Authentication and Authorisation for Research and Collaboration

AAI Overview

Training from AARC

Section 1 - AAI Overview

- Authentication and authorization processes
- Externalizing authentication (LDAP)
- Federated authentication (SAML and OIDC)
- Federations in R&D
- The inter-federation: eduGAIN

Access to applications

When you manage an application, you must guarantee it is **accessed only by people that have the right** to use it.

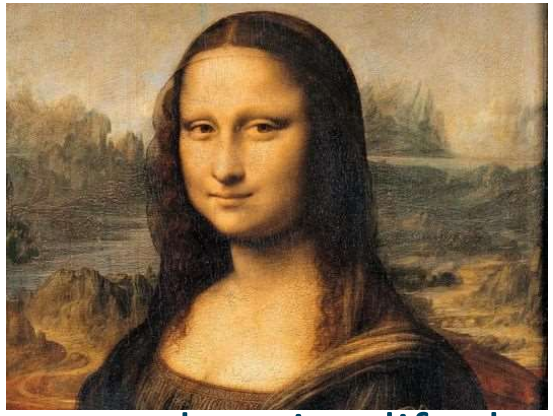
To achieve successfully this goal, two distinct processes need to be implemented:

- **authentication** of a user
- enforce of **authorization** rules for the user

Authentication (AuthN)

- Authentication is the **act of confirming the truth** of an attribute of a single piece of data or entity (the user of an application, for instance).

- Example (in the real world): authenticating the Mona Lisa.



- In the digital world we tend to simplify the confirmation by using **username and password** *(the assumption is that password is known only by the intended user, so specifying the right password you're demonstrating you actually are who you pretend to be)*.

Authorization (AuthZ)

- Authorization is the function of **specifying access rights** to resources related to information security and computer security in general and to access control in particular.
  - More formally, ”to authorize” is to define an access policy.

- Example: going to a concert.

Authorization (AuthZ)

- In the digital world, defining the access rules user by user can be impractical.
- Authorization is often implemented with the so called **Role-Based Access Control** (RBAC)
  - **users are pooled** in groups based on their organizational role (e.g. payroll manager, project group A, …)
  - access **rights** are then **associated to roles**
- When a user access an application:
  - authenticates himself/herself
  - activate one or more roles (depending on the groups of belonging)
  - access application/services by leveraging RBAC authorization

Managing AuthN and AuthZ

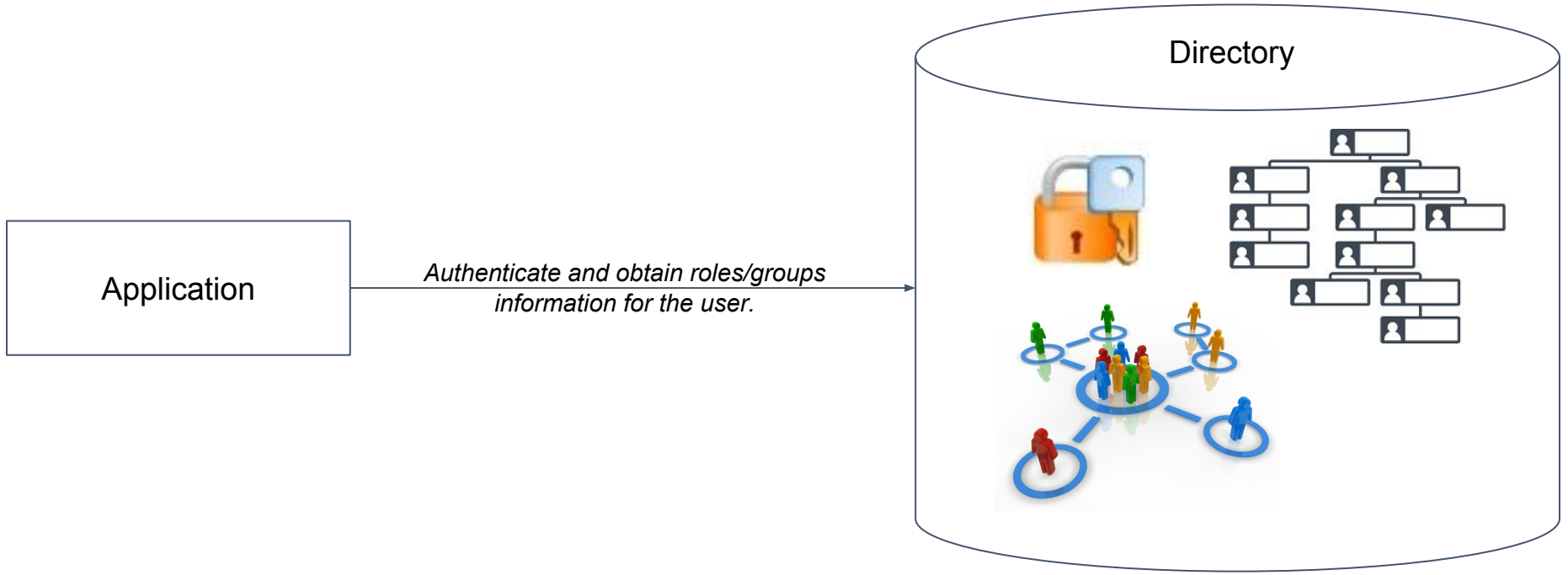
As we have seen, an application to deal with authentication and authorization has to manage the following information:

1. **usernames** and associated **passwords**, *to authenticate users and verify they are who they pretend to be;*
2. institutional **roles**, *to describe the roles within the group or organization (used for RBAC);*
3. user **groups**, *to pool together users that have the same role in the organization (groups are associated to roles);*
4. **access policies**, *rules in the form of (role name -> access right) to describe which operation each role is entitled to perform and which not inside the application.*

Externalizing authentication

For simplicity, and not to duplicate information, usually a **Directory** is used to collect username, password, roles and groups for the whole organization.

Directory services play an important role by allowing the sharing of information about users, systems, networks, services, and applications throughout the network.

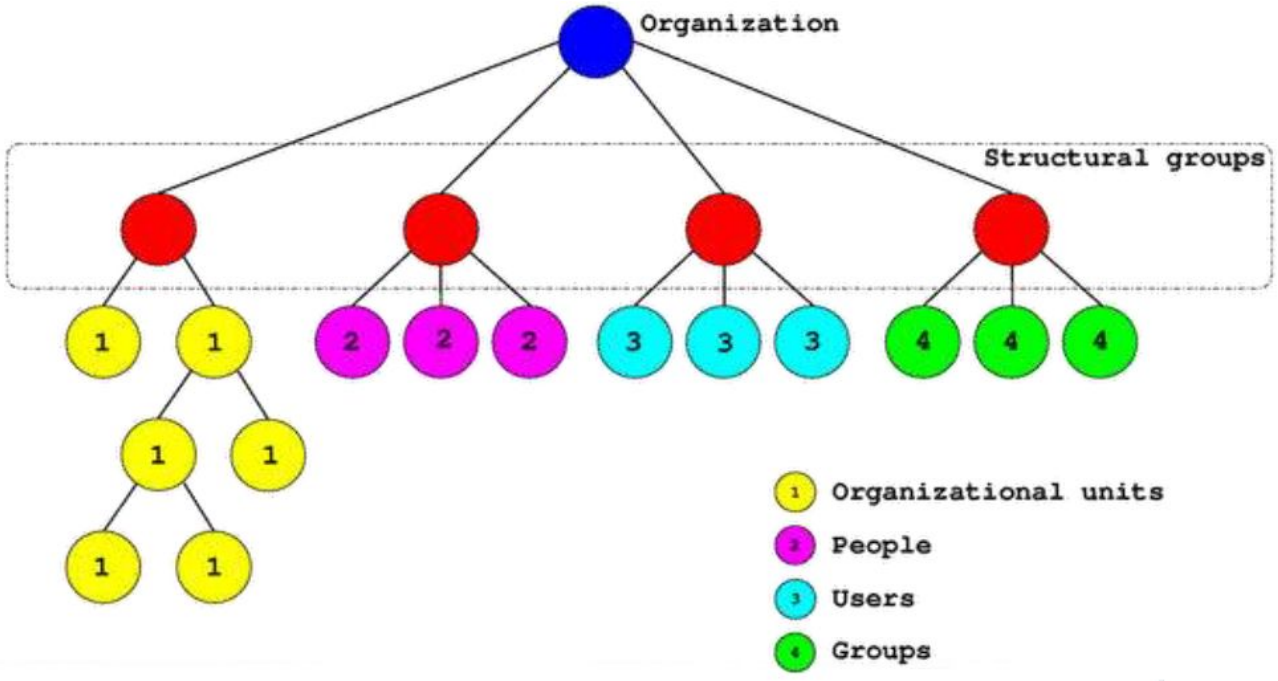


LDAP

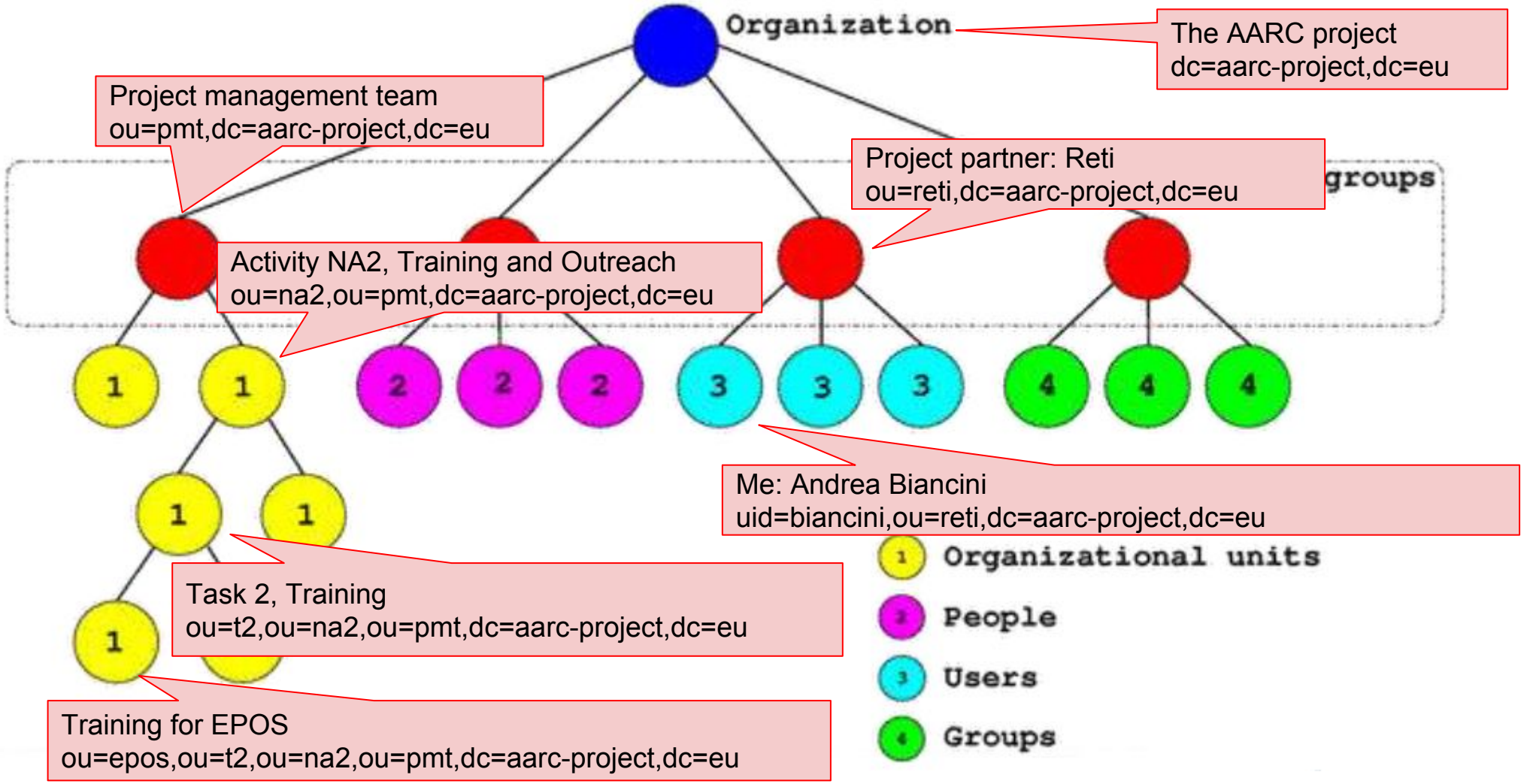
The **Lightweight Directory Access Protocol** (LDAP) is an open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services. LDAP is specified in a series of IETF RFCs.

Information in LDAP are structured as a **tree representing the organizational structure**.

Groups, people and users are then represented as nodes or leaves of the tree.



LDAP - Example



Federated authentication

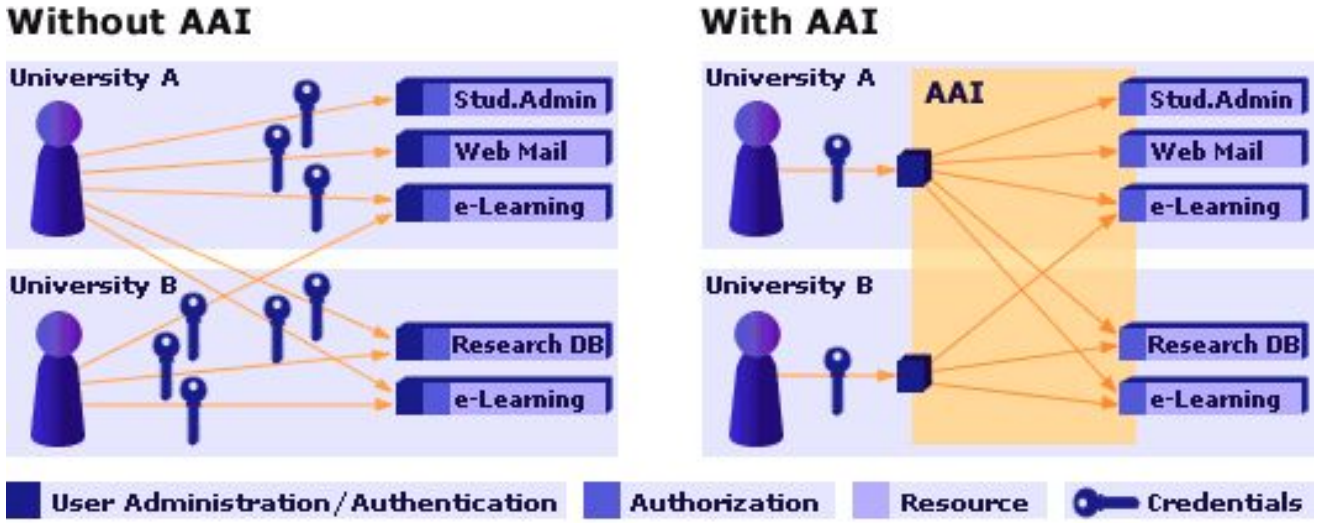
LDAP is widespread and quite always used to maintain authentication information of an organization. Often to facilitate collaboration is useful to enable access to service to **users that belong to different organisations** compared to the one that operate the service. This can be possible by either asking people to create an account with that services or by enabling federated access.

To permit these users to access the application, we use a **federated identity**:

- users can authenticate on different identity provider (IdP) services on the network;
- the different IdPs use similar protocols and user definitions so that applications can deal with users belonging to different organization in a similar manner.

Federated authentication

The objective of the AAI is, in a nutshell, to **simplify inter-organizational access** to resources. With a single login, for instance, a researcher can access applications at multiple organizations (universities or research institutions).



Benefits of federated authentication

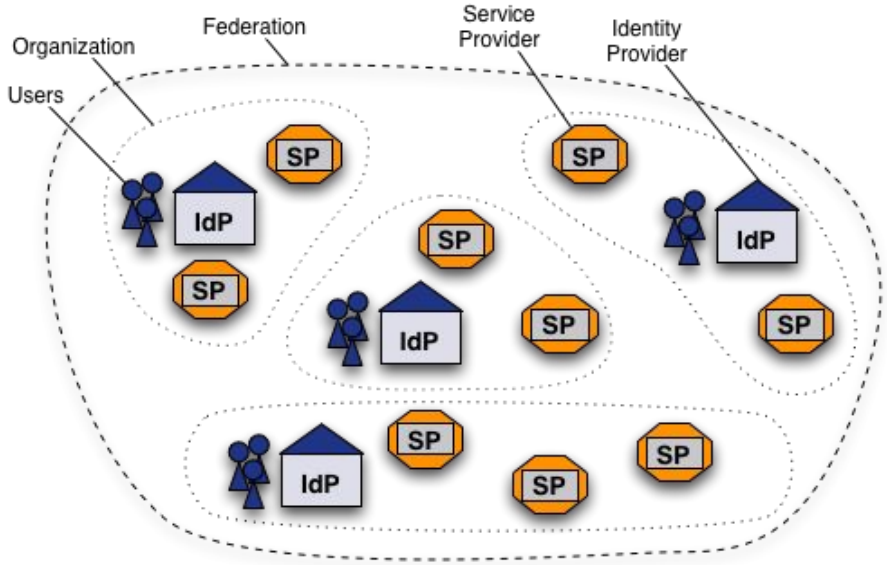
- **A user registers only once** - namely with his/her so-called home organization to which the user is affiliated. This Home Organization is responsible for maintaining the user related information and provides the user with the credentials. Home Organizations can be institutions like universities, libraries, university hospitals etc.
- **Authentication** is always **carried out by** the **user's Home Organization**, which can also provide additional information about the user to the Resource upon Resource's request and user's consent.
- **All AAI-enabled Resources are available to a user** with a single set of credentials. At the same time, there is no need for Resource operators to register new users, because they get the required information directly from the user's Home Organization.
- An **access control decision** (authorization) is made by the Resource based on the retrieved information about the user.

What is a Federation

A federation is a **collection of organizations** that agree to interoperate under a certain rule set. Federations will usually define trusted roots, authorities and attributes, along with distribution of metadata representing this information.

In general each organization participating in a federation operates:

- one **Identity Provider** (IdP) for their users, and
- any number of **Service Providers** (SP) or applications.



SAML

SAML (Simple Assertion Markup Language) is a **standard** that **facilitates the exchange of security information**. Developed by OASIS, SAML is an **XML-based framework**. SAML enables different organizations (with different security domains) to securely exchange authentication and authorization information.

To create a SAML infrastructure:

- an **IdP** must be **installed on top of each organization directory** to permit user authentication in the federation
- an **SP** must be **installed on top of each application** to consume authentication and authorization information obtained from the federation

OpenID Connect

OpenID Connect (OIDC) is a **standard** that **facilitates the exchange of security information**. Published by the OpenID Foundation, OIDC is a framework that uses **REST APIs** and **JSON** format. OIDC, as SAML, enables different organizations (with different security domains) to securely exchange authentication and authorization information.

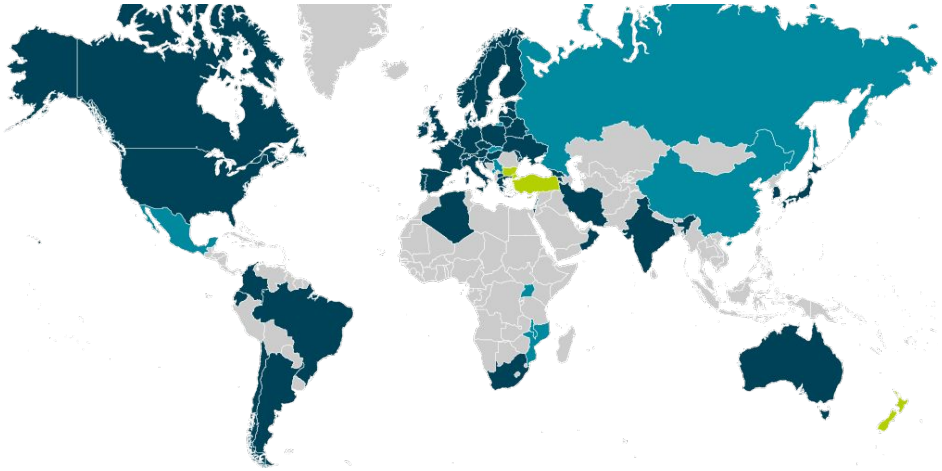
To create an OIDC infrastructure:

- an **OP** must be **installed on top of each organization directory** to permit user authentication in the federation (similarly to IdP in SAML)
- an **RP** must be **installed on top of each application** to consume authentication and authorization information obtained from the federation (similarly to SP in SAML)

Why these two frameworks?

In our communities, we need to consider at least these two framework:

- **SAML** was implemented by different Research and Education Networks all over the world and has a very **strong legacy in R&E**;



- **OIDC** is sustained by **many internet companies**.

Google KDDI Microsoft
NRI The Office of the National Coordinator for Health Information Technology ORACLE
PayPal PingIdentity Symantec.
verizon

Federations in R&D

A **federation operator** is an organisation that operates an identity federation. Operation typically includes at minimum:

- Collecting, processing and republishing federation metadata *(metadata permits to create a trust between IdPs and SPs to communicate securely)*
- Common policies and legal frameworks that all federation participants adhere to
- Guidelines and deployment instructions to operate services in the federation
- Helpdesk to assist with deploying services and debugging issues

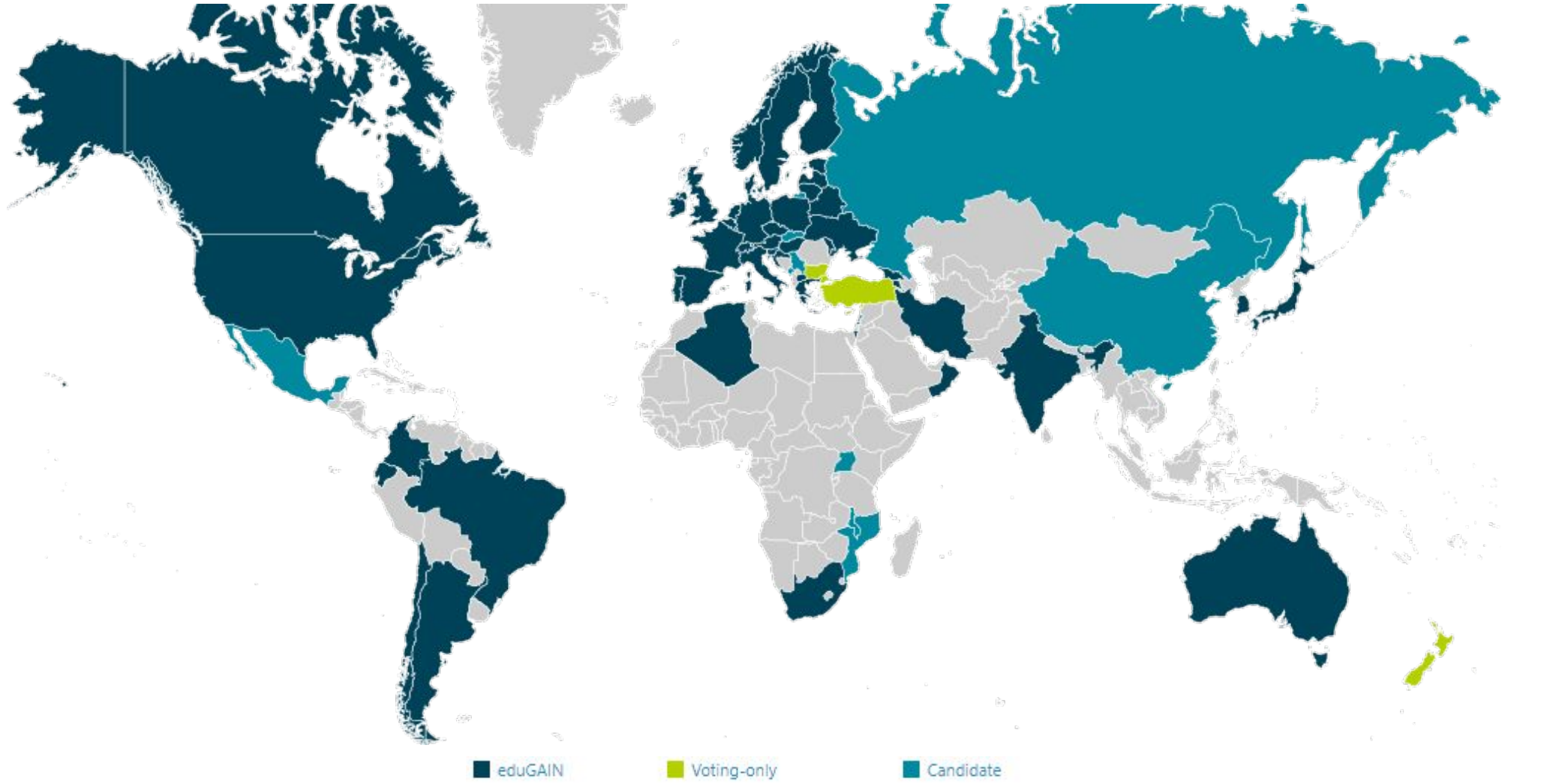
Most academic federations are operated by the **national research and education network** (NREN). These organisations typically also operate the network connecting the universities and research organisations within a country.

The inter-federation

NRENs usually operates federation within a country.
To scale to a global level, R&E introduced the concept of **interfederation**.

Interfederation takes place if a **user from one federation accesses a service which is registered in another federation**. eduGAIN is the most known and largest academic Interfederation service to exchange trusted identity information across boundaries of (national) identity federations.

eduGAIN



What we have learnt

- ★ What is authentication and what is authorization
- ★ Which are the most common ways of doing authentication and authorization
- ★ What is a directory for an organization
- ★ How the information used for AAI are stored in a directory
- ★ How to create a federation to create trust between different organizations
- ★ What is an inter-federation

Section 2 - How to start federating

- What is an SP, what is an IdP
- Trust between SPs and IdPs
- How to deal with authentication
- How to deal with authorization

What is an IdP

An **identity provider** (IdP) is a system entity that creates, maintains, and manages **identity information** for principals while providing authentication services to relying party applications within a federation or distributed network.

An identity provider offers **user authentication as a service**. Relying party applications, such as web applications, outsource the user authentication step to a trusted identity provider.

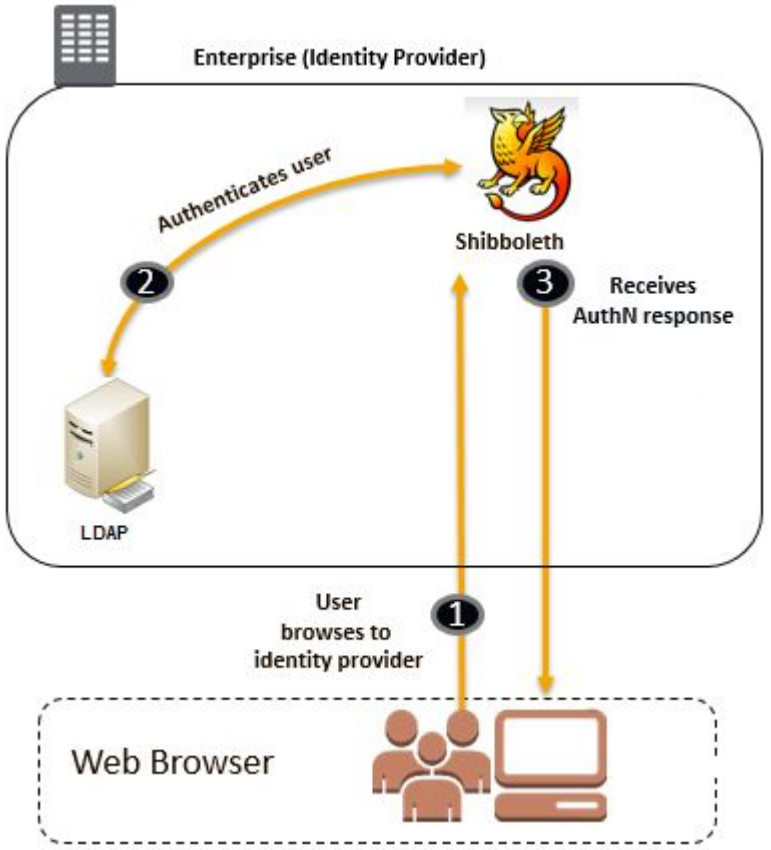
The IdP can also provide additional **attributes** and convey them to the service providers. Examples include user data (name, email, …) or affiliation with the IdP's organization

How to transform a directory into an IdP

An IdP is usually created by using a specific software package configured to provide identities to a federation:

- for **SAML**: e.g. **Shibboleth IdP** or **SimpleSAMLphp**
- for **OIDC**: e.g. **pyoidc** library

The IdP uses a backend datasource to authenticate users and retrieve attributes. A LDAP directory is commonly used.



What is a SP

A **service provider** (SP) is a system entity that **receives and accepts authentication assertions** released by an Identity Provider.

The SP stays in front of user facing applications and takes the **responsibility to receive authentication information** and to verify them according to federation rules.

Examples of SPs:

- SAML Service Provider
- OpenIDConnect Relying Party

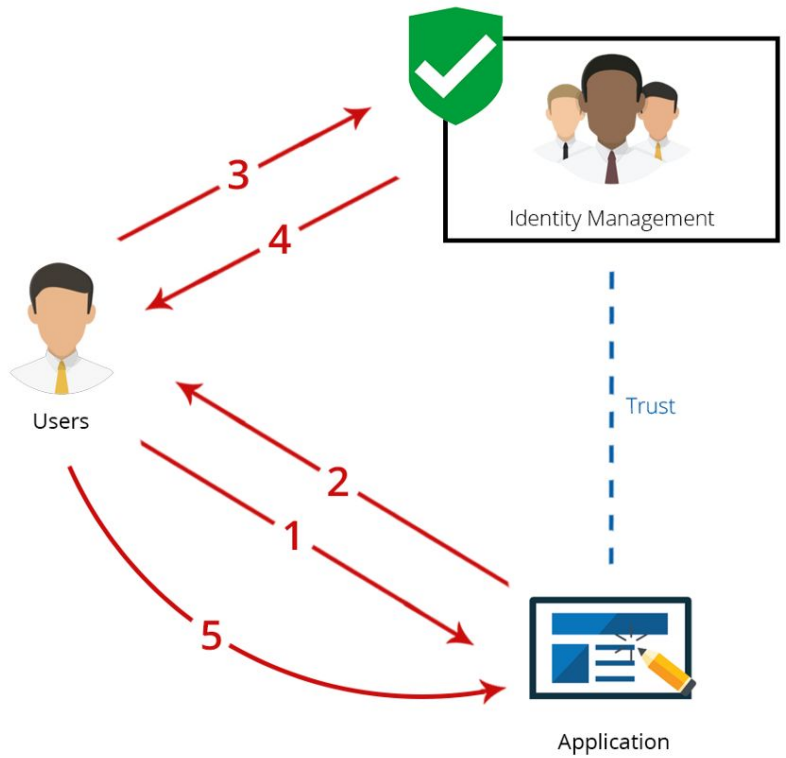
How to federate an application

To federate an application, specific library or software must be configured to handle federated authentications:

- for **SAML**: e.g. **Shibboleth SP** or **SimpleSAMLphp**
- for **OIDC**: e.g. **pyoidc** library

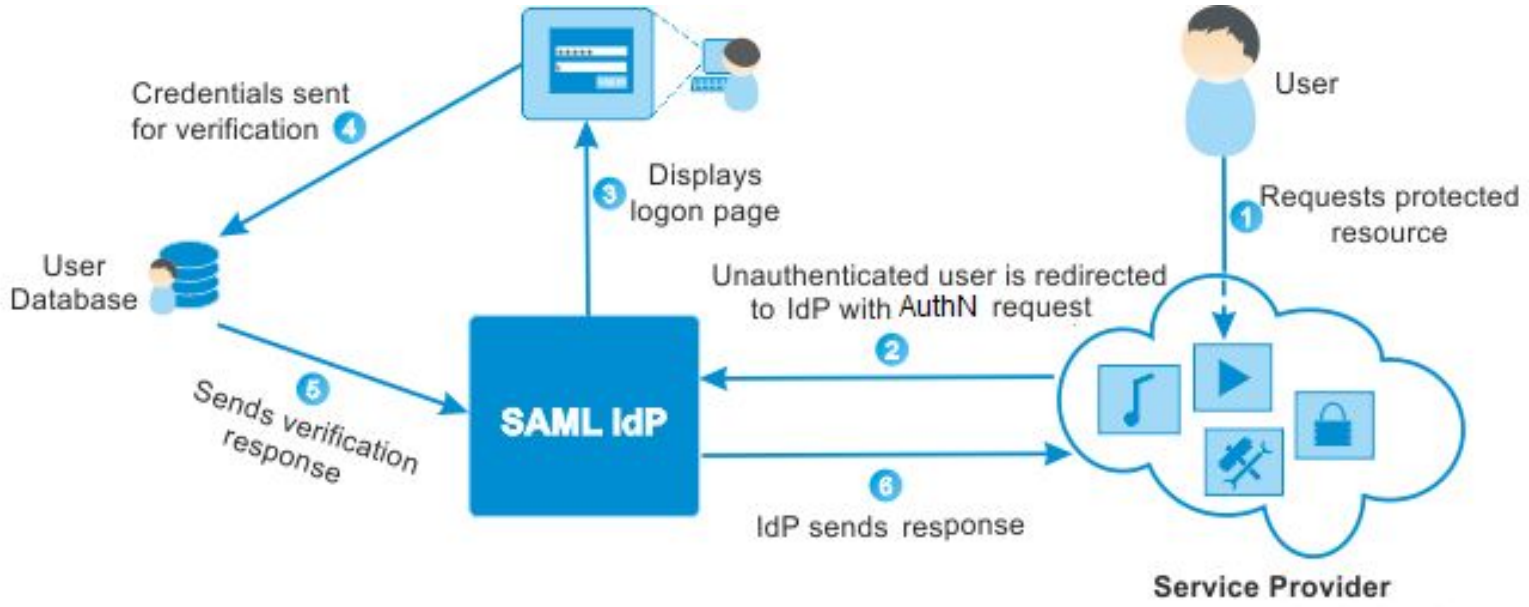
All these software takes the responsibility to interact with entities at the federation level (=IdPs) to exchange digital identities information.

The application itself can then consume the user attributes obtained from an IdP.



How to deal with authentication

Example of the authentication flow with SAML:



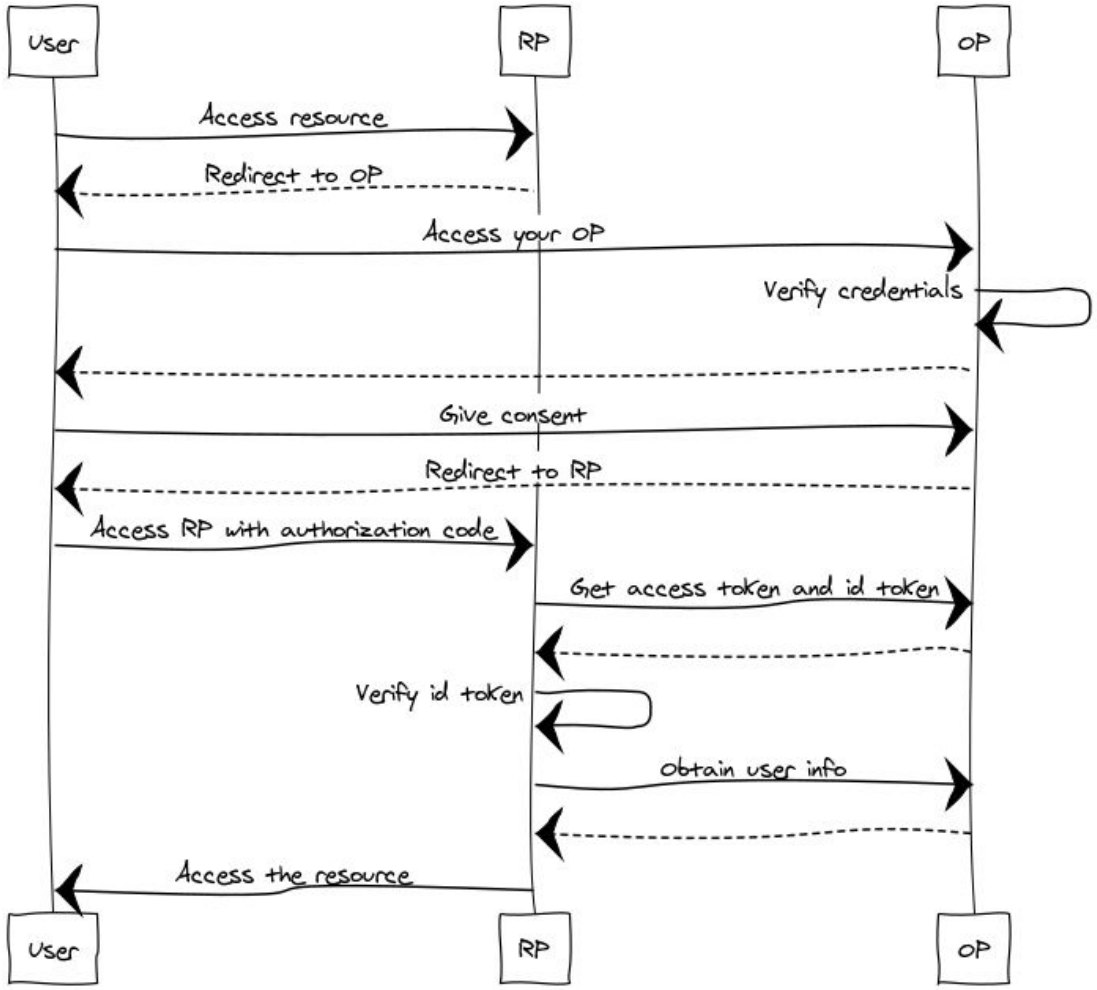
How to deal with authentication

Example of the authentication flow with OIDC

This shows the **Authorization code flow**

The general flow is very similar to SAML

Note that the access token and id token are requested via backchannel between RP and OP

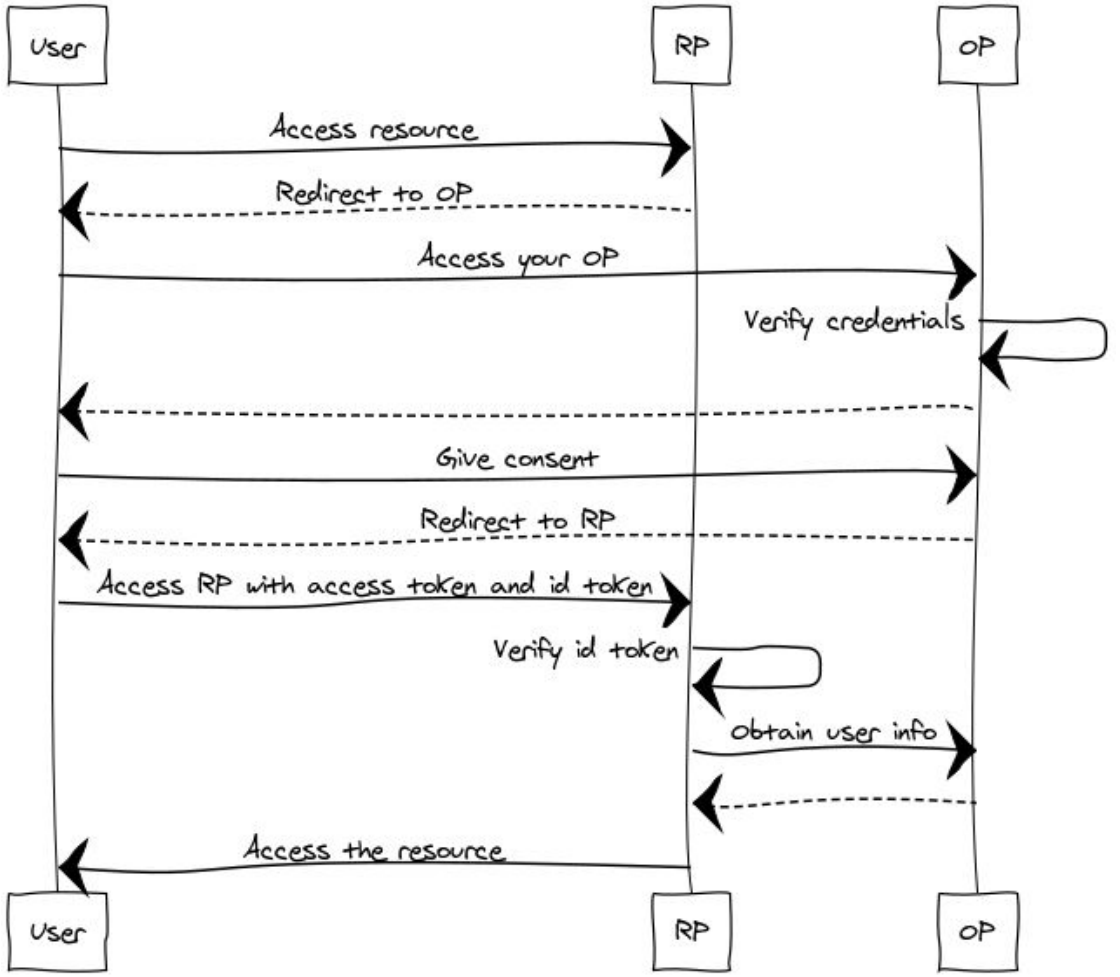


How to deal with authentication

Example of the authentication flow with OIDC

This shows the **Implicit Flow**

Access token and id token are transmitted directly to the RP



Trust between IdPs and SPs

Trust between IdPs and SPs is established with **technical and policy** means.

IdPs and SPs usually use **PKI** to **communicate in a secure way**. They tend to use:

- **encryption**: to prevent someone unauthorized from reading the messages;
- **signature**: to guarantee sender and receiver are exactly the entities they pretend to be *(avoid man in the middle attack)*

Usually this information needs to be configured beforehand (SP and IdP explicitly add each other's metadata and create trust that way)

This might work when you have one SP and only a couple of IdPs. Want to offer a service to all IdPs in europe? Good luck...

Trust between IdPs and SPs

The **trust** between IdPs and SPs is very commonly **created by federations** that:

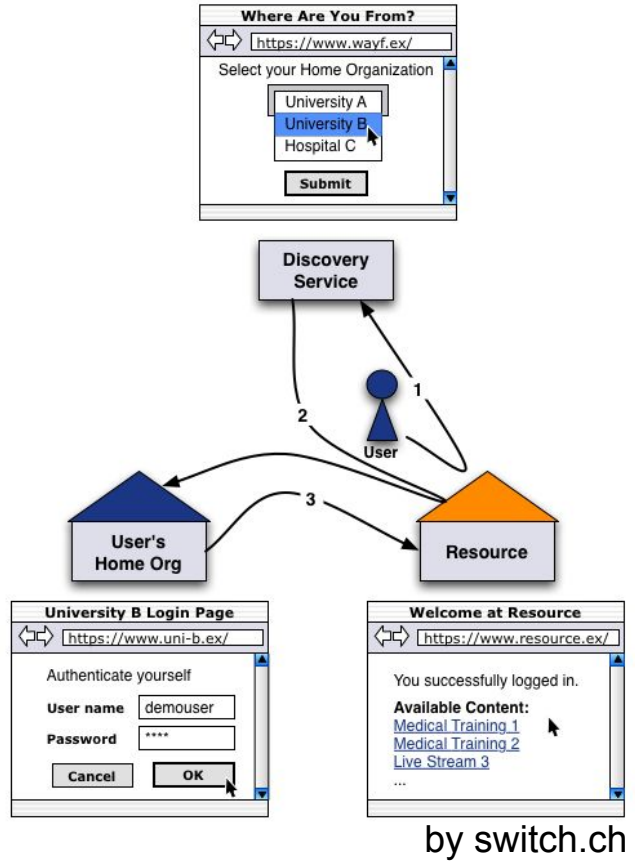
- accept **registration** for all entities in the federation
- **verify** their adherence to federation policies
- create and distribute a set of **trusted metadata** to all entities willing to operate in the federation.

Every entity in the federation trusts the federation operator and consumes the combined metadata of all participants

Discovery Service

Since there are several IdPs in the combined metadata of the federation, the SP needs to choose which IdP to use

This is done by a Discovery Service (DS), which is commonly run by the federation operator

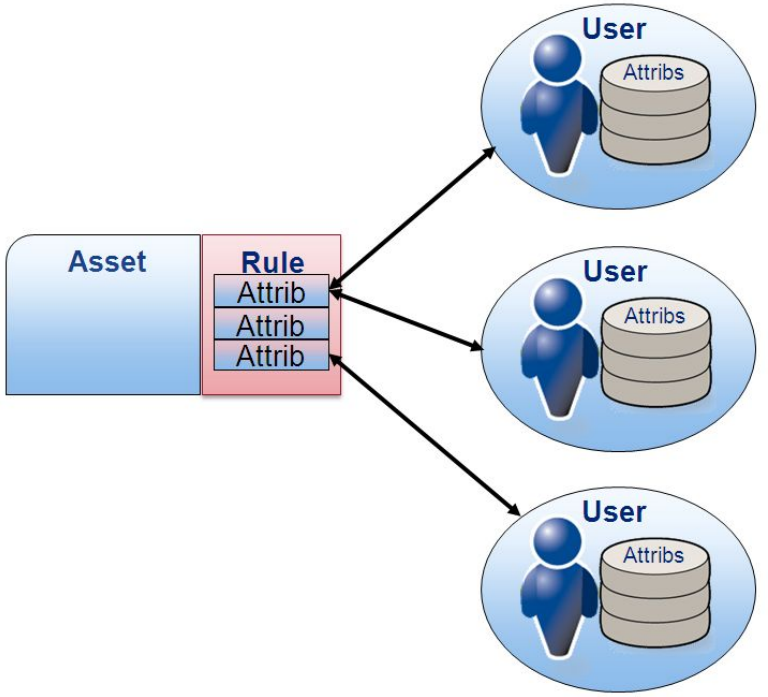


by switch.ch

How to deal with authorization

The **authentication response** contains **information about the groups** the user belongs to. These groups can be used to permit or deny access to specific operations.

The **rule enforcement** is performed **in the application** with these information coming from the federation.



From application managed authentication to federated authentication

1. **create a Directory** (with LDAP) at the organization level;

2. modify all applications **removing username/password management** and leveraging the identity provided by the external directory and by implementing **RBAC access policies**;

3. **install an IdP** on the directory and federate it *(for instance: install Shibboleth IdP and register the IdP on the NREN identity federation and in eduGAIN)*;

4. **install and configure SP** on each application so that they can consume credentials coming from the federation *(for instance: install Shibboleth SP, register the SP on the NREN identity federation and in eduGAIN and use credentials coming from the SP)*.

Benefits for users

Using federated access gives users a set of benefits:

1. users do have to register (and remember) only **one username/password**: the one on the IdP that can be used to access different applications from different organizations;

2. users can experience **Single Sign On** (SSO) instead of having to authenticate different times to different applications;

3. users are ensured their **personal information** is **handled with care**, in a single place and according to privacy policies.

Benefits for application operators

Federated access offers benefits also to application operators:

1. they do not have to deal with **user maintenance problems** (password, information updates, …)

2. they can **trust users** they have never seen, given the trust provided by federation operators;

3. with **RBAC authorization** is strongly simplified by permitting access only to specific roles and groups and not having to deal with single users.

What we have learnt

- ★ What is an SP, what is an IdP
- ★ How an IdP extends the directory
- ★ Which is the role of an SP in an application
- ★ Which is the interaction for authentication
- ★ How to implement authorization
- ★ Which are the main benefits a federated AAI gives users and application operators
- ★ How is it possible to simplify attribute release within an AAI

Section 3 - Examples of AAIs

- Why an AAI for community
- eduGAIN
- ELIXIR AAI
- BBMRI AAI
- DARIAH AAI

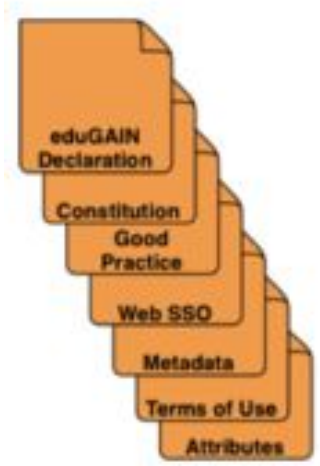
Why an AAI for a research community

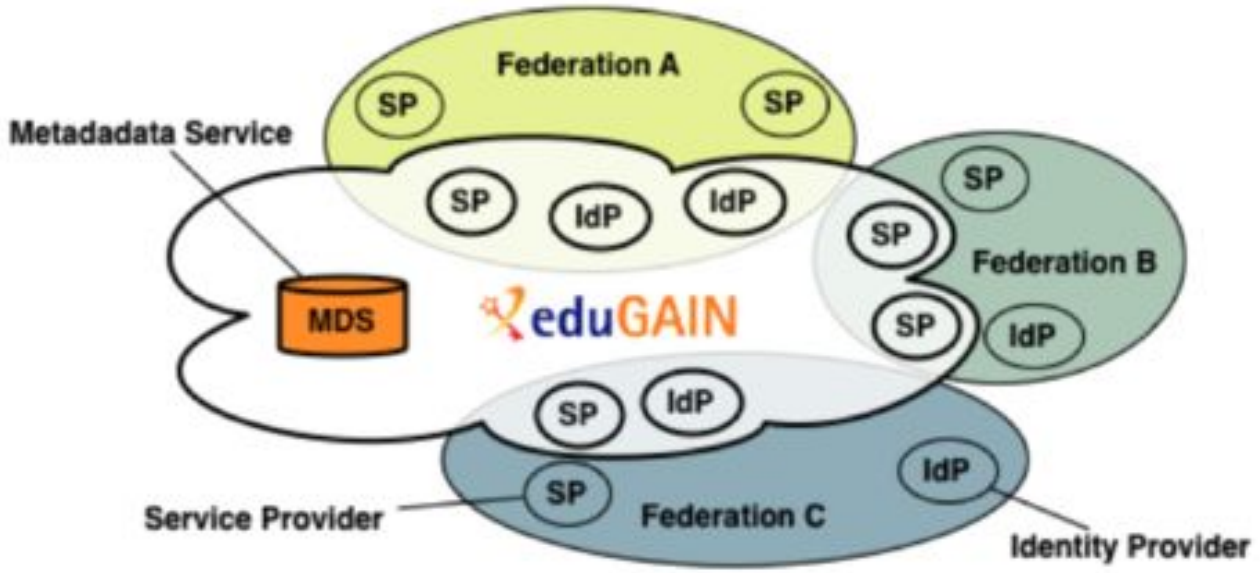
By creating an AAI at a research community level, many benefits can be achieved, for instance:

- service operators do not have to deal with authentication but can leverage existing processes;
- users do have a common way for accessing services;
- users can benefit from SSO;
- problems like policy compliance can be implemented once for the whole community;
- ...

eduGAIN

- The eduGAIN service **interconnects identity federations** around the world, simplifying access to content, services and resources for the global research and education community. eduGAIN enables the trustworthy exchange of information related to identity, authentication and authorisation (AAI).





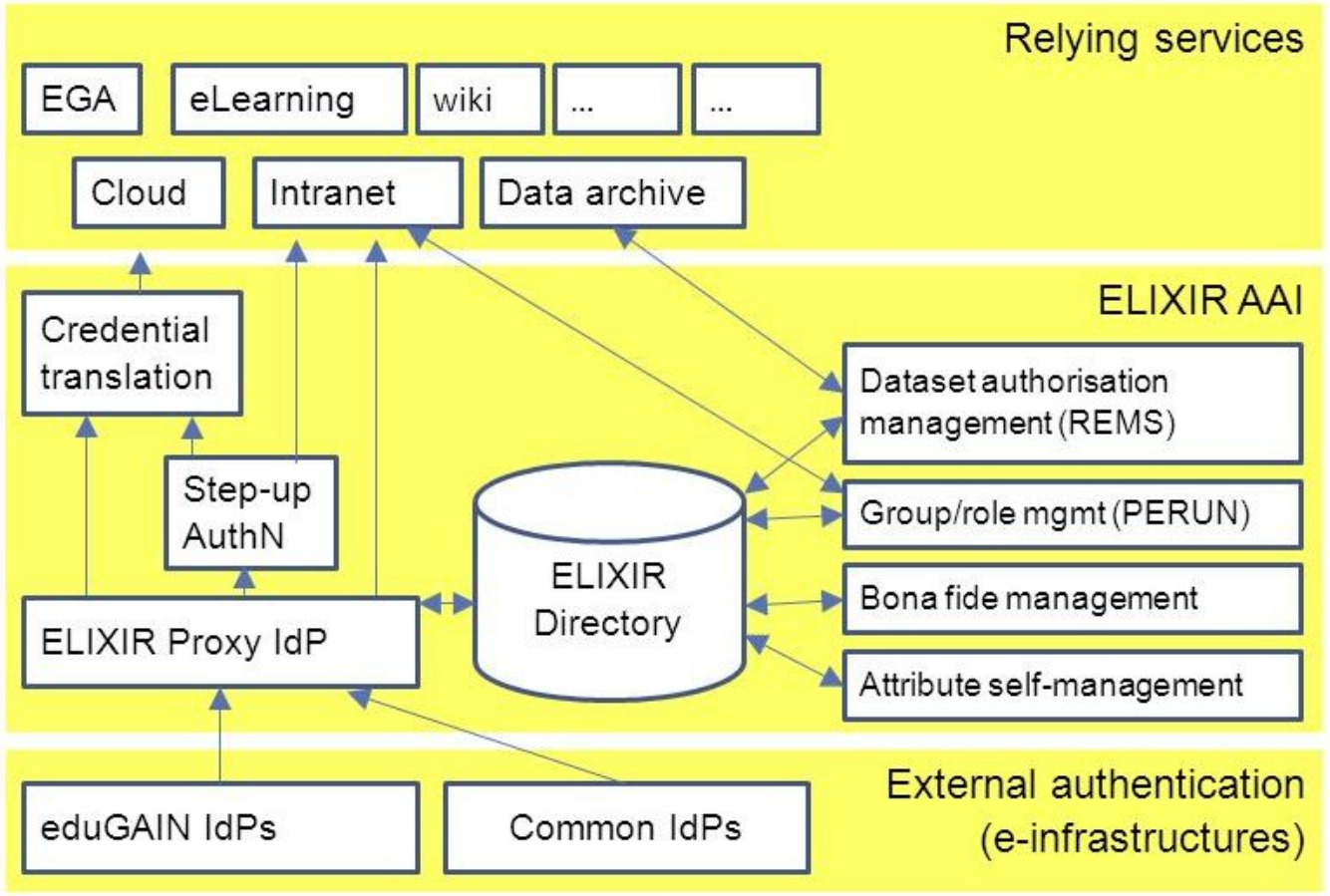
Technical details

- eduGAIN is built upon **national identity federations** usually managed by NRENs and built on **SAML protocol**.

- The user attributes can be different on different national federations. eduGAIN defines the attributes that can be used as: the **attributes defined** in the **eduPerson** or the **SCHAC** schema. Additional attributes can optionally be used.
  - displayName
  - cn
  - mail
  - eduPersonAffiliation
  - eduPersonScopedAffiliation
  - eduPersonPrincipalName
  - eduPersonTargetedID
  - schacHomeOrganization
  - schacHomeOrganizationtype

ELIXIR AAI

- The ELIXIR Authorization and Authentication Infrastructure (AAI) allows single sign-on to services across ELIXIR.
- ELIXIR leverages **eduGAIN** authentication.
- It adds:
  - group/role management
  - step up authentication
  - credential translation
  - user management of attributes

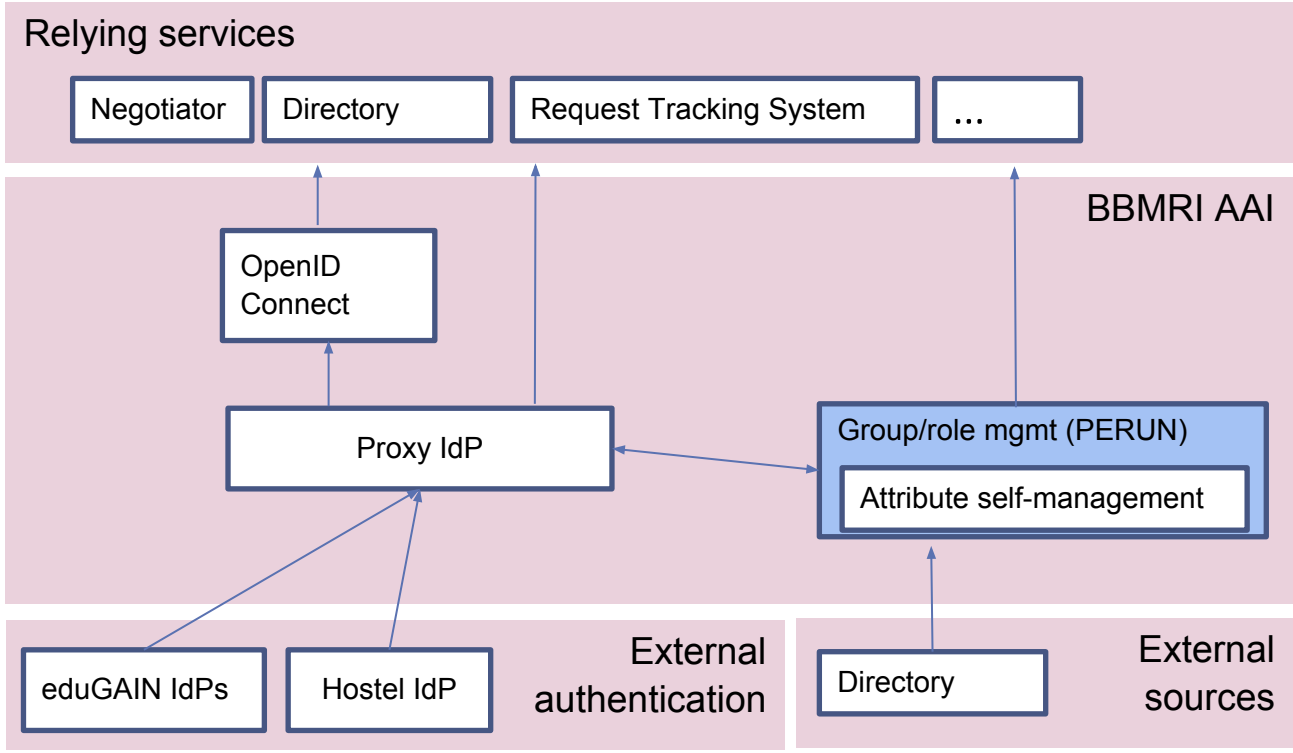


ELIXIR AAI - Main requirements

- **Unique identity**: Each ELIXIR identity should represent a single natural person and should be uniquely identified by two independent identifiers.
- **Up--to--date affiliation information**: Each ELIXIR identity can be associated with one or more affiliations, known as home organisations.
- **User--managed identity attributes**: A user should be able to self--manage some of their attributes through a web--based UI.
- **Multiple authentication providers**: A user should be able to link one or more external authentication providers to their ELIXIR identity.
- **Level of Assurance**: the eduGAIN IdPs must provide sufficient LoA.
- **Common attribute policy framework**: All participating entities in the AAI ecosystem should commit to a common policy framework for the processing of personal data.
- **Step--up authentication**: The ELIXIR AAI should provide a step--up authentication service.
- **Groups and roles**: Each user can belong to one or more groups.

BBMRI AAI

- The BBMRI AAI uses the same concept and components as ELIXIR AAI.
- BBMRI leverages **eduGAIN** authentication and also IdP of the last resort.
- It adds:
  - group/role management
  - user management of attributes
  - synchronization with external sources

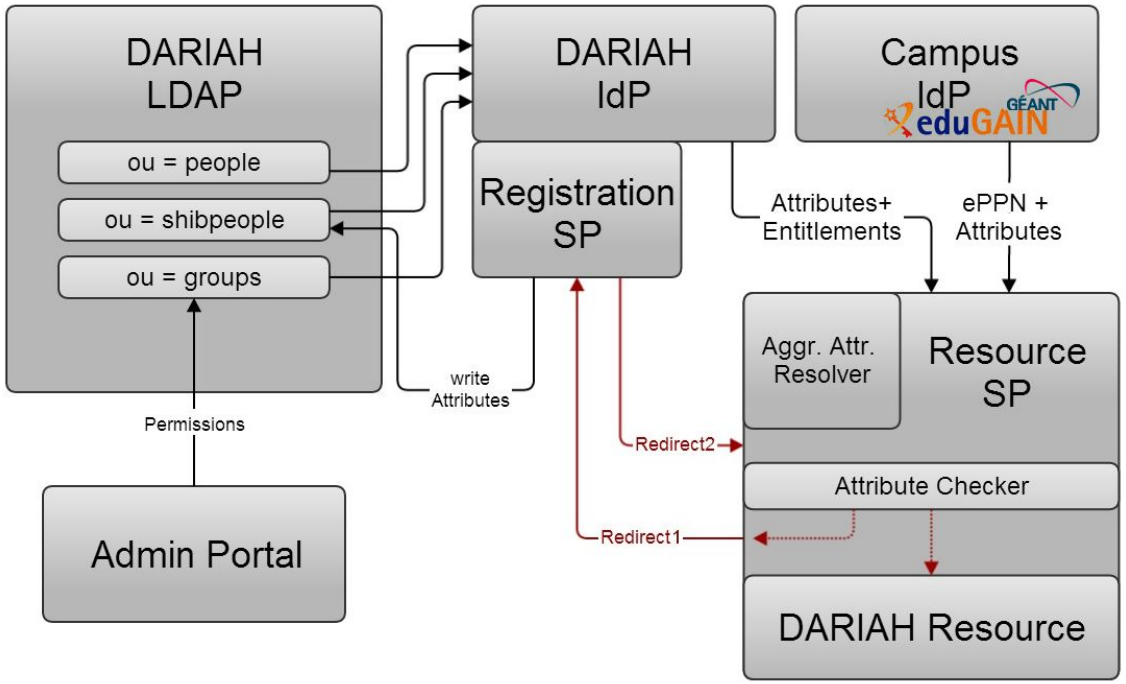


BBMRI AAI - Main requirements

- **Unique identity**: Each BBMRI identity should represent a single natural person and should be uniquely identified in a trusted way.
- **Up--to--date affiliation information**: Each BBMRI identity can be associated with one or more affiliations, known as home organisations.
- **User--managed identity attributes**: A user should be able to self--manage some of their attributes through a web--based UI.
- **Multiple authentication providers**: A user should be able to link one or more external authentication providers to their BBMRI identity.
- **Level of Assurance**: the eduGAIN IdPs must provide sufficient LoA.
- **Groups and roles**: Each user can belong to one or more groups.
- **AUP management**: Track changes in AUP and force users to approve them.
- **External synchronization**: Import of groups from external sources.

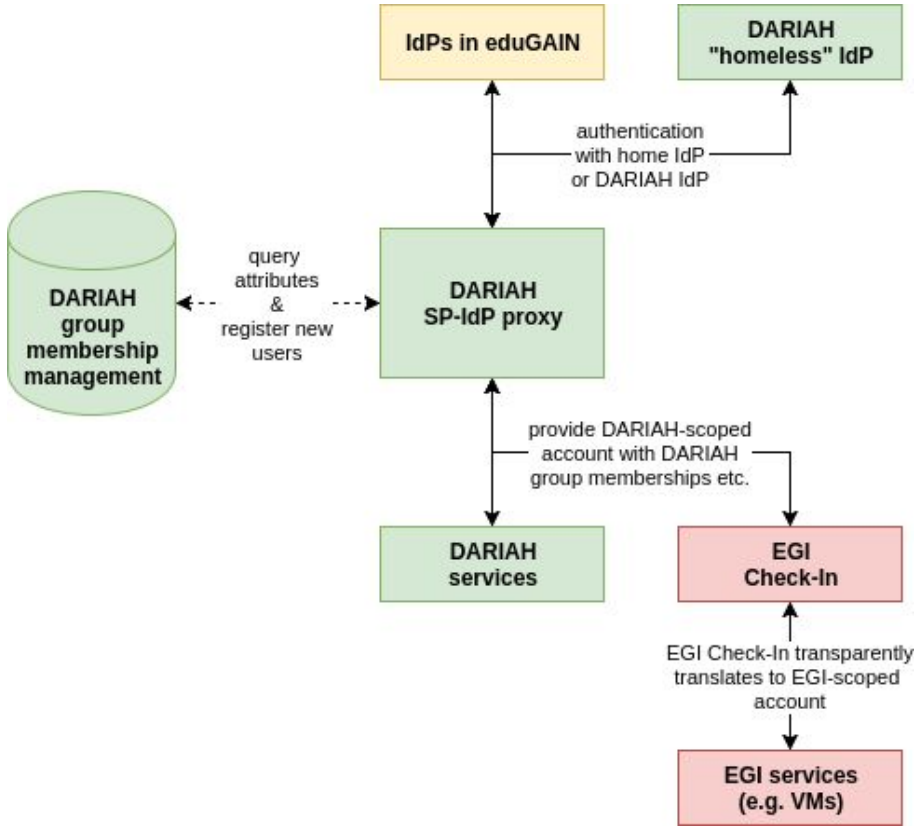
DARIAH AAI

- The DARIAH Authentication and Authorization Infrastructure (DARIAH AAI) is based on **SAML** and Shibboleth in the European higher education identity inter-federation **eduGAIN**.
- Currently SPs have to connect directly to all eduGAIN IdPs and deal with attribute aggregation, checking of policies etc. themselves
- DARIAH also provides a homeless IdP



DARIAH AAI

- The current approach does not scale very well and involves a lot of tedious work for SPs
- DARIAH is currently transitioning to a proxy architecture



DARIAH AAI - Main requirements

- The AAI service must be **easy to use**, ideally using their own institutional credentials (if available) on a Europe-wide and international Level.
  - DARIAH AAI leverages eduGAIN but integrates it with a productive solution based on homeless-IdP
- **Single sign-on** to all (DARIAH) resources, tools and services.
- **Authorisation granularity**, e.g. access to 'sensitive data'
  - DARIAH AAAI manages this with a specific attribute authority that adds authorization information to the users as additional attributes.
- Authorization of **non-web-based services**.

What we have learnt

- ★ We have seen some examples of AA infrastructures
- ★ eduGAIN for the Education and Research world
- ★ ELIXIR
- ★ BBMRI
- ★ DARIAH

Thank you

Any Questions?

https://aarc-project.eu

The work leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 730941 (AARC2).