Call for action on federated identity

Many of the EU’s emerging research infrastructures need a solid technical layer to ensure the authentication and authorization of their users. Especially distributed infrastructures, like CLARIN and DARIAH, depend on a well-functioning federated identity management, as their users are spread all over the European academic landscape.

In most of the EU countries SAML-based academic trust federations have been or are being established. For Germany e.g. the DFN-AAI is connecting academic institutions and services. At supranational level the eduGAIN project aims at cross-linking the national identity federations.

The general goal of federated identity management is to provide a service (hosted at a so-called Service Providers) to a multitude of trusted users (who authenticate by entering a username and password to the Identity Provider of their own university). This approach provides advantages in terms of:

- **Usability.** Only one user name and password are needed. Moreover, when accessing different services the user needs to authenticate only once (the single-sign-on experience).

- **Security.** The password does not leave the home organisation. Instead the only information passed on are the user attributes: information like the user name and the email address.

Practice has shown that there are however some factors impeding the acceptance and use of federated identity management. The main problem is that Identity Providers are worried or even unwilling to pass on personal attributes to Service Providers. Consequently the Service Providers from the research infrastructures cannot deliver a satisfying service due to the lack of personal information.

Unlike most commercial Service Providers (e.g. publishers) who usually only need anonymous attributes, research infrastructures require personal information. This information is necessary to enable collaboration in scholarly contexts, cf. [2]:

“As is the standard practice in the higher education and research world, collaboration on these sites involves knowing who the collaborators are: name, email, institutional affiliation.”

The information requested encompasses the following attributes that are needed to enable collaboration within both research infrastructures and to ensure a good end-user experience:
<table>
<thead>
<tr>
<th>Attribute name</th>
<th>Description</th>
<th>Example</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>eduPersonPrincipalName</td>
<td>Unique user identifier</td>
<td><a href="mailto:jdoe@uni-foo.de">jdoe@uni-foo.de</a></td>
<td>Unambiguous identification of a user(^1)</td>
</tr>
<tr>
<td><strong>cn (common name)</strong></td>
<td>Full user name</td>
<td>John Doe</td>
<td>Personalization of the user account</td>
</tr>
<tr>
<td><strong>mail</strong></td>
<td>Email address</td>
<td><a href="mailto:john.doe@uni-foo.de">john.doe@uni-foo.de</a></td>
<td>Communication with other researchers and the infrastructure's operating team</td>
</tr>
<tr>
<td>o (organizationName) or schacHomeOrganization</td>
<td>Name or Domain name of the home organization</td>
<td>University of Foo or uni-foo.de</td>
<td>Personalization of the user account</td>
</tr>
<tr>
<td>eduPersonScopedAffiliation</td>
<td>User's affiliation within home organization</td>
<td><a href="mailto:member@uni-foo.de">member@uni-foo.de</a></td>
<td>Personalization of the user account</td>
</tr>
</tbody>
</table>

It is generally understood that releasing personal attributes to academic Service Providers is less troublesome than releasing them to commercial ones. Therefore CLARIN and DARIAH hope that declaring that the information acquired will only be used for user-requested academic services will address the privacy concerns of the Identity Providers.

This approach fits well with similar initiatives taken by eduGAIN \(^1\), the InCommon federation in the US \(^2\) and the proposed new EU regulation on data protection \(^3\).

CLARIN and DARIAH declare adherence to eduGAIN’s **Data Protection Code of Conduct** \(^4\) as a sign of goodwill towards the Identity Providers. They hope that this signal will encourage national identity federations to approach their Identity Providers with the request to pass on the relevant user attributes.

At the same time Identity Providers should also be invited to support the **SAML Enhanced Client or Proxy Profile**, which would ensure that new technologies provided by the research infrastructures, such as authentication of web services, could be implemented \(^5\).

We strongly hope that this approach will lead to the right balance between user data protection at one hand and bringing a useful and usable infrastructure to a broad academic user group at the other hand. After all, all actors involved serve the same goal: providing the scientific community with the right tools to further research.

**References**

\(^1\) Since the infrastructure can be accessed via many services, a pseudonymous identifier like eduPersonTargetedId doesn’t work because they are only unambiguous in respect of one particular Service Provider