DARIAH WG FIM4D

2nd Meeting, Berlin, April 26, 2017

Agenda

- Intro to the WG
- Intro to DARIAH-AAI
- Statusreports on the single pilot projects
  - EU: C/D Course registry
  - EU: D Inkind Contribution Management Tool
  - BE: VRE.SI
  - AT: Vocabulary Repository
  - FR: NAKALA service
  - LU: cvce.eu – Digital Toolbox

FIM4D WG

- Federated Identity Management for DARIAH
- chairs: Peter Gietz and Lars Wienecke
- Scope:

  1. Introduce the DARIAH AAI to all DARIAH countries

  2. Enhance the number of DARIAH services accessible via the AAI

  3. Guide on how to „shibbolize“ services

  4. Pilots in AT, BE, FR., LU, NL, the DARIAH EU course registry and DARIAH EU In-kind contribution tool, DARIAH-EGI WG

FIM4D WG news

- We already had two workshops for Service Providers, but with little participation from DARIAH SPs :-(
- Basically the working group means free consulting on how your service be part of the interoperable world
- The EGI DARIAH CC is now a DARIAH WG now called „Cloud Services for DARIAH“
- VCC1 will define Criteria for the DARIAH Hallmark (Formerly seal of approval), and AAI integration will be one of the criteria. The FIM4D WG is to specify such AAI criteria. The earlier the better.
- We have a new pilot from DARIAH-ES

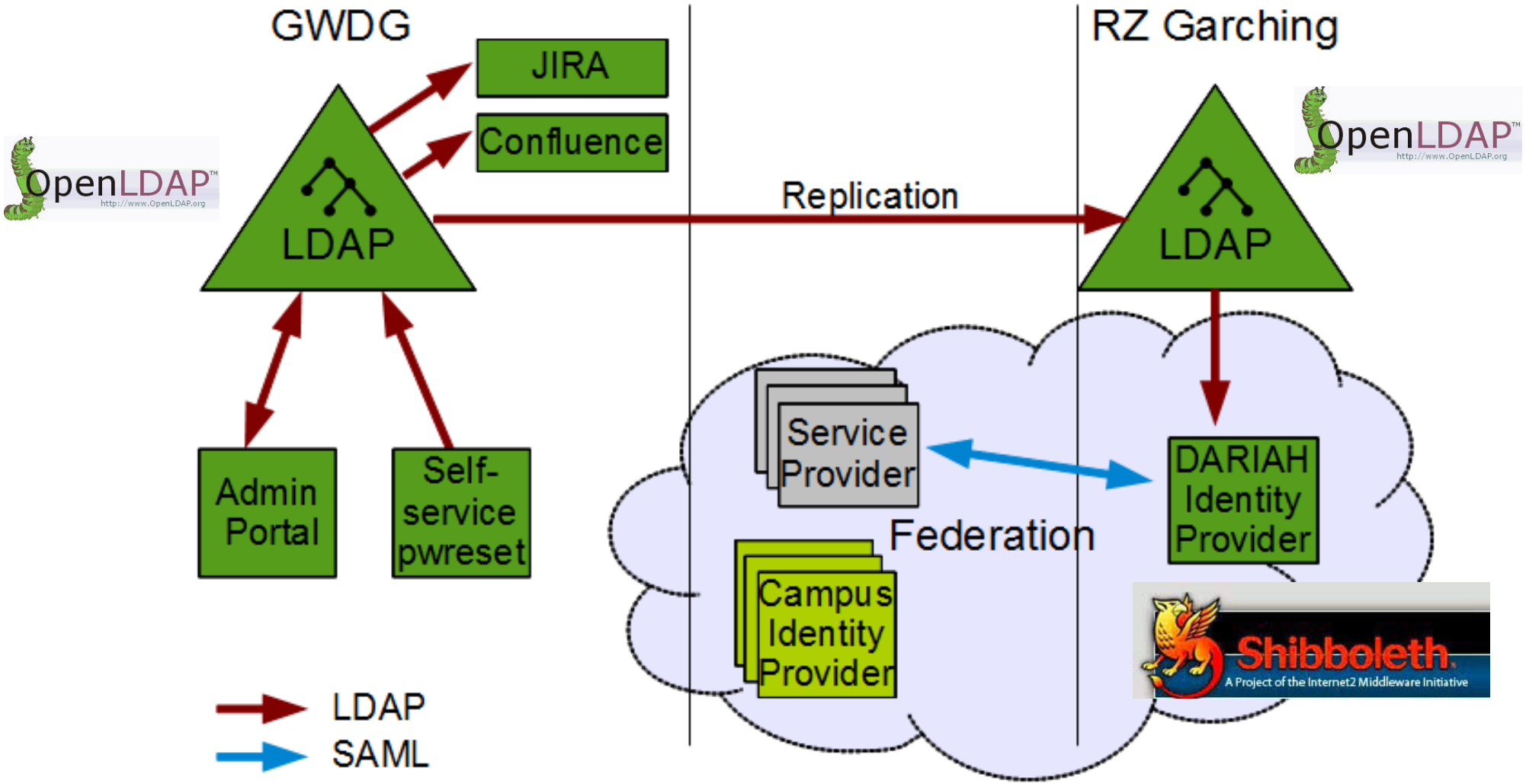
Standards and groups

- Standards like LDAP, SAML, OAuth2 / OpenID Connect are key for interoperability
- Relevant groups that are involved:
  - NRNs, GEANT, eduGain, single campus user managements
  - Research infrastructures (e.g. DARIAH, CLARIN, ELIXIR, …)
  - Research Infrastructure provider (e.g. EGI, EUDat, etc.)
  - FIM4R and FIM SIG in RDA
  - EU Project AARC brings all these groups together
- An interoperable world already exists, and DARIAH is part of it
  - Based on the DARIAH-AAI developed in DARIAH-DE that can reuse campus accounts, but also operates a DARIAH user management for the „homeless“ (and there are a lot of reasons to be homeless ...)

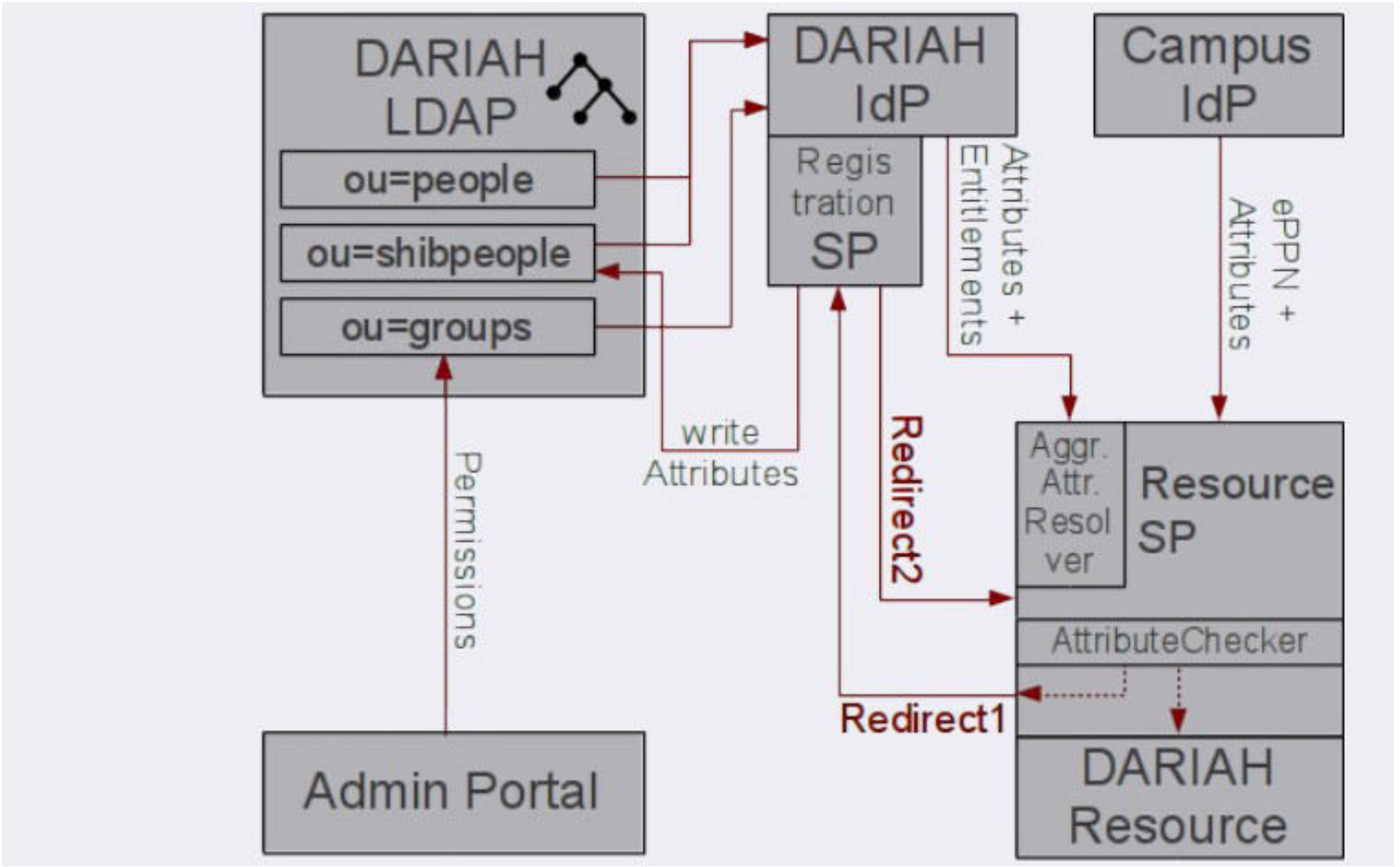
DARIAH AAI Practice

- Current AAI set-up: a first version of an AA infrastructure has been running productively since many years
- based on two standards:
  - LDAP (Lightweight Directory Acess 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

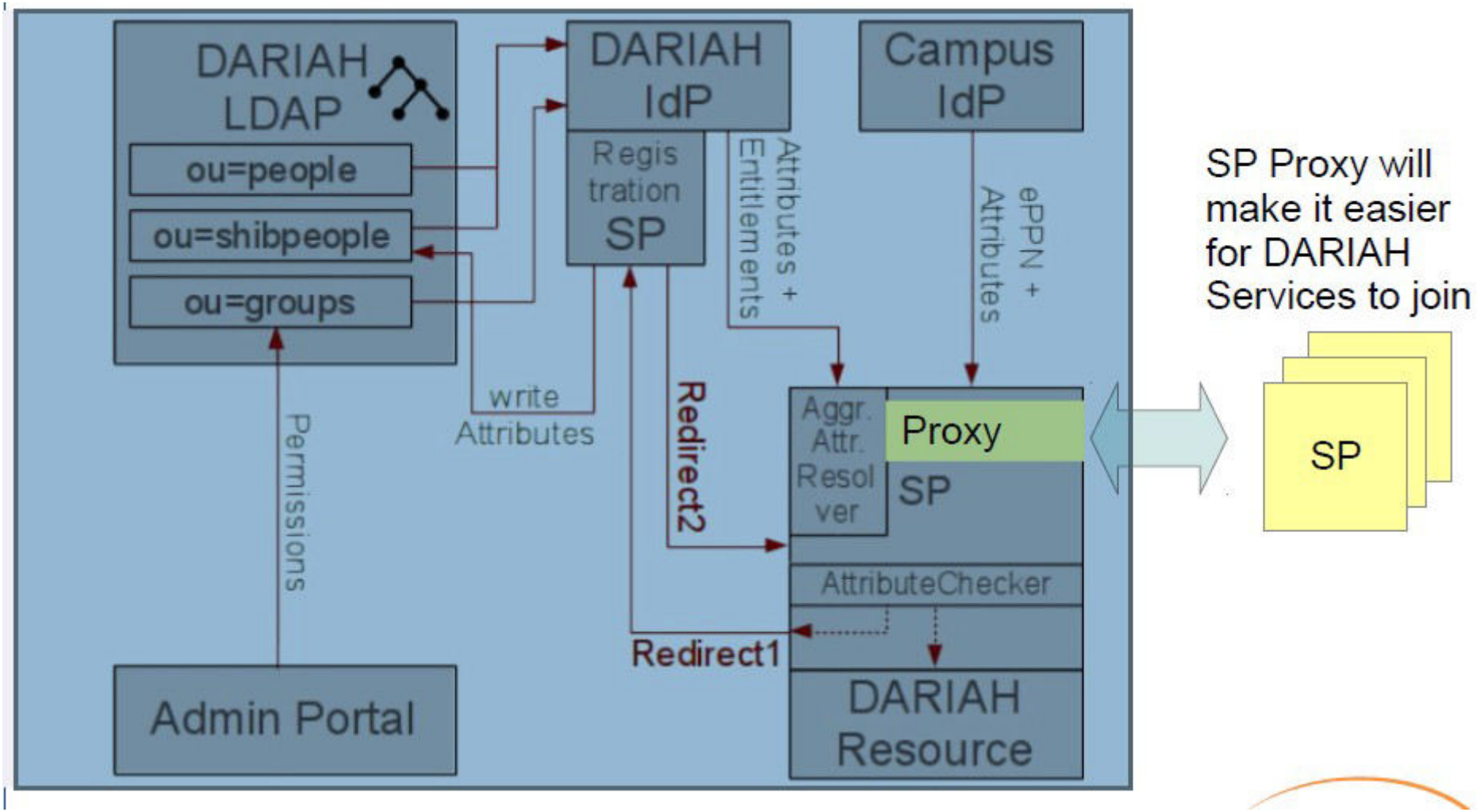
DARIAH AAI Setup



DARIAH VO management



DARIAH VO management NG



Current Challenges

- 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)
- Fine grained access control on file level , observed within a data replication federation (= non web SSO)

Current figures

- We currently have > 3700 DARIAH Users (still growing, 600 more than November 2015)
- Still most do not log in via their home IdP
  - It's easier (and sort of familiar) to create a new DARIAH account
- We currently have > 280 different user groups (10 more)
- Every project usually uses three or four priviledge groups, thus ca. 80 projects (5 more)
  - X-users, X-contributors, [X-developpers], X-admins

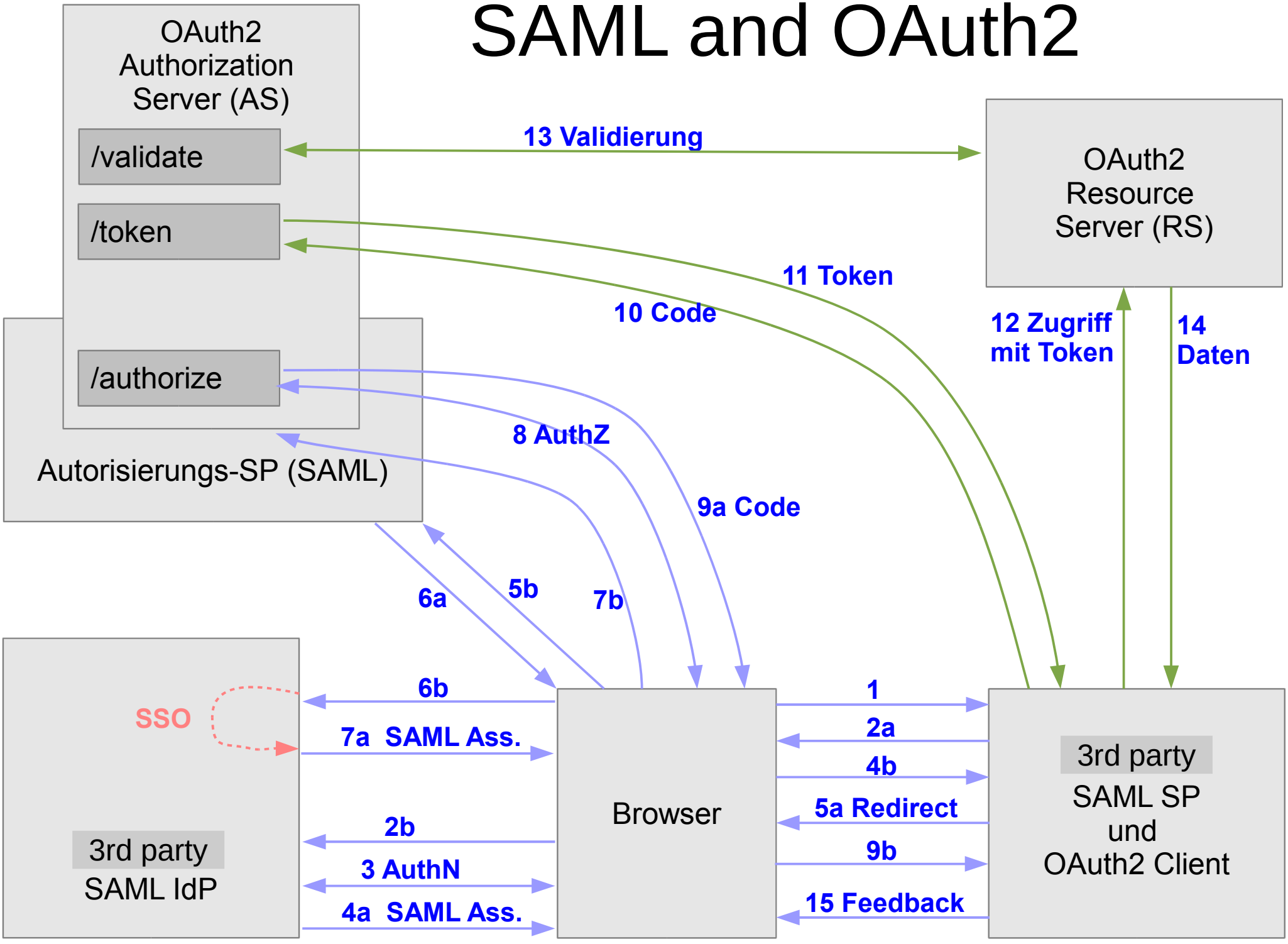
How to make this a European-wide Infrastructure

- 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
- Production ready Administration interface is there currently operated at DARIAH-DE CO in Tübingen
- Who will take over roles?

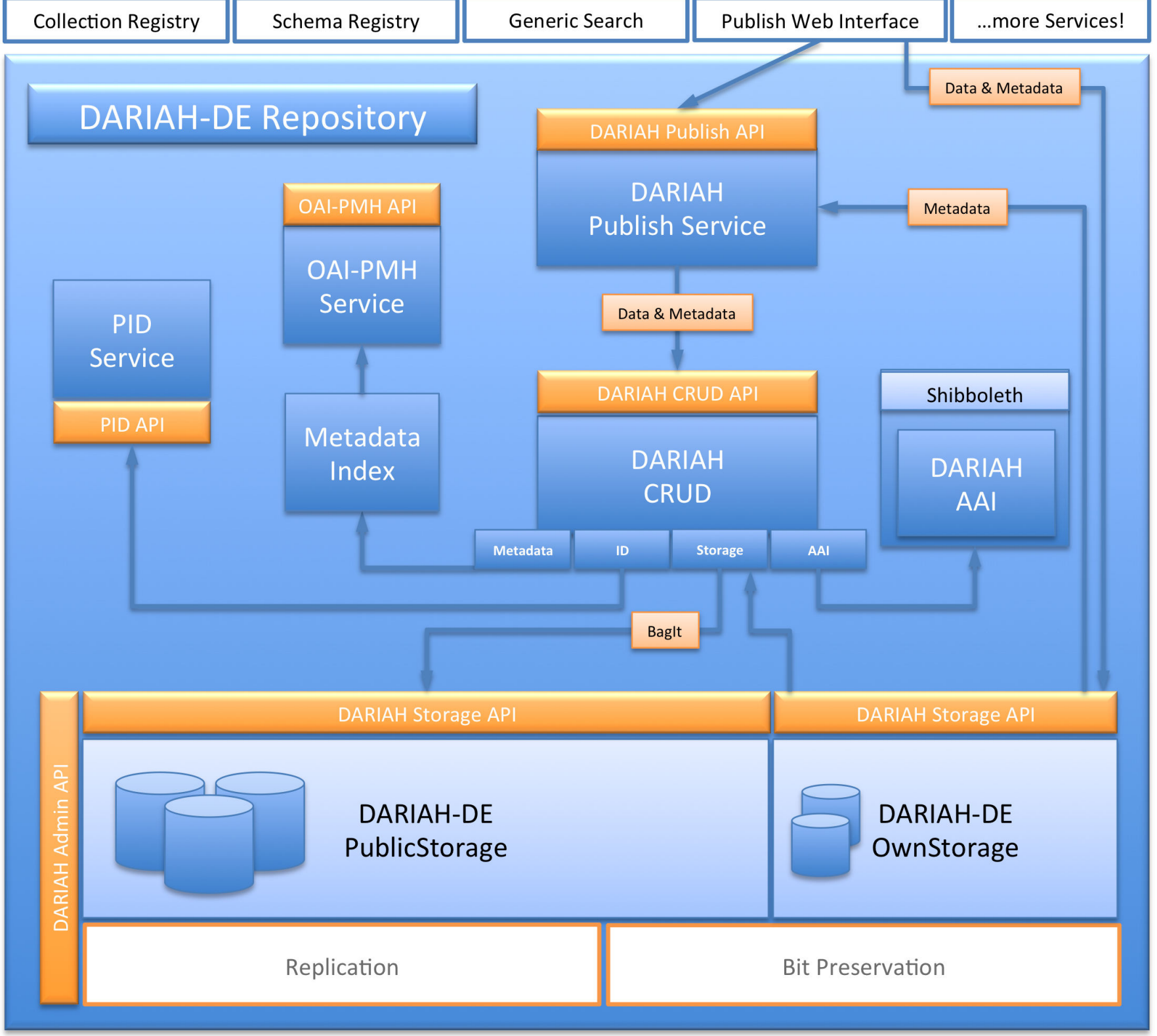
How to make this a European-wide Infrastructure

- The Web-based administration and self-service interfaces have been improved, e.g.
  - Distributed user management
  - 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, there is slow but continous ongoing development work

SAML and OAuth2

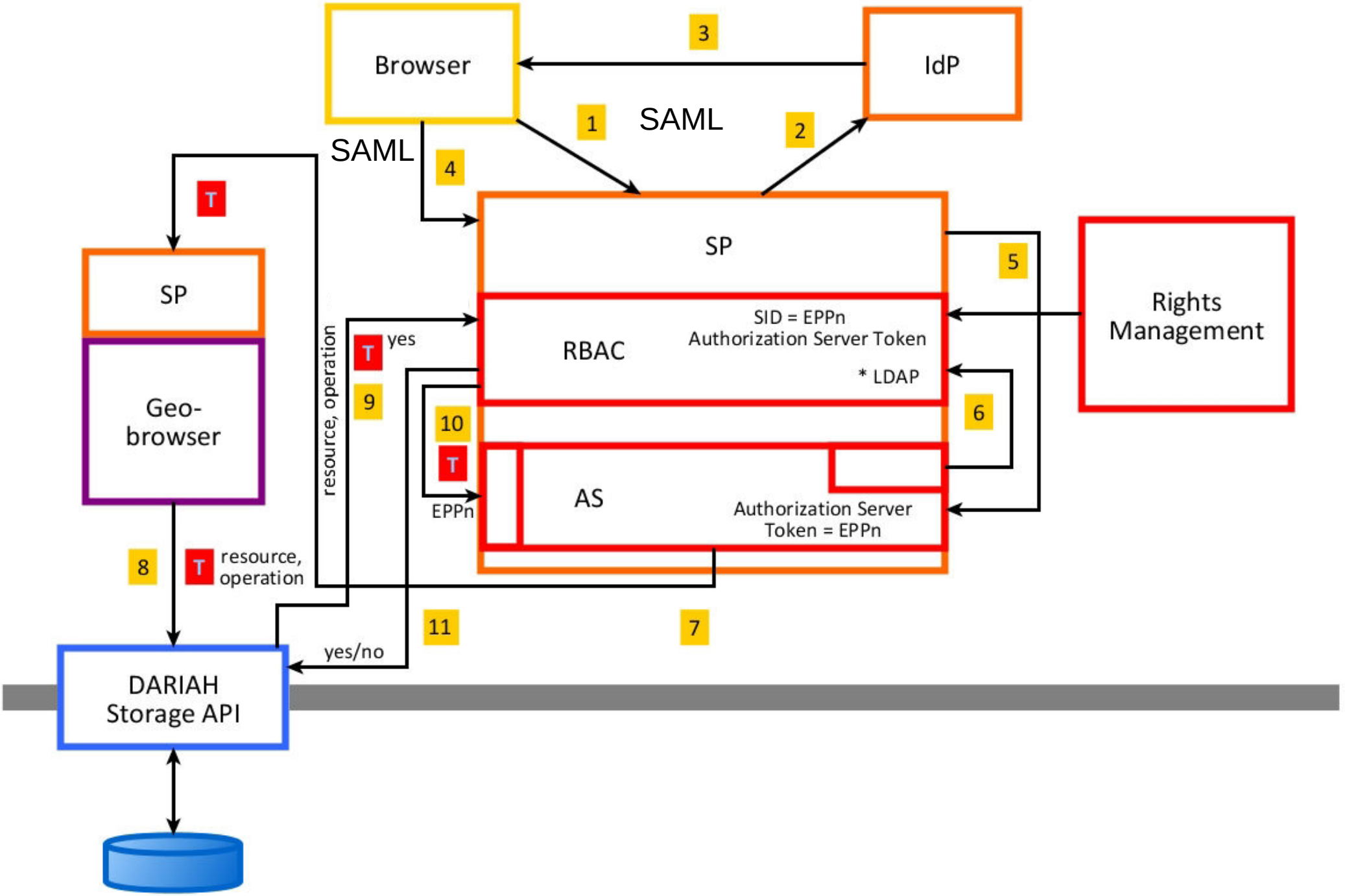


- 1-4: Normal SAML AuthN flow
- 5: 3rd party SP redirects to Authorization endpoint (AS), which is protected by an SAML SP. HTTP Get contains
  - 1. Scope (z.B. “read”, “write”, etc),
  - 2. Client ID,
  - 3. Redirect URI,
  - 4. Identitfier for den Status
- 6-7: Re-authN in Non-SSO case
- 8: oAuth2 AuthZ
- 9: redirect to the 3rd party SP, including authorization code
- 10-11: retrieval of the actual OAuth2 access token by showing the authorization code
- 12: request resource with access token
- 13: validation of the access Code
- 14-15: actually get the data and give back a success message to user
- 16-17: reauthorization if access token is not valid any more


Collection Registry
Schema Registry
Generic Search
Publish Web Interface
...more Services!
DARIAH-DE Repository
Data & Metadata
DARIAH Publish API
DARIAH Publish Service
Metadata
OAI-PMH API
OAI-PMH Service
PID Service
PID API
Data & Metadata
Metadata Index
DARIAH CRUD API
DARIAH CRUD
Shibboleth
DARIAH AAI
Metadata
ID
Storage
AAI
BagIt
DARIAH Admin API
DARIAH Storage API
DARIAH Storage API
DARIAH-DE PublicStorage
DARIAH-DE OwnStorage
Replication
Bit Preservation

Central Policy Decision Point

- If more than one system trusts an access policy it makes sence to have a central policy decision point
  - Manage the access rules only once
  - Policy Enforcement Point
    - doesn‘t have to deal with managing and evaluating access policy
    - Just needs to get access decisions from the PDP via simple and standardized protocols
- DARIAH uses the RBAC compliant open-source PDP „didmos Decision Point“


Browser
IdP
3
1
SAML
2
SAML
4
T
SP
SP
5
Rights Management
SID = EPPn
Authorization Server Token
RBAC
* LDAP
6
Geo-browser
T
yes
9
10
T
EPPn
AS
Authorization Server Token = EPPn
resource, operation
8
T
resource, operation
11
7
yes/no
DARIAH Storage API

PDP flow

- O: service has a session with the user
- 1: If service needs to access protected resources user gets redirected to the SP protecting the PDP
- 2: SP initiates normal SAML AuthN flow
- 3-4: SAML assertion gets to the PDP
- 5: OAuth2 token for user which is valid in Storage API
- 6: PDP creates a session connected with OAuth2 token and user ID (EPPN), and with the roles of that ID
- 7-9 OAuth2/SAML flow
- PDP checks token validity
- PDP sends policy decision to storage API

Summary

- DARIAH has a productive solution based on homeless-IdP and attribute authority
- Distributed user and priviledge administration
- Roadmap for a sustainable service unit
- Policies that allow for integration into DFN-AAI and thus into eduGain
- DARIAH is actively co-operating with AARC