list of colon-separated **<CHILD-RESOURCE>** components represents a specific branch of the hierarchy of resources under the identified **<RESOURCE>.**
- An optional list of comma-separated **<ACTION>**s MAY be included, which, if present, MUST be prefixed with the literal string "**act**". This component MAY be used for further specifying the actions a user is entitled to do at a given resource. Note that the list of **<ACTION>**s is scoped to the rightmost child-resource; if no child-resource information is specified, actions apply to the top level resource. The interpretation of a capability without actions specified is an implementation detail.
- <AUTHORITY> is a mandatory and non-empty string that indicates the authoritative source of the capability. This SHOULD be used to further specify the exact issuing instance. For example, it MAY be the FQDN of the service that issued that specific capability. The <AUTHORITY> is specified in the f-component [RFC8141] of the URN; thus, it is introduced by the number sign ("#") character and terminated by the end of the URN. All characters must be encoded according to [RFC8141]. Hence, the <AUTHORITY> MUST NOT be considered when determining equivalence (Section 3 in [RFC8141]) of URN-formatted capabilities.

3.4 Affiliation information

This section describes how affiliation information should be expressed when transported across AARC BPA-compliant AAIs (see [AARC-G025]). Two different types of affiliation have been identified, namely Affiliation within the Home Organisation, such as a university, research institution or private company; and Affiliation within the Community, such as cross-organisation collaborations. Both affiliation types should be communicated to the service providers that rely on affiliation information in order to control access to resources. Note the use of the word "within," suggesting that the affiliation is not necessarily just membership but could also include the type of membership or role in the organisation.

3.4.1 Types of affiliation information

3.4.1.1 Affiliation within Home Organisation

Each user can be affiliated with one or more Home Organisations (such as, a university, research institution or private company) and the user's affiliations may change over time. The user's Home Organisation expresses affiliation information typically through the eduPersonScopedAffiliation attribute (ePSA) [EPSA] defined in the 200312 version of the eduPerson schema. ePSA is a multi-valued attribute that:

> "specifies the person's affiliation within a particular security domain in broad categories such as student, faculty, staff, alum, etc. The values consist of a left and right component separated by an "@" sign. The left component is one of the values from the eduPersonAffiliation controlled vocabulary [EPA]. The right-hand component of ePSA is the "scope" whose value MUST be the administrative domain to which the affiliation applies."

The affiliation within the user's Home Organisation is typically used by Service Providers for controlling access to resources, or for accounting purposes[5]. After receiving a scoped attribute from the IdP of the Home Organisation, SPs are expected to filter the attribute values by comparing the asserted scope to the scope value(s) in the IdP SAML metadata or

[5] Most services that authorise access based on institutional subscriptions will need to know whether a person is a "member" (as opposed to, say, "affiliate" or "library-walk-in", who would not be authorised to use the service through an institutional subscription.) For "members" (who are authorised), cases have been proposed where the service needed to know whether they were "students" specifically, in order to report the percentage of authorised users who were students.

to a locally defined list. Therefore, a BPA-compliant proxy SHOULD NOT release affiliation with Home Organisation information using ePSA because the SAML IdP metadata of the proxy typically does not include the scopes of the proxied Home Organisation IdPs. Instead, the proxy SHOULD ensure that the affiliation of the user within their Home Organisation (as released by the Home Organisation through the ePSA attribute) is conveyed to Service Providers via the voPersonExternalAffiliation (vPEA) attribute [VPEA]. The vPEA was defined in version 1.1.0 of the VO Person schema. The syntax and semantics of the vPEA attribute follows the ePSA described above. In particular, vPEA attributes values are scoped, but SPs SHOULD NOT verify the scope value against the list of acceptable scopes as asserted by the proxy in its SAML IdP metadata. As long as vPEA is not used for other purposes, the original authority of the asserted value can be gleaned from the scope of the value. An example flow of the attributes conveying the affiliation within the home organisation is illustrated in Figure 3.1.

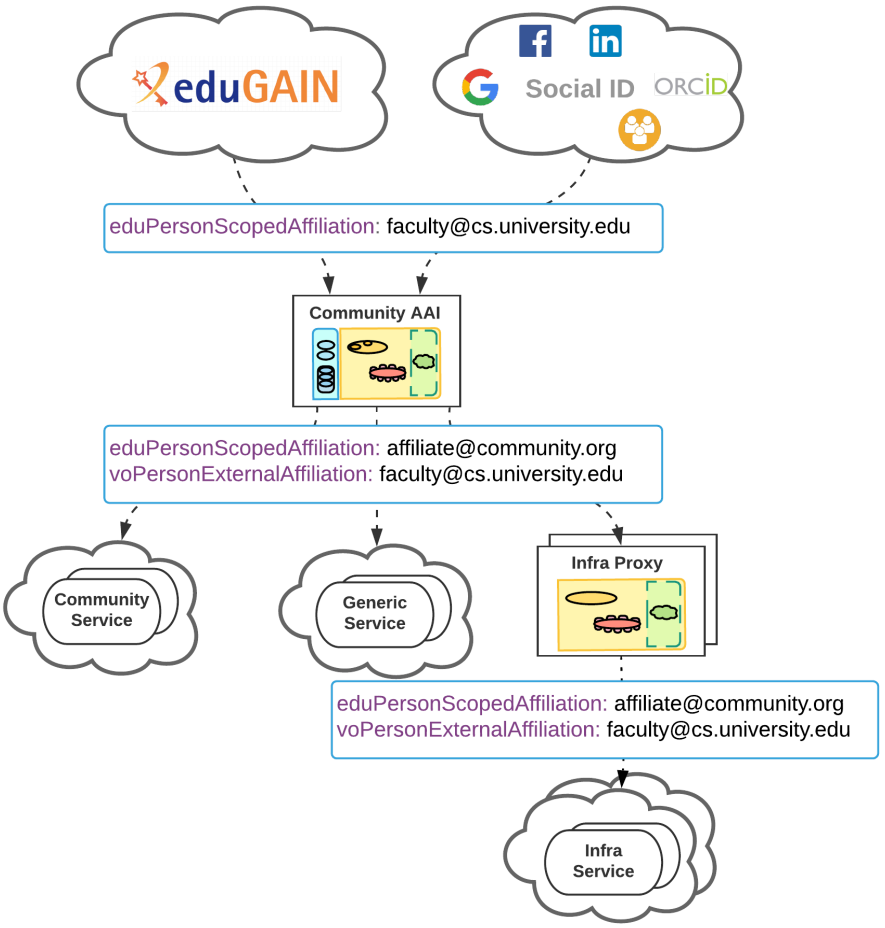

Figure 3.1: Flow of affiliation information across AARC-BPA compliant AAIs: Affiliation with Home Organisation is typically released to the BPA-compliant proxy of the community/research infrastructure/collaboration by the IdP of the Home Organisation through the eduPersonScopedAffiliation attribute. Services connected to the community SP-IdP-Proxy consume the Affiliation with Home Organisation information through the voPersonExternalAffiliation attribute. The Affiliation within Community is made available through the eduPersonScopedAffiliation attribute.

3.4.1.2 Affiliation within Community

Communities typically grant their members access to services and resources as expressed through each member's community identity (see Glossary). The SP-IdP-proxy that is serving the Community SHOULD release the affiliation within the Community using the eduPersonScopedAffiliation attribute [EPSA]. To allow the SPs behind the IdP proxy to consume the ePSA attribute values, the security domain(s) of the Community should be included as allowed scope values in the IdP

proxy metadata. An example flow of the attributes conveying the affiliation within the community is illustrated in Figure 3.1.

3.4.2 Representation of affiliation information

This section specifies how the types of affiliation information presented in Section 3.4.1 shall be represented using federated identity protocols.

Affiliation type	SAML attribute	OIDC claim	Example value
Affiliation within Community	eduPersonScopedAffiliation	eduperson_scoped_affiliation	`affiliate@community.org`
Affiliation within Home Organisation	voPersonExternalAffiliation	voperson_external_affiliation	`faculty@cs.university.edu`

Table 3.1. Example values for the different types of affiliation information

3.4.2.1 Security Assertion Markup Language 2.0 (SAML)

In SAML, affiliation information is represented as follows (see Table 3.1 for example values):

1. Affiliation within Home Organisation is represented using the multi-valued voPersonExternalAffiliation attribute, as defined in voPerson [VPEA].
2. Affiliation within Community is represented using the multi-valued eduPersonScopedAffiliation attribute, as defined in eduPerson [EPSA].

3.4.2.2 OpenID Connect (OIDC)

In OIDC, affiliation information is represented as follows (see Table 3.1 for example values)

1. Affiliation within Home Organisation is represented using the multi-valued voperson_external_affiliation claim, as defined in voPerson [VPEA], following the naming conventions specified in [OIDCRE].
2. Affiliation within Community is represented using the multi-valued eduperson_scoped_affiliation claim, as defined in eduPerson [EPSA], following the naming conventions specified in [OIDCRE].

3.4.3 Expression of affiliation information freshness

These guidelines have adopted the definition of freshness from version 1.0 of [RAF] which defines hierarchical values for expressing the "freshness" of affiliation information. "Freshness" here does not mean the actual freshness of the attribute, i.e. the time when the home organisation validated it, but rather the target *time window* within which the published value must change following a change in the user's affiliation. Specifically, when asserting `$RAF-PREFIX$/ATP/ePA-1d` for a given user, `$RAF-PREFIX$/ATP/ePA-1m` MUST also be asserted. Note that RAF is limited to the eduPersonAffiliation, eduPersonScopedAffiliation and eduPersonPrimaryAffiliation attributes defined in [EDUPERSON]. Additionally, the freshness of the attribute is further limited by the RAF specification to apply only to the following attribute values: "faculty", "student" and "member". Other values and attributes are out of scope of the RAF specification. Therefore, AARC has introduced additional values for expressing the freshness of affiliation information, which have no restriction on the values of the ePSA attribute. If the ePSA value is one of the values covered by the RAF specification (i.e. the left component of the value is one of "faculty", "student" or "member"), the affiliation freshness values MAY be expressed by asserting both the AARC and the RAF values. Note that the AARC values (listed in Table 3.2) are expressed as URIs which have the following prefix:

`$AARC-PREFIX$=https://aarc-community.org/assurance`

Value	Description
$AARC-PREFIX$/ATP/ePA-1m	eduPersonScopedAffiliation (SAML) / eduperson_scoped_affiliation (OIDC) (if populated and released to the RP) reflects user's departure from the Community within[6] 31 days time. If the value of affiliation is one of "faculty", "student" and "member" then $RAF-PREFIX$/ATP/ePA-1m MAY be asserted in addition to $AARC-PREFIX$/ATP/ePA-1m.
$AARC-PREFIX$/ATP/ePA-1d	eduPersonScopedAffiliation (SAML) / eduperson_scoped_affiliation (OIDC) (if populated and released to the RP) reflects user's departure from the Community within one day. If the value of affiliation is one of "faculty", "student" and "member" then $RAF-PREFIX$/ATP/ePA-1d MAY be asserted in addition to $AARC-PREFIX$/ATP/ePA-1d.
$AARC-PREFIX$/ATP/vPEA-1m	voPersonExternalAffiliation (SAML) / voperson_external_affiliation (OIDC) attributes (if populated and released to the RP) reflect user's departure from the Home Organisation within 31 days time. $AARC-PREFIX$/ATP/vPEA-1m SHOULD only be released if a) the Home Organisation released the eduPersonScopedAffiliation value within the same authentication session and b) the HO follows procedures in line with the $RAF-PREFIX$/ATP/ePA-1m policy, which is asserted by the HO to the proxy either via the release of the $RAF-PREFIX$/ATP/ePA-1m or by other means).
$AARC-PREFIX$/ATP/vPEA-1d	voPersonExternalAffiliation (SAML) / voperson_external_affiliation (OIDC) attributes (if populated and released to the RP) reflects user's departure from the Home Organisation within one day. $AARC-PREFIX$/ATP/vPEA-1d SHOULD only be released if a) the Home Organisation released the eduPersonScopedAffiliation value within the same authentication session and b) the HO follows procedures in line with the $RAF-PREFIX$/ATP/ePA-1d policy, which is asserted by the HO to the proxy either via the release of the $RAF-PREFIX$/ATP/ePA-1d or by other means).

Table 3.2. AARC values for expressing the freshness of affiliation information

Note that the term departure is used according to the definition from Section 2.3 in version 1.0 of [RAF].

[6] Since we follow RAF's definition of freshness, we have adopted the wording from the RAF specification. The use of the word "within" is ambiguous as it can suggest past or future, but the use here copies the usage in RAF. The intended meaning is that if at some point there is an event that leads to the user no longer being entitled to the attribute as originally published, this change is reflected by the attribute being changed accordingly, or removed, after at most 31 days following the event.

4 Authorisation

This section provides information for efficiently implementing access restrictions that are required by the individual communities and e-Infrastructures (see AARC-I047]). The provided information covers two topics: (a) classification of authorisation information and (b) models for community-based authorisation.

4.1 Classification of authorisation information

Authorisation information can be classified into two types:

1. User-attributes (often aggregated from different sources) such as:
    - Affiliation within the Home Organisation and/or the Community
    - Assurance, i.e. how well attribute assertions can be trusted
    - Group and role information (these primarily come from the Community)
2. Capabilities such as:
    - Information describing what actions a user is entitled to perform on a specific resource

It should be noted that, technically speaking, groups, roles and capabilities can all be expressed using the same attributes or claims (for example using the eduPersonEntitlement SAML attribute). The distinction between entitlements used for describing user-attributes as opposed to those that are used for describing capabilities will be clarified further below. The different authorisation models described in Section 4.2 rely on this distinction of authorisation information.

4.1.1 Expression of user-attributes

Expression of authorisation information for user-attribute-based information is described in the REFEDS Assurance Framework [RAF], [AARC-G021] and [AARC-G002] (see Section 3.2). For example, group membership information SHOULD be expressed as:

```
<NAMESPACE>:group:<GROUP>[:<SUBGROUP>]...[:role=<ROLE>]#<GROUP-AUTHORITY>
```

The following example describes membership of a top-level group, “parent-group”:

```
urn:example:example-ri.org:group:parent-group#auth-x.example-ri.org
```

The example below expresses a membership with a specific role[7], i.e. “manager”, in a group named “child-group” which is a subgroup of “parent-group”:

```
urn:example:example-ri.org:group:parent-group:child-group:role=manager#auth-x.example-ri.org
```

4.1.2 Expression of capabilities

Expression of capabilities follows the AARC Specification for expressing resource capabilities [AARG-G027] (see Section 3.3):

```
<NAMESPACE>:res:<RESOURCE>[:<CHILD-RESOURCE>]…[:act:<ACTION>[,<ACTION>]…]#<AUTHORITY>
```

[7] Note that the `role` component is scoped to the rightmost (child)group.

For example, the right to perform the actions create and delete on the storage resource identified as `vm_dashboard` could be issued by the example-ri.org as follows:

```
urn:example:example-ri.org:res:vm_dashboard:storage:act:create,delete#auth-x.example-ri.org
```

4.2 Authorisation models

Authorisation models describe the organisational flow of authorisation information. Any other information needed by the service to fulfil actions such as personalisation, accounting, traceability, is out of the scope of this document. The organisational flow of authorisation information follows this lifecycle:

- Definition of authorisation information at one or more Attribute Authorities (AA)
- Aggregation of authorisation information
- Use of authorisation information for making an authorisation decision
- Enforcement of the authorisation decision

The information provided in this section is based on the analysis of the authorisation architectures from nine different use cases detailed in [AARC2-DJRA1.2]. Based on this analysis, we have identified three main authorisation models[8] that make use of an SP-IdP-Proxy, as illustrated in Figure 4.1.

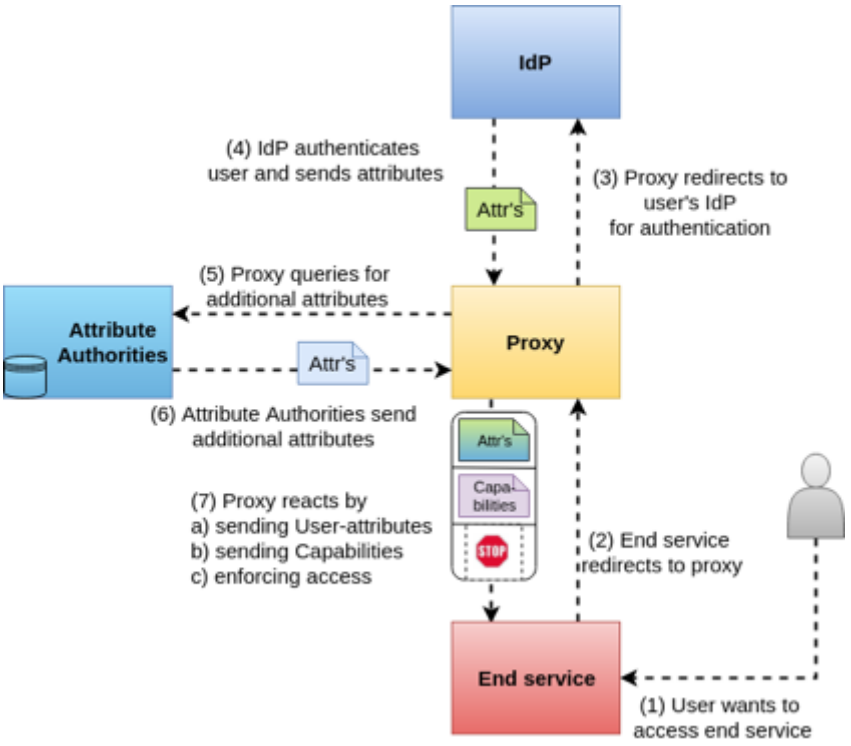

Figure 4.1: Flow of authorisation information for a user who wants to access an end service in a BPA-compliant infrastructure. There are three different alternative paths in Step 7, where the proxy either (a) sends user-attributes, such as group or role information, to the end service, which can then make a decision based on that information; or (b) takes

[8] Please note that this document does not make use of the "P*P" terminology adopted in [AARC2-DJRA1.2].

the authorisation decision and re-expresses it as, for example, capabilities; or (c) takes the authorisation decision and enforces it by denying access to the end service.

An outline of the authorisation models is provided below:

1. Centralised Policy Information Point: the proxy aggregates user-attributes and makes them available to the end-services (Step 7a in Figure 4.1)
2. Centralised Policy Management and Decision Making: the proxy conveys the authorisation decision to the end-services (Step 7b in Figure 4.1)
3. Centralised Policy Management and Decision Making and Enforcement: the proxy enforces the decision directly at the proxy (Step 7c in Figure 4.1)

It should be mentioned that user-attributes and/or capabilities in the models above can be communicated to end services following either a "push" or "pull" approach [AARC-G006]. For example, in the SAML 2.0 Web SSO flow, attributes are pushed to end-services, whereas in the OpenID Connect authorisation code flow, the end-services can either query the `UserInfo` endpoint or rely on the pushed `id_token`. These three models are described in more detail in the following sections.

4.2.1 Centralised Policy Information Point

In this model, the proxy aggregates the information and makes it available to the end services so they can make the authorisation decision. This allows the service to perform fine-grained access control, because all information necessary for an informed decision is available. However, scalability may become an issue for large deployments. For example, it may become non-trivial to consistently update authorisation across a large number of services, as the authorisation policy needs to be replicated to every service. Additionally, services may see user-specific authorisation data, such as group membership, that might be intended for other services. This may be problematic with regard to the "data minimisation principle". Furthermore, this puts the onus on the services to correctly interpret and act on the obtained authorisation information.

4.2.2 Centralised Policy Management and Decision Making

In this model the proxy makes the authorisation decision and encodes this decision into resource-specific authorisation information, typically in the form of capabilities. This allows the decision at the proxy to be based on additional information which the proxy might prefer not to send to the services. This is generally simpler for the end services to implement, since the complexity of interpretation of the authorisation information is handled by the proxy. In contrast to the approach described in Section 4.2.1, this puts the onus on the proxy to correctly interpret and act on the authorisation information. Note that in this model:

1. the proxy is creating and/or translating authorisation statements
2. the proxy may need to make a mix of capabilities and user attributes available for the service to be able to properly enforce the authorisation decision.

4.2.3 Centralised Policy Management, Decision Making and Enforcement

In this model, the proxy makes the authorisation decision, as in the case of Centralised Policy Management and Decision Making. Furthermore, the proxy is responsible for enforcing that decision. This allows the integration of services that might not be capable of doing any authorisation, with only little modification. However, it requires the proxy to understand the authorisation policy of the end services. Often this type of authorisation enforcement is only used for certain parts (e.g. a global black- or whitelist) while using the other models for the rest of the authorisation. For example, in case the proxy grants the user access to the end service, this model may be followed by either of the models described in sections 4.2.1 and 4.2.3.

4.3 Summary of recommendations

This section provides a summary of the recommendations in this chapter.

- The release of attributes to SPs SHOULD follow the data minimisation principle" (GDPR Article 5 (1.e) [GDPR, GDPR-INFO]. This might influence the choice between attribute- and capability-based access-control.
- To support traceability, as required by [SIRTFI, SNCTFI], implementations SHOULD at least do one, but preferably both, of the following:
    - maintain and send a pseudonymous unique identifier for the user from proxies down to the services.
    - maintain and send a unique ID that identifies the job or associated session.
- Group and subgroup membership and roles (where applicable):
    - SHOULD be expressed using [AARC-G002]
    - Subgroups (where applicable) SHOULD be used for expressing finer grained access permissions.
    - Roles (where applicable) SHOULD be used to specify additional rights inside the corresponding (sub)group (see [AARC2-DJRA1.3])
- Resource Specific Capabilities (where applicable) SHOULD be expressed according to [AARC-G027].
- Assurance information (where applicable) SHOULD be expressed:
    - Using the REFEDS Assurance Framework [RAF] and [AARC-G021] which extends [RAF] with additional assurance profiles recommended to be used between infrastructures.
    - In conjunction with specifications focusing on authentication, such as the REFEDS Single Factor Authentication (SFA) [REFEDS-SFA] and the REFEDS Multi-Factor Authentication (MFA) [REFEDS-MFA] profile.
- Affiliation information (where applicable) SHOULD be expressed according to [AARC-G025].

4.3.1 Considerations on the different models

1. Authorisation implementations SHOULD support the Centralised Policy Information Point model for end services that require full control over the authorisation process. Authorisation implementations MUST be aware that in this model it is easy to send more data than required to end service. Filtering MAY be a solution.
2. Authorisation implementations SHOULD support the Centralised Policy Management and Decision Making model for simplifying the authorisation process for the end services. Authorisation implementations MUST be aware that the onus for correctly interpreting and acting upon authorisation information is put on the proxy.
3. Authorisation implementations SHOULD only use the Centralised Policy Management, Decision Making and Enforcement model for a partial authorisation decision (e.g. central suspension), and combine it with one of the two models above.
4. Depending on the requirements of the Service Providers reached through the proxy, it is possible to use a hybrid approach, combining any of the three models above, in a single authorisation flow. In all these flows the proxy can supplement the attributes from the authenticating IdP with information from AAs. The three different approaches address whether and how this information is passed on to the end services.

5 Evaluation and combination of the assurance of external identities

The AARC BPA allows users to authenticate through more than one external identity provider (external to and independent of the infrastructure), be they home organisation, social media, community managed virtual organizations, etc. Each identity provider provides different personal identity attributes (name, email), affiliation (organisational affiliation, community membership) and assurance information which the proxy combines together to create the community[9] identity (see definition in the Glossary). When multiple external identities are linked to the community identity, the user has different authentication options. In this context, we will refer to the identity used to authenticate as the *effective identity*.

Infrastructures also define one or more assurance profiles [AARC-G021], or a combination of assurance components, tailored to a specific risk assessment. To assign an assurance profile (or a set of assurance components values) to their users, the Infrastructure needs to evaluate the assurance of the linked identity, or identities, used to register with the Infrastructure's AAI. The remainder of this chapter provides guidelines for combining assurance information and for assessing assurance component values in the absence of assurance information from the external identity provider (see [AARC-G031]).

5.1 Combined assurance evaluation

The guidelines provided in this section adopt the definition of Assurance as specified in the REFEDS Assurance Framework [RAF]. Along the lines of other recent assurance guidelines [NIST.SP.800-63-3] and proposed standards [RFC8485], the RAF does not use the concept of level(s) of assurance, rather it splits assurance into separate components. The RAF considers the following three components:

- Identity uniqueness
- Identity proofing and credential issuance, renewal and replacement
- Attribute quality and freshness

The assurance values are represented using the eduPersonAssurance attribute [EDUPERSON] in case of SAML 2.0, or using the eduperson_assurance claim as defined by the REFEDS OIDCre working group [REFEDS-OIDCre] in case of OIDC.

A requirement for the assurance evaluation is that assurance components related to the same individual, but coming from different IdPs, are defined along the lines of the RAF, or, when expressed through other assurance frameworks such as the eIDAS assurance levels [EIDAS-LOA], can be translated into those definitions. When no assurance information is directly provided by the IdP during the authentication, the Infrastructure SHOULD NOT make any assumption on the assurance of the external identities, but it can rely on other evidence and compensatory controls to assess the relevant assurance features of the incoming identity, as it will be shown in the following sections on a component by component base.

The components SHOULD eventually be collapsed to compose assurance profiles, each consisting of a set of values for one or more of these components. Please refer to [AARC-G021] for the available assurance profiles.

[9] Note that the term "infrastructure identity" used in [AARC-G031] has been replaced with "community identity" in this document. This change in the terminology is a result of the community-first approach (see Section 2.2).

5.1.1 Identifier uniqueness (ID)

The RAF ID component describes "how a CSP (see Glossary) expresses that an identifier represents a single natural person and if that person remains the same over time" [RAF]. When an external identity provider asserts the ID component value `unique`, no further evaluation is to be made by the Infrastructure, and the value SHOULD be treated verbatim. The evaluation SHOULD be performed at the time of the identity linking.

5.1.1.1 Combined evaluation

When combining ID component values that belong to two or more linked identities, the value for the Infrastructure identity SHOULD be calculated with an AND operation where a value `unique` is equal to TRUE and a value N/A (not available value) is equal to FALSE. As in:

```
ID_value = ID_value_1 AND ... ID_value_n
```

Possible combinations and values with two linked identities are listed in Table 5.1.

Linked Identity 1 ID value	Linked Identity 2 ID value	Infrastructure Identity ID value
`unique`	`N/A`	`N/A`
`N/A`	`unique`	`N/A`
`unique`	`unique`	`unique`

Table 5.1. ID component combinations with two linked identities

Effectively, the value unique for the Infrastructure Identity cannot be asserted when any of the linked identities lacks it. This is required to prevent linking the Infrastructure Identity with shared or reassignable accounts.

5.1.1.2 Compensatory controls

When an external identity provider does not assert the ID component value unique the Infrastructure SHOULD perform compensatory controls as defined by Expression of REFEDS RAF assurance components for identities derived from social media accounts [AARC-G041]. Failure to do so will expose the Infrastructure to unreasonable risks (for example non-traceability of users or the use of shared accounts). The compensatory controls required to assert the ID component value `unique` are listed in Table 5.2.

External identity provider	Compensatory controls (short name[10])
Any IdP (including social media IdPs)	`R&S_EC || (im_a_person && contacts)`

Table 5.2. Compensatory controls to assert the ID component value unique

5.1.2 Identity proofing and credential issuance, renewal and replacement (IAP)

The RAF IAP component describes the quality of the identity proofing, credential issuance, renewal and replacement processes. The possible values are:

- `https://refeds.org/assurance/IAP/low`

[10] See Section 5.2 for technical details about the compensatory controls.

- `https://refeds.org/assurance/IAP/medium`
- `https://refeds.org/assurance/IAP/high`

The IAP component value MUST be asserted incrementally, that is: when asserting a value medium, the value low MUST be asserted too; when asserting a value high, the values medium and low MUST be asserted too [RAF]. When an external identity provider asserts the IAP component value, no further evaluation is needed.

5.1.2.1 Combined evaluation

When combining IAP component values that belong to two or more linked identities, the value for the Infrastructure identity will be equivalent to the value of the effective identity.

5.1.2.2 Compensatory controls

When an external identity provider does not assert any IAP component values, the Infrastructure SHOULD perform compensatory controls defined in [AARC-G041]. The controls listed in Table 5.3 can be used to raise the assurance of the IAP component from no value to `low`.

External identity provider	Compensatory controls (short name)
Any IdP (including social media IdPs)	`conf_email`

Table 5.3. IAP component compensatory controls to assert the value low

The controls included in Table 5.3 allow for asserting the value `low` for the IAP components without the need to manage policies per IdP and/or Identity Federation. However, it may well be the case that an eduGAIN IdP would qualify for higher IAP values. To assert IAP component values above `low`, the Infrastructure SHOULD consider evaluating both the Identity Federation policy and the assurance information published in the metadata of the incoming IdP. All the policies of the Identity Federations that belong to eduGAIN are published on the eduGAIN Technical site [eduGAIN-TECH].

5.1.3 Attribute quality and freshness (ATP)

The ATP component describes the quality and the freshness of the attributes the IdP delivers to the SP (in this case the SP side of the IdP/SP proxy of the Infrastructure). Current values are limited to represent the freshness of the affiliation attributes defined in [eduPerson]: eduPersonAffiliation, eduPersonScopedAffiliation and eduPersonPrimaryAffiliation. The permitted values are:

- `https://refeds.org/assurance/ATP/ePA-1m`
- `https://refeds.org/assurance/ATP/ePA-1d`

The values reflect the latency in updating the affiliation status of a user in case of departure or role change. The values are hierarchical, that is when asserting ePA-1d then ePA-1m MUST also be asserted.

The Infrastructure MAY rely on the ATP component value expressed by an external identity provider to compute its own [AARC-G041].

5.2 Compensatory controls

The list of compensatory controls proposed in this section is not meant to be exhaustive. The Infrastructure can define additional control if deemed necessary and according to its own policy and risk assessment.

5.2.1 I'm a person

When users register with the Infrastructure, they will be required to confirm that they are a single natural person and that they will not share the account with other people. Those requirements MAY also be included in the Infrastructure AUP.

Rationale	Ensure that the user is a single natural person, and have a simple way to ban users that share their account for policy/AUP violation.
RAF requirement	The "I'm a person" statement is meant to meet one of the four requirements for asserting the value unique of the ID component: the "User account belongs to a single natural person" [RAF].
Enforcement	The "I'm a person" statement itself cannot prevent bad actors and misbehaviour, but it gives a solid ground for banning or suspending malevolent or careless users. Failure to confirm the statement will prevent the user to access the Infrastructure.
Short name	`im_a_person`

5.2.2 Contacts

When users register with the Infrastructure, their (external) identity providers will be required to release contact information such as email or mobile phone number. The "Confirmation mail" compensatory control can substitute "Contacts", but not vice versa.

Rationale	Have a mean to contact the user.
RAF requirement	The "Contacts" control is meant to meet one of the four requirements for asserting the value unique of the ID component: the "CSP can contact the person to whom the account is issued" [RAF].
Enforcement	The failure to release contact information by the external IdP can have two different outcomes: the users cannot access the Infrastructure or they will be asked to supply the missing information.
Short name	`contacts`

5.2.3 Research and Scholarship entity category

eduGAIN IdPs asserting the support for the REFEDS Research and Scholarship entity category [REFEDS-R&S] commit to release a set of attributes following specific rules on the quality of the identifier.

Rationale	Reuse the entity category rules about the identifier.
RAF requirement	Support for REFEDS R&S meets all the requirements of the value `unique` of the ID component.
Enforcement	Failure to detect support for the entity category in the IdP metadata should activate the other compensatory controls.
Short name	`R&S_EC`

5.2.4 Confirmation email

When users register with a service, it is common practice to send an email to the provided address with a confirmation link. Once received, the user will follow the link to complete the registration process. This process guarantees that the email is both valid and in control of the user. The Infrastructure will embrace the same process for the users' registration.

Rationale	Obtain a verified email address for each user registering to the Infrastructure.
RAF requirement	The confirmation email is the basic requirement for the value `low` of the IAP component.
Enforcement	Failure to provide a valid email address, or to follow the link sent via the confirmation email, will prevent the user from accessing the Infrastructure.
Short name	`conf_email`

5.3 Authentication assurance

The RAF does not cover the assurance quality of the authentication process. However, the REFEDS Assurance Working Group [REFEDS-AWG] has defined two authentication assurance profiles that MAY be paired with the RAF assurance components:

- REFEDS Multi-Factor Authentication (MFA) Profile [REFEDS-MFA]
- REFEDS Single-Factor Authentication (SFA) Profile [REFEDS-SFA]

Whether or not the Infrastructure will evaluate the authentication assurance expressed by an external identity provider, the authentication assurance values cannot be combined. The authentication value of the Infrastructure identity MUST always be equal to the one associated with the effective identity.

6 IdP discovery for SPs in multi-BPA environments

Authentication to a service in a multi-IdP environment requires that the service redirect the incoming user to their home identity provider (IdP). Currently this is often accomplished by discovery services (often also called “where are you from” or WAYF services), where the user chooses their home IdP. The AARC BPA introduces new IdPs, i.e. IdP-SP-Proxies, that may be chosen by the service instead of sending the user directly to a home IdP. Often, users have to choose between a list of IdP-SP-Proxies. This makes it increasingly difficult for users to understand which IdP is the best choice for authentication.

In this section we focus on enabling Service Providers / OIDC-Relying-Parties / WAYF Services to obtain a hint about the IdP to which the user should be sent for authentication (see [AARC-G049]). We define a portable and technology-agnostic way to allow services to receive hints about which IdP to use. This mechanism can greatly simplify the discovery process for the end-user, by either narrowing down the number of possible IdPs to choose from or by making the actual selection process fully transparent. Furthermore, the described concept includes the possibility of chaining, so that hints can be nested. This allows creating URLs that point to an SP, with a hint trail that leads via an IdP-SP-Proxy to a given home IdP.

Finally, we want to stress that this hinting process takes place before any authentication has happened. The flow of information is therefore independent of the underlying protocol used. The hints themselves, however, may contain protocol specific information. We also stress that it is only a hint. Whether the proxy or service actually honours the hint depends on the list of locally configured trusted IdPs.

6.1 Context

The IdP hinting mechanism described in this document is based on the following assumptions:

- **Web**: We focus on web, but do not a priori exclude non-web scenarios.
- **Context**: IdPs may be home-IdPs or IdP-SP-Proxies.
- **Trust**: Services trust IdPs based on a trust relation that is out of scope of this document. Therefore, we use the term “hinting”, to emphasize that it is certainly possible for the SP or proxy to decide not to follow a hint.
- **AARC Blueprint Architectures (BPA):** Our definition supports, but by no means requires, that services are operated in a BPA context. I.e. in addition to the previous point, end services in a BPA context would only accept hints towards supported proxies.
- The service that obtains a hint can either process it itself, or decide to pass the hint to its WAYF for filtering the list of potential IdPs. Details for this are out of scope of this document.
- Services that want to send users to a specific home-IdP for reauthentication, will need to keep track of the necessary identifier to do so.

Multiple technologies can benefit from IdP hinting. IdP hinting should at least work for SAML2 and OAuth2/OIDC based services.

6.2 Specification

- The identifier of the hinted IdP MUST be passed through the “idphint” GET parameter.
- POST parameters MAY be supported in addition to the GET parameter.
- The service MUST interpret the parameter “idphint” of a request as the URL-encoded identifier of the IdP to which the creator of the url intends to send the user for authentication.
- Implementations MUST also encode slashes (‘/’).
- The hinted identifiers MUST be well-defined URIs [RFC3986]:
    - For SAML it MUST be the EntityID

    - For OAuth2.0 and OIDC it MUST be the issuer
- Multiple IdPs MAY be provided, which MUST be encoded as a comma separated list of URL-encoded identifiers.
- Case sensitivity MUST follow the underlying specification of the URL-decoded identifier.

Conclusions

The AARC Blueprint Architecture provides a reference architecture for implementing an AAI that supports common use cases within research collaborations. The community-first approach focuses on interoperability across BPA-compliant AAIs and provides a broader view for addressing an increasing number of use cases from research communities requiring access to federated resources offered by different infrastructure providers. It should be stressed that the community-first approach retains compatibility with previous versions of the BPA which have already been adopted by many e-infrastructure providers, research infrastructures and collaborations.

The current iteration of the BPA is accompanied by a set of guidelines and informational documents. Specifically, there are documents that provide guidance on the interoperable expression of information that includes community user identifiers, group membership and role information, resource-specific capabilities and affiliation information.

Furthermore, the updated BPA includes authorisation models that allow implementers to delegate many of the complex authorisation decisions to central components. These models support services with different requirements with respect to their level of involvement in the authorisation process.

There are also guidelines for combining the assurance information which is associated with the external identities linked to the community identity. The provided guidelines include compensatory controls for assessing assurance component values in the absence of assurance information from the external identity provider.

A portable and technology-agnostic mechanism has been specified to allow services to receive "hints" about which identity provider to use. This can greatly simplify the discovery process for the end-user, particularly in the presence of one or more SP-IdP-Proxies.

Work on the BPA will continue beyond AARC2, focusing on:

- guidelines for integrating OIDC- and OAuth2-based services, covering the following topics:
    - scalable & trusted registration mechanisms based on the OpenID Connection Federation specification
    - standardised OIDC profiles and claims for Research and Education (see OpenID Research & Education working group [OIDC-RANDE])
    - validation of OAuth2 tokens in multi-proxy/Authorisation server environments
- streamlining the process for sharing community services with other communities, following the community-first approach
- best practises for (de)provisioning of user account information across BPA-compliant AAIs

References

[AARC-BPA-2016]	AARC Blueprint Architecture 2016 (AARC-G011); https://aarc-project.eu/guidelines/aarc-g011/
[AARC-BPA-2017]	AARC Blueprint Architecture 2017 (AARC-G012); https://aarc-project.eu/guidelines/aarc-g012/
[AARC-G002]	AARC guidelines: Expressing group membership and role information (AARC-G002); https://aarc-project.eu/guidelines/aarc-g002
[AARC-G006]	AARC informational document: Best Practices for managing authorisation (AARC-G006); https://aarc-project.eu/guidelines/aarc-g006
[AARC-G016]	AARC guidelines: Recommendations on the exchange of personal data in accounting data sharing (AARC-G016); https://aarc-project.eu/guidelines/aarc-g016
[AARC-G021]	AARC guidelines: Exchange of specific assurance information between Infrastructures (AARC-G021); https://aarc-project.eu/guidelines/aarc-g021
[AARC-G025]	AARC guidelines: Expressing affiliation information *(to appear)*; https://aarc-project.eu/guidelines/aarc-g025
[AARC-G026]	AARC guidelines: Expressing community user identifiers *(to appear)*; https://aarc-project.eu/guidelines/aarc-g026
[AARC-G027]	AARC guidelines: Specification for expressing resource capabilities; https://aarc-project.eu/guidelines/aarc-g027
[AARC-G031]	AARC guidelines: Evaluation and combination of the assurance of external identities; https://aarc-project.eu/guidelines/aarc-g031
[AARC-G041]	AARC guidelines: Expression of REFEDS RAF assurance components for identities derived from social media accounts; https://aarc-project.eu/guidelines/aarc-g041
[AARC-G042]	AARC guidelines: Data Protection Impact Assessment – an initial guide for communities; https://aarc-project.eu/guidelines/aarc-g042
[AARC-G049]	AARC guidelines: A specification for IdP hinting; https://aarc-project.eu/guidelines/aarc-g049

[AARC-I047]	AARC informational document: Implementing scalable and consistent authorisation across multi-SP environments; https://aarc-project.eu/guidelines/aarc-i047
[AARC2-DJRA1.2]	AARC2 Deliverable: Scalable, integrated authorisation models for SPs; https://aarc-project.eu/wp-content/uploads/2018/07/AARC2-DJRA1.2_V4-FINAL.pdf
[AARC2-DJRA1.3]	AARC2 Deliverable: VO Platforms for Research Collaboration; https://aarc-project.eu/wp-content/uploads/2018/10/AARC2-DJRA1.3-v2.pdf
[CORMACK1]	A. Cormack, "Federated Access Management and GDPR"; https://community.jisc.ac.uk/blogs/regulatory-developments/article/federatedaccess-management-and-gdpr
[CORMACK2]	A. Cormack, "Legitimate Interests and Federated Access Management"; https://community.jisc.ac.uk/blogs/regulatory-developments/article/legitimate-interests-and-federated-access-management
[EDUPERSON]	eduPerson Object Class Specification (201602); http://software.internet2.edu/eduperson/internet2-mace-dir-eduperson-201602.html
[EIDAS-LOA]	European Commission, "Commission Implementing Regulation (EU) 2015/1502 of 8 September 2015 on setting out minimum technical specifications and procedures for assurance levels for electronic identification means"; http://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=OJ:JOL_2015_235_R_0002
[EPA]	eduPersonAffiliation attribute definition (eduPerson 1.0); http://software.internet2.edu/eduperson/internet2-mace-dir-eduperson-201602.html#eduPersonAffiliation
[EPE]	eduPersonEntitlement attribute definition (eduPerson 200210); http://software.internet2.edu/eduperson/internet2-mace-dir-eduperson-201602.html#eduPersonEntitlement
[EPSA]	eduPersonScopedAffiliation attribute definition (eduPerson 200312); http://software.internet2.edu/eduperson/internet2-mace-dir-eduperson-201602.html#eduPersonScopedAffiliation
[GDPR]	General Data Protection Regulation on eur-lex; https://data.europa.eu/eli/reg/2016/679/2016-05-04
[GDPR-INFO]	Informational website for the General Data Protection Regulation; https://gdpr-info.eu/
[NIST.SP.800-63-3]	NIST Special Publication 800-63-3, Digital Identity Guidelines, June 2017; https://doi.org/10.6028/NIST.SP.800-63-3
[OIDC-CORE-v1.0]	N. Sakimura, J. Bradley, M. Jones, B. de Medeiros, B., and C. Mortimore, "OpenID Connect Core 1.0", November 2014; http://openid.net/specs/openid-connect-core-1_0.html
[RAF]	REFEDS Assurance Framework; https://refeds.org/assurance
[REFEDS-AWG]	REFEDS Assurance Working Group; https://wiki.refeds.org/display/GROUPS/Assurance+Working+Group
[REFEDS-MFA]	REFEDS Multiple Factor Authentication Profile; https://refeds.org/profile/mfa
[REFEDS-OIDCre]	REFEDS OpenID Connect for Research and Education Working Group; https://wiki.refeds.org/display/GROUPS/OIDCre
[REFEDS-R&S]	REFEDS Research and Scholarship Entity Category; https://refeds.org/category/research-and-scholarship
[REFEDS-SFA]	REFEDS Single Factor Authentication Profile; https://refeds.org/profile/sfa
[RFC3986]	T. Berners-Lee, R. Fielding, L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", RFC 3986, DOI:10.17487/RFC3986, January 2005; https://www.rfc-editor.org/info/rfc3986
[RFC8141]	P. Saint-Andre, J. Klensin, "Uniform Resource Names (URNs)", RFC 8141, DOI:10.17487/RFC8141, April 2017; https://tools.ietf.org/html/rfc8141
[RFC8485]	J. Richer, L. Johansson, "Vectors of Trust", RFC 8485, DOI 10.17487/RFC8485, October 2018; https://www.rfc-editor.org/info/rfc8485
[SAML-SubjectID-v1.0]	SAML V2.0 Subject Identifier Attributes Profile Version 1.0. Edited by Scott Cantor. 16 January 2019. OASIS Committee Specification 01; https://docs.oasis-open.org/security/saml-subject-id-attr/v1.0/cs01/saml-subject-id-attr-v1.0-cs01.html
[SIRTFI]	Security Incident Response Trust Framework for Federated Identity; https://refeds.org/sirtfi
[SNCTFI]	Scalable Negotiator for a Community Trust Framework in Federated Infrastructures; https://www.igtf.net/snctfi
[VPEA]	voPersonExternalAffiliation attribute definition (v1.1.0); https://github.com/voperson/voperson/blob/1.1.0/voPerson.md#vopersonexternalaffilation-attribute-definition
[WISE-SCI]	Security for Collaborating Infrastructures (SCI) Trust Framework; https://wise-community.org/sci/
[X.1254]	International Telecommunication Union. Series X. Data Networks, Open System Communication and Security. Cyberspace security – Identity management. Entity authentication assurance framework. Standard X.1254; https://www.itu.int/rec/T-REC-X.1254


AARC

Glossary

AAI	Authentication and Authorisation Infrastructure
AAI service	A service that enables authenticated and authorised access to resources
AUP	Acceptable Use Policy
CA	Certification Authority
Community	A group of users, organised with a common purpose, and jointly granted access to resources. It may act as the interface between individual users and the resources. (see also [WISE-SCI])
Community AAI	An AAI service that also enables the use and management of community identities for access to resources. It comprises three (3) AARC BPA component layers: the Access Protocol Translation, the Community User Attribute Services, and the Authorisation.
Community identity	A user's digital identity that may be enriched by the community with additional attributes such as a shared user identifier, profile information, and community attributes such as group membership and role information (see [REFEDS-R&S] and [SIRTFI]).
Community service	A service provided only to members of a specific community.
Credential	A set of data presented as evidence of a claimed identity and/or entitlements [X.1254]
Credential Service Provider	A trusted actor that issues and/or manages credentials [X.1254]. In the context of the [RAF] specification, Credential Service Provider refers to the Identity Provider and the associated Identity Management system that manages the user identities and attributes observed by the Relying Parties.
CSP	Credential Service Provider
Digital identity	Information that represents an entity (subject) within a domain. It contains information about the subject's attributes and relationships
eduGAIN	International interfederation service interconnecting research and education
Generic service	A service provided to members of different communities
GDPR	General Data Protection Regulation [GDPR]
IdP	Identity Provider
Infrastructure proxy	An AAI service of a research infrastructure or e-Infrastructure (hereafter termed infrastructure) that enables access to resources offered by Service Providers connected to that infrastructure. This AAI service does not provide community membership management. Specifically, the infrastructure proxy comprises two (2) AARC BPA component layers: the Access Protocol Translation and the Authorisation.
Infrastructure service	A service provided by a research infrastructure or e-Infrastructure to members of one or more Community AAI which receives the required attributes through an Infrastructure Proxy
OIDC	OpenID Connect
Relying Party	Actor that relies on an identity assertion or claim [X.1254]
RP	Relying Party
SP	Service Provider
VO	Virtual Organisation


AARC