

DARIAH Update

7th FIM4R Workshop

23.-24.4. 2014, Frascati

Peter Gietz, DAASI International GmbH



www.dariah.eu

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



The screenshot displays the TextGrid Lab application window. The main workspace shows a handwritten document with several rectangular annotations in purple and red. A Toolkit panel on the right provides editing tools. Below the workspace, an XML editor displays the following code:

```
<sp>
<speaker><anchor xml:id="a1" />ANDRES<anchor xml:id="a2" /></speaker>
<stage><anchor xml:id="a3" />(nach einer Pause.)<anchor xml:id="a4" /></stage>
<p><anchor xml:id="a5" />Woyzeck<anchor xml:id="a6" />
<anchor xml:id="a7" />hörst<anchor xml:id="a8" />
<anchor xml:id="a9" />du's<anchor xml:id="a10" />
<anchor xml:id="a11" />noch?<anchor xml:id="a12" /></p>
</sp>
<sp>
<speaker>WOYZECK</speaker>
<p>Still, Alles still, als wär die Welt todt.</p>
</sp>
<sp>
<speaker>ANDRES</speaker>
<p>Hörst du? Sie trommeln drin. Wir müssen fort.</p>
</sp>
```

The status bar at the bottom indicates 'Rectangle mode', 'image/jpeg', and coordinates 'a: 255, r: 165, g: 165, b: 165'. The bottom right corner shows the email address 'testuser@textgrid.de'.

DARIAH VCC Virtual Competence Centers

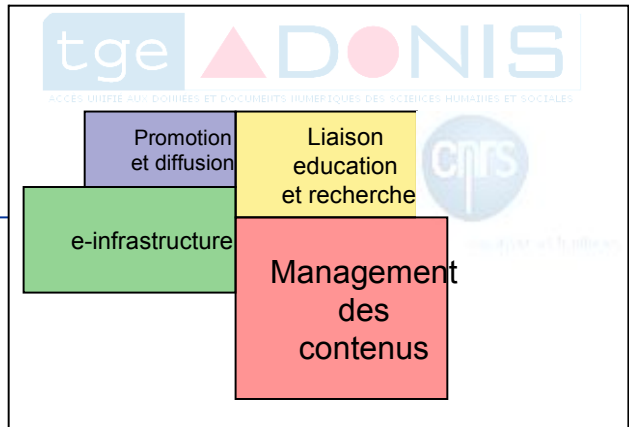


VCC Advocacy

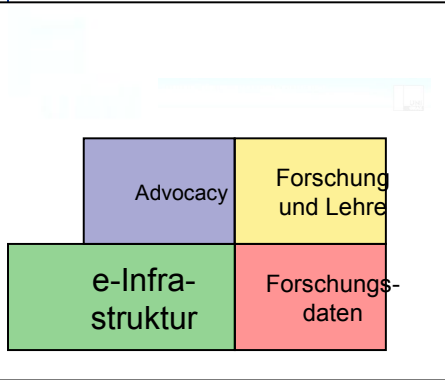
VCC Research
and Education

VCC
e-Infrastructure

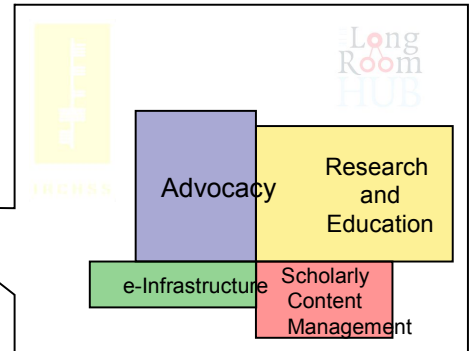
VCC Scholarly
Content
Management



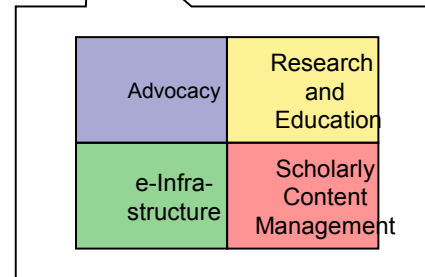
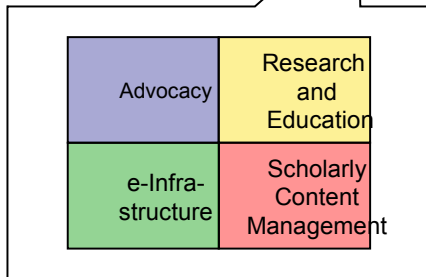
DARIAH-FR



DARIAH-AT



DARIAH-IE

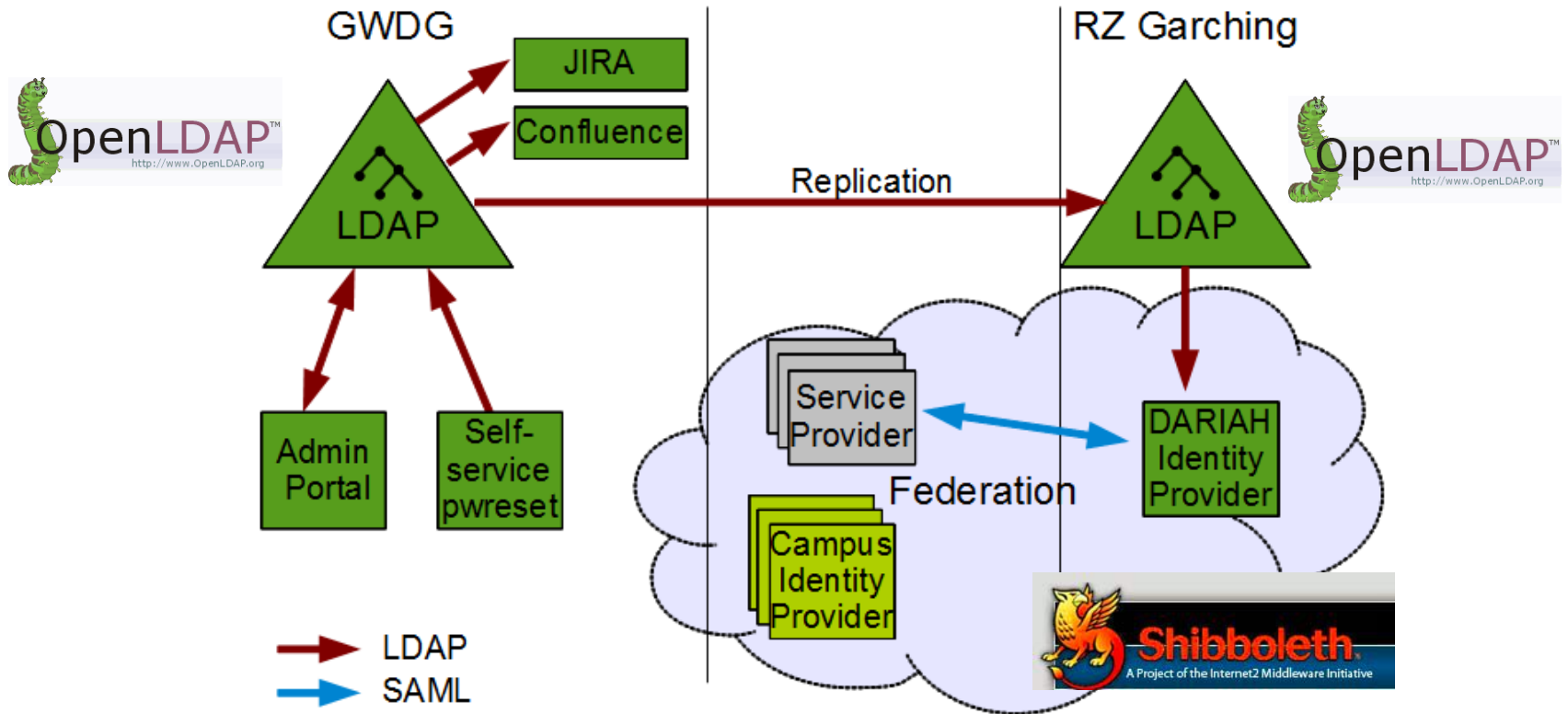


DARIAH AAI Practice

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

DARIAH AAI Setup

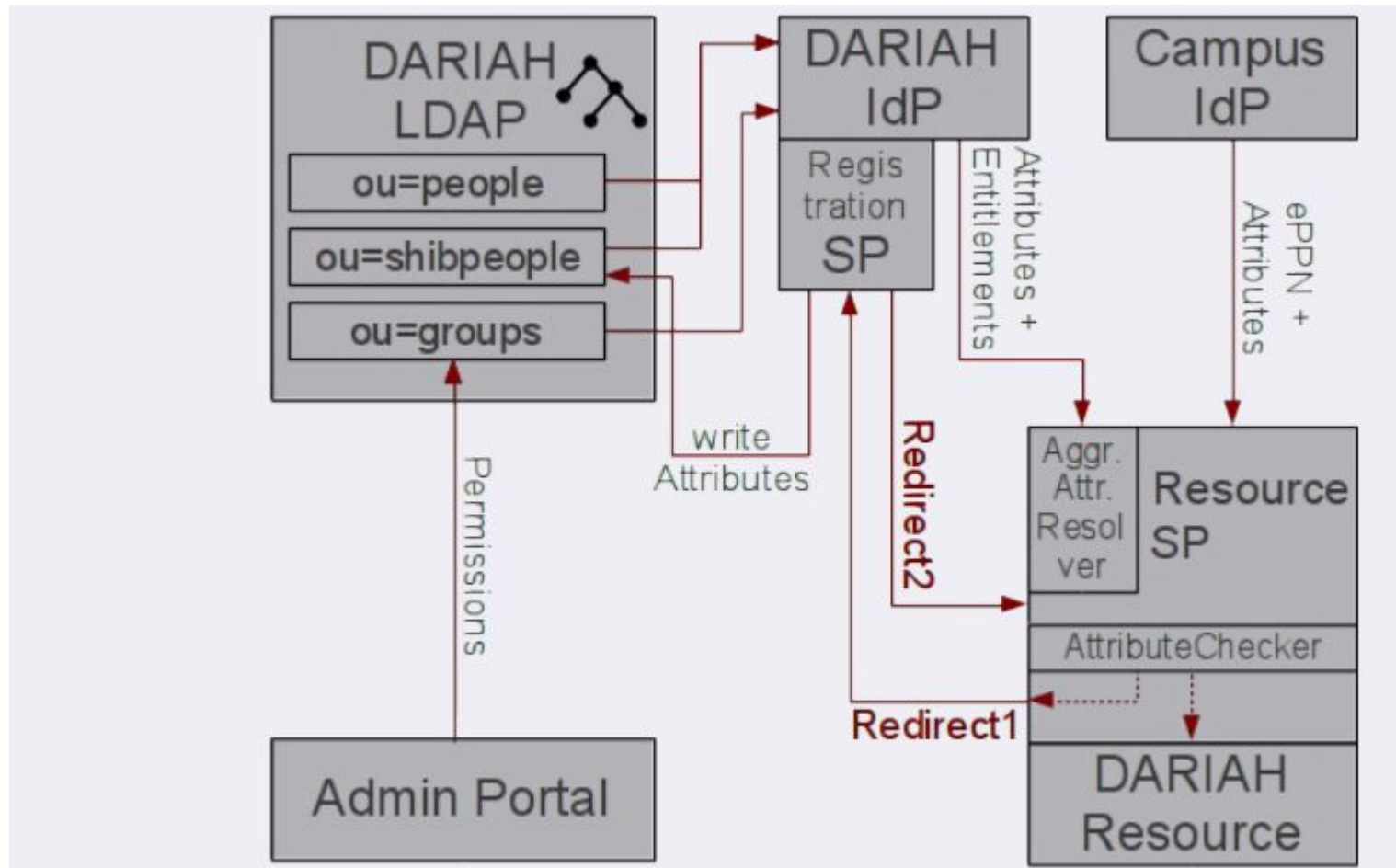


DARIAH Authorization

- No change to campus IdPs except trust / attribute filters
- Standard Shibboleth SP to protect DARIAH applications, however with special configuration:
 - aggregates attributes from campus and central IdP
 - require minimum set of attributes, otherwise redirect to registration application at central SP
- Central LDAP with authZ groups managed by admin portal
- Central IdP gets data from central LDAP and as Attribute Authority releases both user attributes and entitlements (based on groups) to SPs
- Central Registration SP writes manually completed user attributes to central LDAP



VO Management and FIM in DARIAH



Current Challenge

- European-wide federation eduGain has too little outreach
- Not every institution signs federation contracts
- Not every Identity Provider releases personal attributes
- Technologies for non-web-based access only
“almost there” (ECP, STS, Moonshot, oAuth2)

How to make this a European-wide Infrastructure?

- We have productive a 'flat' Group based Authorization:
 - You are member of group
 - EHRI-users allows to access the EHRI part of the DARIAH wiki
 - collection-registry-users allows to use the collection registry
 - collection-registry-editors allows to input data into the registry
 - collection-registry-admins allows to configure the registry
 - Collection-registry-groupadmins allows 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 :
- **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

D1

JIRA

D2

Confluence

D3

Etherpad

D4

NER

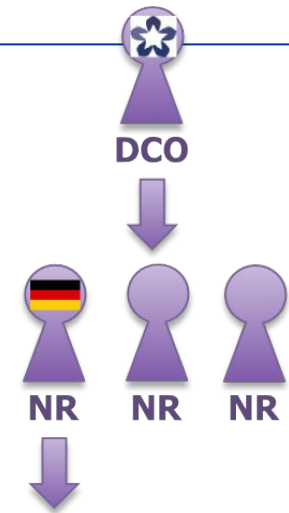
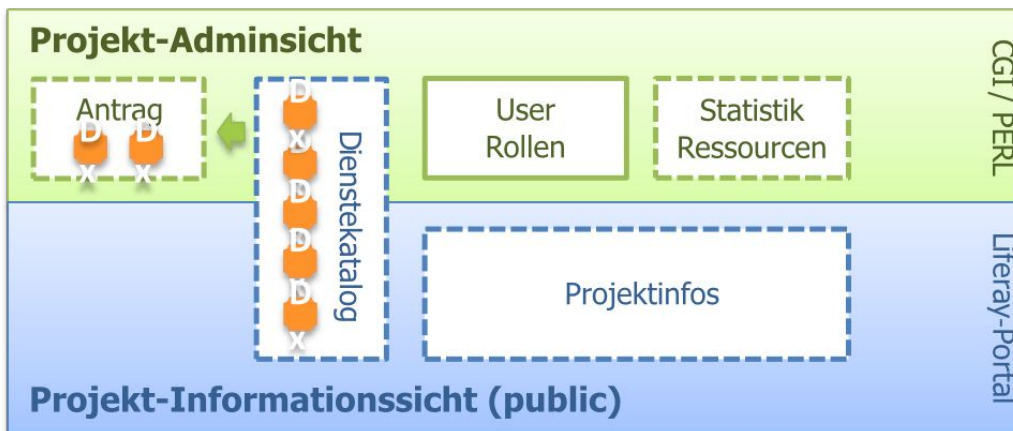
D5

Geo-Browser

Dx

...

Forscherdienste



How to make this a European-wide Infrastructure

- So software there, now we currently are doing the organizational work:
 - 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?

- The infrastructure is accessible via DFN-AAI or via a dedicated DARIAH and TextGrid ('homeless')-Account.
- Ready for EduGain but not depending on it
- With our Homeless IdPs we are operating productively since years

How to get the urgently needed European humanities federation?

Two ways forward:

- A) DARIAH IdPs and SPs either participate via the national federations
- B) 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 is currently running to evaluate these options

At the Humanities federation Workshop with DASISH partners October 17/18, 2013 in Cologne, we decided to go for A for now and have B as Plan B



How to get into DFN-AAI

- DFN is very open for this
- Homeless IdP needs a documented policy
 - Policy and practice is almost in place
 - Contractual work is also on its way
- SPs just needed to sign the contracts
 - This was finalized this month

Policy for Homeless IdP

- DARIAH-DE and TextGrid (which is migrating to DARIAH-AAI) have together about 1800 users in Homeless IdPs.
- They came into the homeless IdPs under an implicit policy
- Now we have an explicit policy and are in the process of purging the accounts
- This is currently being done while merging the TextGrid homeless IdP with the DARIAH IdP

Policy for Homeless IdP: Registration

- If the requester has an email address from the scientific community (e.g. xxx@uni-yyy.de)
 - ok, no questions asked
- If email address is from a commercial provider:
 - Ask to use another email address
 - If this is not possible we have two problems:
 - A) Is the person the one she claims to be
 - B) Does the person belong to the scientific community?

Policy for Homeless IdP: Registration

- A) Identity vetting
 - Either the person has to show the ID card to a DARIAH trusted person, e.g. at a conference (simple DARIAH registration authority)
 - Or the person sends a copy of the ID card
 - Or use the German postident

Policy for Homeless IdP: Registration

- B) Part of the scientific community
 - The person has to show that she is doing research, even without being part of a research organization
 - This could be done by sending a list of publications or by a confirmation of a research project

Policy for Homeless IdP: Data update

- We also have to prove that the status has not changed
- Automated mail to the users
 - If they mails to research email addresses return with „user unknown“ we delete them all others stay in Homeless IdP
 - The users with commercial email addresses have to actively write back and state that they are still part of the scientific community

Data Protection Code of Conduct

- We want to support CoC
- We are part of the CoC uptake activity
 - DARIAH has already signed the support letter
 - Configuration of DARIAH SPs is on its way

Sustainability

- There is the strong will to make DARIAH a sustainable infrastructure
- One of the 6 Project Cluster of DARIAH-DE is DARIAH eHumanities Infrastructure Service Unit (DeISU)
 - We will be working on organizational model and business model
 - That service unit could well also operate AAI services

Summary

- DARIAH has a productive solution based on homeless-IdP and attribute authority
- Distributed user and privilege administration
- Roadmap for a sustainable service unit
- Policies that allow for integration into DFN-AAI and thus into eduGain

Thank you for listening!

Questions?

Comments?