Introduction - Federated Identity and Virtual organizations

Daan Broeder
Max-Planck Institute for Psycholinguistics
CLARIN EU / WP2

Workshop on Virtual Organizations and Federated Authentication, Mannheim, March 2011
A virtual organization or company is one whose members are geographically apart, usually working by computer e-mail and groupware while appearing to others to be a single, unified organization with a real physical location.

- *WhatIs*

In Grid computing, a *Virtual Organization (VO)* refers to a dynamic set of individual and/or institutions defined around a set of resource-sharing rules and conditions.

- Wikipedia

A *Virtual Organization (VO)* is a group of individuals that have something in common, in example working on a project. A *Virtual Organization* may span multiple real organizations, from multiple federations and multiple countries.

- Feide website
EU SSH e-infrastructures

- A limited number of ‘capable’ centers ±100
  - Hosting (in total) very many >10M resources, that should be individually accessible
  - many services (web applications & services) 1000-10000

- A large number of research institutes, universities
  - potentially very large number of users >100k
  - Organized in different national IDFs

- Access to resources & services based on:
  - Identity, EULA signing, affiliation, ?

- Access protocols
  - Web & desktop clients using HTTP(S), SOAP, REST, …
AAI for Resources & Services

Single Virtual Domain
Resources & services

User DB

Org. A

Desktop client

Web browser

Org. B

AuthZ DB

AuthZ DB

SAML2

HTTP(S)
one central AuthZ DB is a possibility, however
• centers want control
• millions of resources
• SPOF
who wants to administrate a user store of >100k users?

- no single user identity
- SPOF
SSH centers

- Center situation evolution
  - Expect consolidation of the number of ‘capable’ centers
  - Expensive to offer reliable persistent services
  - Outsourcing of services and storage to specialized data and compute centers
  - Outsourcing to complementary infrastructures
Community Infrastructure coexistence

**DASISH**

- **CLARIN**
  - community specific

- **DARIAH**
  - CLARIN LT web service infrastructure

- **CESSDA**
  - common SSH metadata catalog

**SSH communities wide - DASISH**

**Data Preservation – (EUDAT)**

**NETWORK Services – NRENS & GEANT**

**Federated AAI**

© DGB
Clients need Authentication too

The application only knows HTTP with basic authentication. It does not understand form-based authentication employed by the Shib. IdP.

User scenario: Copying resources from different repositories to the local machine.

Possible solution: Use certificates obtained from a SLCS and mod_ssl Apache.
- Which SLCS: NREN, EU, community?
- The certificate subject needs to be very specific.
- Apache configuration is difficult.

The application is also not able to profit from the SSO over archives.
WS Security/delegation for workflows

Authorization records are not shown

Composite Web service

federated authentication

(distributed) web-services
The EU dimension

- Originally we planned
  - The CLARIN SPs become members of their national IDF
  - Rely on the eduGain inter-federation to provide the trust between the national IDF

- eduGain had many problems taking off
  - attribute harmonization
  - privacy issues

- homeless users?

Diagram:
- SP1
  - IDF a
- SP2
  - IDF b
- SP3
  - IDF c
- eduGain
  - Metadata & trust
The EU dimension 2

- We have deliverables and prototypes to finish so:
  - Established a CLARIN SP organization as a legal entity able to sign contracts where needed with the national IDF's
  - CLARIN SP organization takes care of exchanging the SP metadata with the national IDF's

Diagram:
- IDF a
- IDF b
- IDF c
- SP1
- SP2
- SP3

Questions:
- homeless users?
CLARIN SP Federation

- Purpose: enable the single user identity & SSO and comply with CLARIN deliverables without waiting on eduGain
- Signs contracts with the national IDF
  - Current SPs: BBAW, CSC, IDS, MPI for Psyl, INL, UFAL, UTU, CNRS-ATILF, Meertens Institute
  - National IDF: HAKA, SURFfed, DFNAAI
  - Not all SPs operational, e.g. use correct WAYF etc.
- Lots of administrative work
- Shortcuts are made
  - E.g. asymmetric solutions. Giving access UK users without getting access to UK CLARIN SPs.
SP requires EULA signed and takes care of this but only for its own domain.

This can break the SSO if the user is required to sign the same EULA many times.

CLARIN will try harmonizing EULAs to a limited number.
Store the EULA info in the user attributes at the IdP

But how does it get there?
• Special app?
• Not every IdP will/can support this
Create special EULA service. This is part of the CLARIN SPF. CLARIN independent of the IDFs.
Group Collaboration

- **Org. A**: Jane (User DB), Sara (User DB)
- **Org. B**: Jane memberOf PrA, Sara memberOf PrA
- **VRE 1 application**: GroupDB

Diagram:
- Center
- VRE 1 application
- Org. A
- Org. B
- Jane
- Sara
- User DB
- GroupDB

 european research infrastructure
Group Collaboration

![Diagram of group collaboration with VRE 1 and VRE 2 applications, user databases, and collaborators from Org. A and Org. B.](image)
VO-platform uses

Anything that requires user attributes/information not provided by the ‘standard’ IdP

- What licenses did a user sign?
- VOs as a basis for collaboration on shared resources: common project workspaces in VREs
- Administ rate roles in the distributed data management workflow process
  - owner, depositor, curator, annotator

Some hesitations

- We have chosen for distributed user account management & distributed authorization, keep communities in charge, scaling, no SPOFs
- What about scaling and SPOF problems with this “central” service
Obstacles for federated identity use & acceptance

- Unfamiliarity of users with the technology
  - WAYF: where do I find my organization, what is my IDF (two step)
  - ARC: prompting user consent for attribute release (uApprove)
  - Need careful guiding of inexperienced users

- Scaling problems
  - Will eduGain have an opt-in policy? Every IdP has to allow its users access the inter-federation (or worse individual SPs)
  - Individual IDF can also have an opt-in policy. Every IdP has to agree to have its users access CLARIN SPs
  - Hopefully they can treat the CLARIN SPs as a single entity

- WAYF SPOF, deploying several will break the SSO
Thank you for your attention

CLARIN has received funding from the European Community's Seventh Framework Programme under grant agreement n° 212230