DARIAH- Update

6th FIM4R Workshop
Helsinki, October 2, 2013
What is DARIAH?

DARIAH: Digital Research Infrastructure for the Arts and Humanities

One of the few ESFRI research infrastructures for the humanities

DARIAH’s mission is to develop, maintain and operate an infrastructure in support of ICT-based research practices

Infrastructure is administration, software and storage services but also Curricula and Methodology

Working with communities of practice: humanities scholars supporting their VREs
Humanities VRE

TextGridLab

Rectangle mode

TextGridLab

DARIAH-EU
DARIAH AAI Practice

Current AAI set-up: a first version of an AA infrastructure has been deployed, based on two standards:

• LDAP (Lightweight Directory Access Protocol)
  – for authentication and authorization attributes
  – deploying Open Source Software OpenLDAP
• SAML (Security Assertions Markup Language)
  – for AAI within a federation
  – including Web Single Sign-On feature
  – deploying Open Source Software Shibboleth
Current Set-Up
VO Management and FIM in DARIAH
Current Challenge

– Not every institution signs federation contracts

– Not every Identity Provider releases personal attributes

– Not every resource provider allows anonymous usage

– A European humanities federation is just at its start (CLARIN federation, DASISH activities)
How to make this a European-wide Infrastructure?

• We have productive a 'flat' Group based Authorization:
  • You are member of group
    • EHRI-users allowes to access the EHRI part of the DARIAH wiki
    • collection-registry-users allowes to use the collection registry
    • collection-registry-editors allowes to input data into the registry
    • collection-registry-admins allowes to configure the registry
    • Collection-registry-groupadmins allowes to manage all groups with names beginning with 'collection registry-
• So how to delegate the groupadmins-rights?
  • We developed and implemented a hierarchical role model to delegate user rights management
How to make this a European-wide Infrastructure

• The management of the delegation is based on organisational roles (not groups) that are structured in a 3 level hierarchy (marked ●):
  • DARIAH Coordination Office as Top of hierarchy
  • Each Country has a National Representative who is allowed to:
    • Create and manage organisations and the organisation admin role
  • Each Organisation in a country has a organisation admin
    • Organisation admin is allowed to:
      • Create and manage groups (of projects the organisation is leading)
      • Create 'homeless'-accounts if needed
How to make this a European-wide Infrastructure

• So software there, now we need to organize it:
  • Who will be National Representative
  • What shall she be able to do except creating organisations and orgadmin roleoccupantships?
  • What will the organizational application process look like?
  • What more data do we need about the users
    • By now: Name, email, preferred language, affiliation
    • Should we add ORCID-IDs?
How to get the urgently needed European humanities federation?

• By now the demonstrated infrastructure is only accessible via DFN-AAI or via a dedicated DARIAH and TextGrid ('homeless')-Account.
• EduGain, the European federation of national federations is evolving
How to get the urgently needed European humanities federation?

Two ways forward:

- DARIAH IdPs and SPs either participate via the national federations
- Or they create (together with CLARIN and other DASISH partners) a humanities federation that can become member of eduGain

A GÉANT 3 Plus Pilot project with DARIAH has started to evaluate these options

There will be a Humanities federation Workshop with DASISH partners to decide on this in October 17/18 in Cologne

DARIAH-EU
Thank you for listening

Is there time for questions?
Further technical Plans

• It is planned to include technologies like OAuth2 and OpenID Connect into the DARIAH SAML based infrastructure
  • It is possible to have a SAML based Authentication within an OAuth-Infrastructure as well as
  • To have an OpenID based authentication in a SAML based infrastructure.
• Experiments on these technologies have been performed successfully
• Main aim is that an application developer only has to support one API for AAI.
Web Services

- Adoption of ECP both in SPs and clients low
- Consideration of OAuth2
  - Simpler implementation for clients, pure HTTP(S) and JSON
  - Authorization Server could be shared for multiple resource servers → presumably less implementation effort on the resource side
  - Allows for 1-tier delegation
  - SAML IdPs can be connected via SAML Bearer Token
  - Access and Refresh Token instead of login/password
  - Natively uses OpenID Connect for AuthN (and other mechanisms possible, instead of SAML, if needed)
- OAuth2 standard pushed by the industry, so probably better bet in the future?
Web Services and Delegation: OAuth2

- New IETF Standard OAuth2 provides both for
  - delegation (1-Hop, not N-Tier) and
  - simple RESTful clients API (compared to ECP)
- OAuth2 is an Authorization Protocol, and leaves the Authentication method open
- Token-based mechanism:
  - one-time authorization code
  - access token
  - refresh token
- Can be integrated in a SAML federation using the OAuth2 SAML Bearer Profile
- Simply by protecting the OAuth2 Authorization Server's /authorize endpoint by a SAML SP within the federation
 OAuth2 AuthZ Code Flow with SAML

OAuth2 Authorization Server (AS)

/validate

/token

/authorize

SAML Service Provider (SP)

11 Optional Validation

OAuth2 Resource Server (RS)

8 Code

9 Token

10 Access w/ Token

12 Content

Browser

Web Service Client or Portal

SAML Identity Provider (IdP)

3a

2b

5b

6 AuthZ

7a Code

4 AuthN

5a SAML Ass.

3b

2a

7b

13 Content

1

2
IdPs that do not release ePPN

- Due to data protection and privacy issues, some IdP maintainers decide to only release a pseudonymous ID that is
  - cryptic
  - unique for that particular user and SP combination
  - e.g. eduPersonTargetedID (ePTID) or persistentID

- We have a solution where user self-asserts any attribute at the DARIAH registration SP

- Use a mapping table
  - SP1' ID1 maps to Registration SP IDX
  - SP2' ID2 maps to Registration SP IDX as well
  - When SP2 sends an Attribute Query for ID2, IdP maps ID2 to IDX, where all user attributes can be found

- This is work in progress!
IdPs that do not release ePPN