

DARIAH Update

8th FIM4R Workshop

CERN, Geneva, 3.-4. February 2015

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 (**ERIC is in place by now**)

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 shows 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

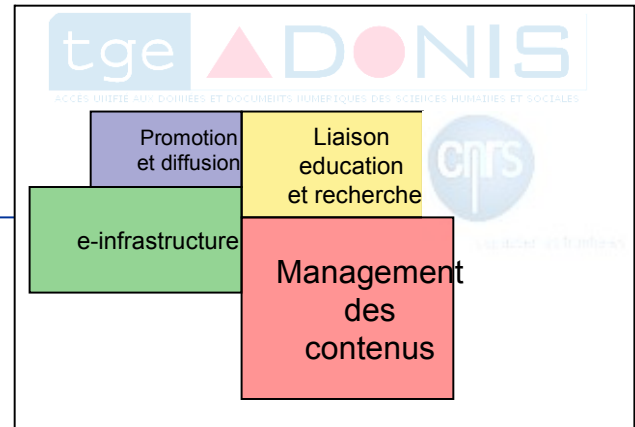
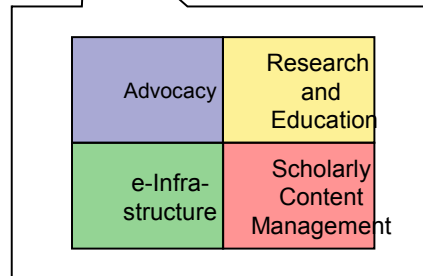
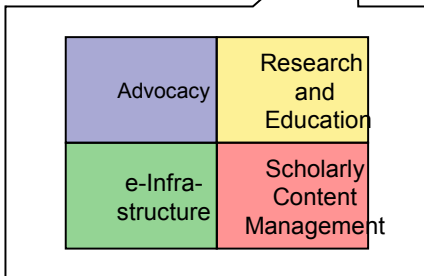
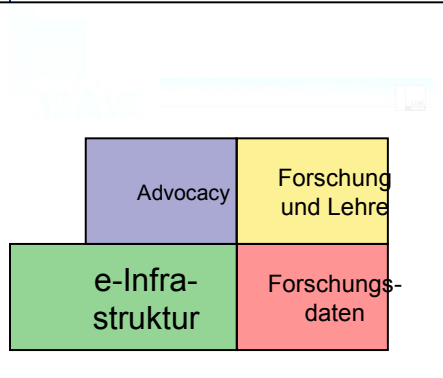
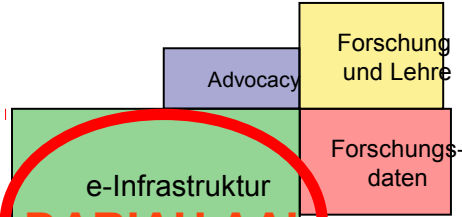
VCC Scholarly
Content
Management

e-Infrastruktur
DARIAH AAI

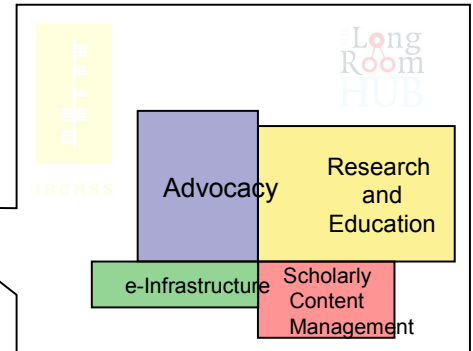
DARIAH-DE

DARIAH-AT

DARIAH-DE



DARIAH-FR



DARIAH-IE

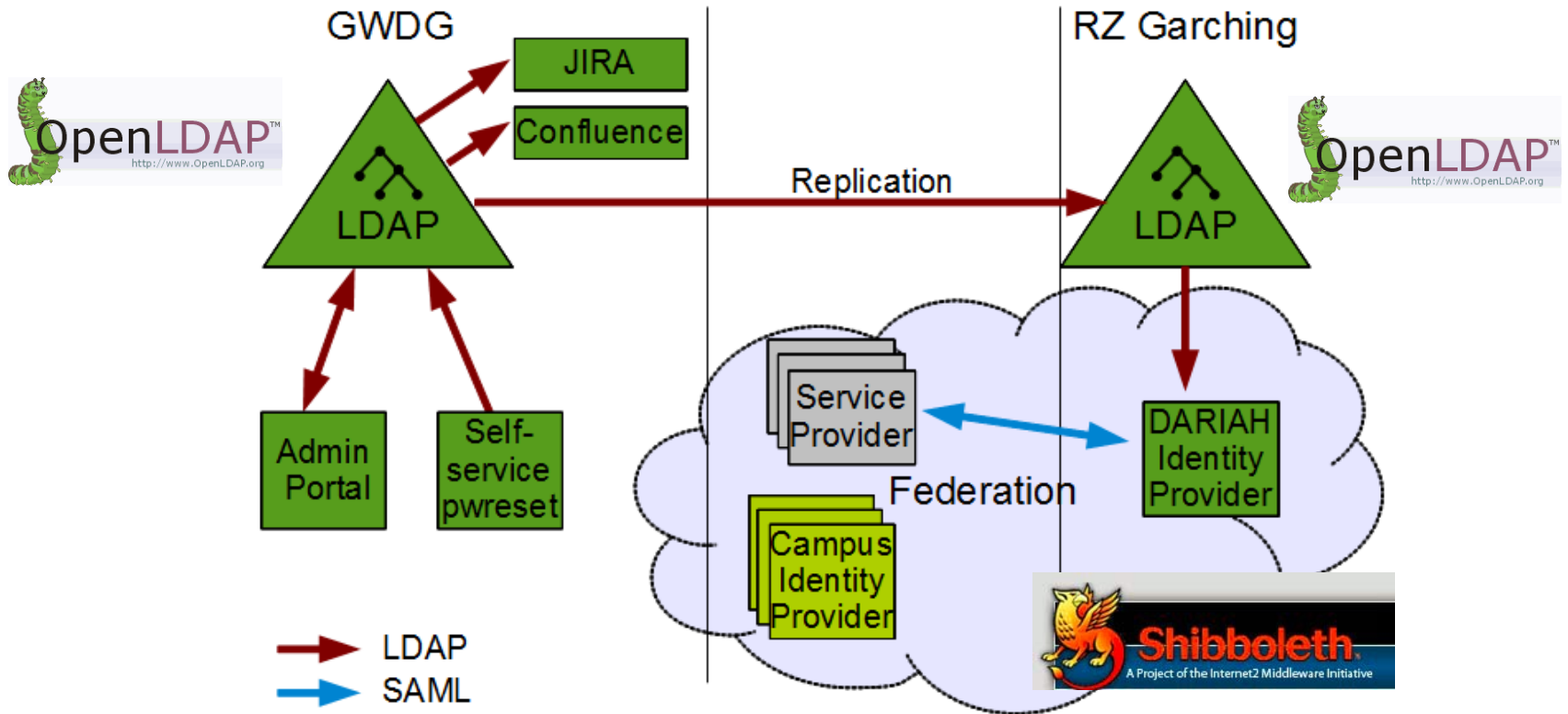
DAASI
International

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

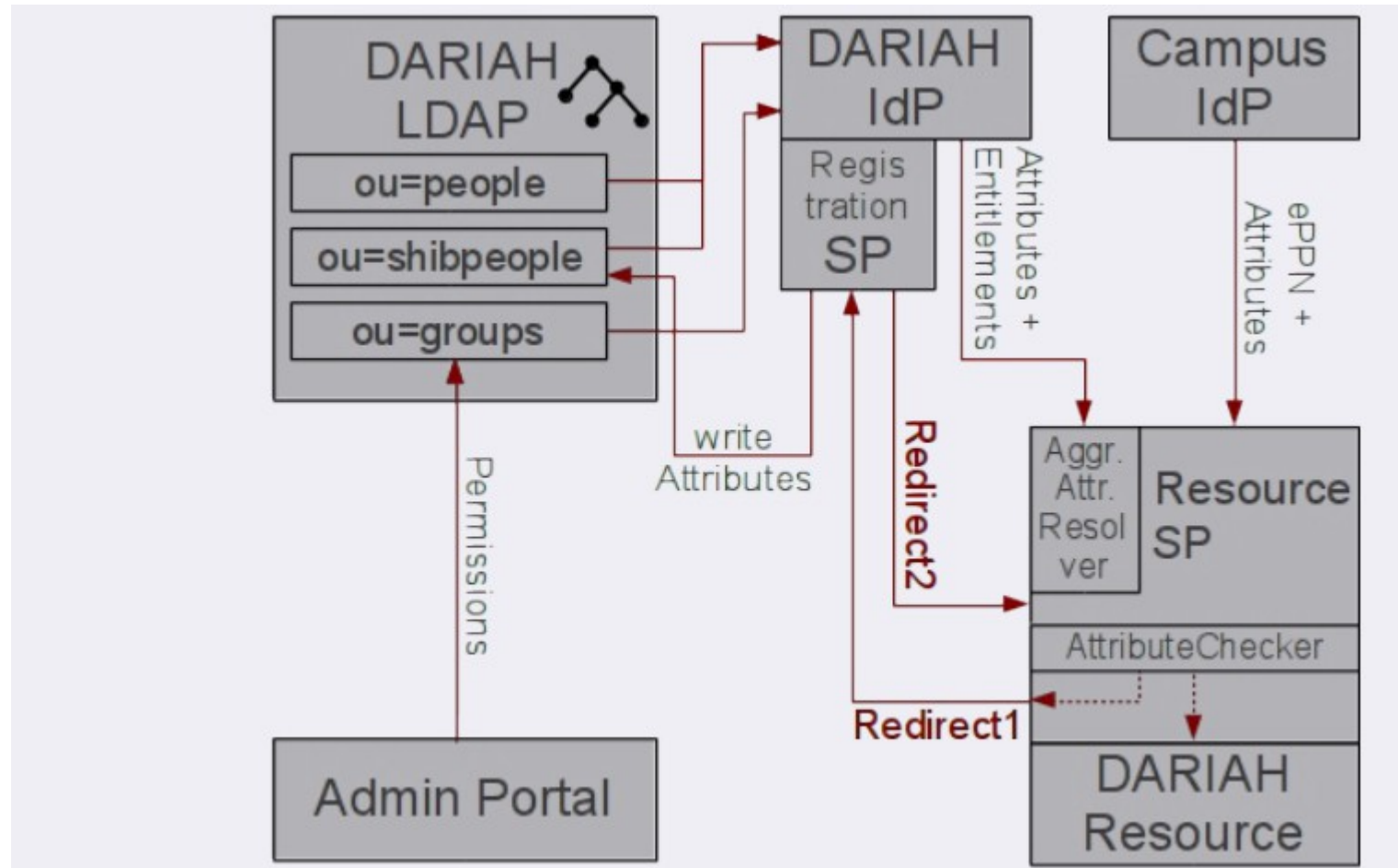


DARIAH Authorization

- 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)

Now, since we have AARC, there is new hope!



Current figures

- We currently have **2380** Users
- Only very few log in via their home IdP
 - It's easier (and sort of familiar) to create a new DARIAH account
- We currently have **210** different user groups
 - Every project usually uses three or four privileged groups (thus ca. **60** projects):
 - X-users
 - X-contributors
 - [X-developoppers]
 - X-admins

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

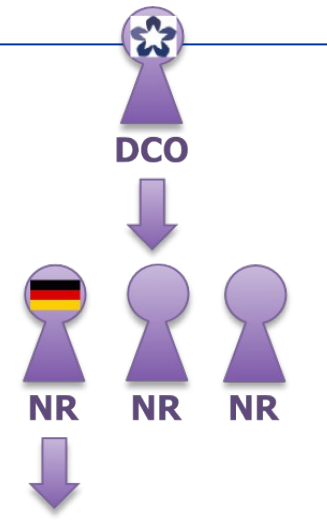
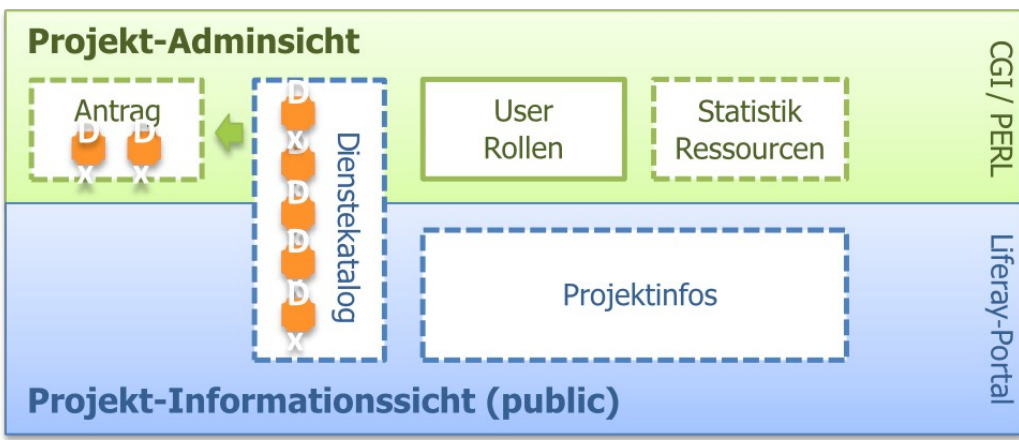
D5

Geo-Browser

Dx

...

Forscherdienste



New Features of User Management

- The Web-based administration and self-service interfaces have been improved, e.g.
 - Better password forgotten processes
 - Completed role based administration
 - Concept of initial group is implemented
- Since the administration interface is actually used, new requirements pop up quite often
 - Our development ticket system is full with such requests, such as bulk-upload from CSV file
 - So development will keep us busy

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 **many** 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

We decided to go for A for now and have B as Plan B

- **All German SPs are integrated in the productive DFN-AAI**
- **The homeless IdP also is, possible because of the enforcement of a policy for entering only members of the research community**



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?

Data Protection Code of Conduct

- DARIAH fully supports CoCo
 - We are part of the CoC uptake activity
 - DARIAH officials have signed the support letter
 - **All productive DARIAH SPs have the entity category configured**

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
- **AARC gives hope that current challenges can be met**

Thank you for listening!

Questions?

Comments?