Identifier Attributes

Introduction
Web services often need to re-identify users in order to present them their profile, or some other user-specific information. From the user's perspective, it is desirable that the service receives only the absolutely needed information about them (data minimization principle). For this purpose, there are different identifier attributes around in the federated AAI world. They differ in their degree of anonymity and reusability as well as in other areas. In the following, this page first describes the properties that help distinguishing these identifiers and then the identifiers themselves.

Identifier Attribute Properties

**Uniqueness:** as the name implies an identifier should identify a user without doubt, no two users should have the same identifier.

**Reassignability:** user identifiers can be reassigned to another user after the first user with this identifier has left the organization. Reassigned identifiers can cause access control and traceability problems.

**Opacity:** an identifier that gives no clue (i.e. it only contains random data and no names) about the identity of the user is called opaque.

**Persistency:** an identifier that stays identical over time is called persistent.

**Targetedness:** an identifier that is intended to be used for a single service only and is specific to that service is called targeted.

**Transientness:** a transient identifier stays the same for a login session, but changes when the user logs in again.

Common Identifier Attributes in eduGAIN

The following list describes identifier attributes that are common in eduGAIN, which means that there is a substantial number of Identity Providers that can release these identifier attributes about their users.

**eduPersonPrincipalName (ePPN) - urn:oid:1.3.6.1.4.1.5923.1.1.1.6**

The eduPersonPrincipalName looks like an e-mail-address with a user part that must be unique followed by an @ sign and the domain of the organization to which the user belongs. The ePPN is designed to be human readable and often contains the users first or last name. Some organisations also use the user's login name (e.g. uid attribute) before the @. The organization domain at the end of the identifier defines a name space (the so-called "scope"), so an organization can assure itself the uniqueness and possible reassignability of the identifier.

**eduPersonTargetedID (ePTID) - urn:oid:1.3.6.1.4.1.5923.1.1.1.10**

The eduPersonTargetedID consists of a privacy preserving, opaque identifier part and a source and target identifier. The target identifier can point to a set of services or a single service. The eduPersonTargetedID is a rather lengthy identifier. Per the SAML format definition, the identifier portion MUST NOT exceed 256 characters, and the source and audience URI values MUST NOT exceed 1024 characters. Due to this enormous (theoretical) length and opaque contents, the ePTID is not intended to be human-readable friendly, but preserves privacy quite well. This also implies that this identifier should not be shown on web pages as it would look quite ugly.

Additionally to sending the ePTID as a SAML attribute in the AttributeStatement of the SAML assertion (which is the only possible way for all other identifiers mentioned on this page), most SAML implementations (besides Shibboleth) be send it as a SAML persistent NameID in the subject of the SAML assertion. Shibboleth IdP can send the value of the eduPersonTargetedID as SAML attribute as well as a SAML NameID in the subject. The Shibboleth developer as well as the SAML2int profile recommend in the long term to send this identifier only in the SAML2 subject.

**eduPersonUniqueid (ePUID) - urn:oid:1.3.6.1.4.1.5923.1.1.1.13**

The eduPersonUniqueid also has the form of an e-mail-address with a scope like the principal name. In difference to the ePPN it is by definition not reassignable and usually opaque, but not necessarily. The ID portion before the @ symbol is restricted to the alphanumeric characters [a-zA-Z][0-9], other characters are not allowed. The eduPersonUniqueid is limited in length and hence better suited for interoperability than eduPersonTargetedID, but it does not preserve privacy that well since it ist he same across different services and user actions can be tracked. This attribute only was introduced in the eduPerson schema in 2013. Therefore, not too many Identity Providers have implemented this attribute and therefore can release it about their users.

**mail - urn:oid:0.9.2342.19200300.100.1.3**
The mail address can sometimes be found as an identifier attribute. Mostly, this is done out of convenience of the SP-developer/administrator, but it has some disadvantages. First, the specification allows multiple values for the attribute (although rarely used in practice), which makes its usage technically unsound. Then, the user's real name and first name are often included which is unwanted for privacy-concerns. Also, given that some users have several federated accounts it can happen that all the users account have the same email address attribute, which in some cases is beneficial for the purpose but might be a problem in other cases (because the email attribute is not unique in such a case). As experience has shown, email addresses are relatively often subject to changes (i.e. because the faculty domain name changes or similar), which then deteriorates the persistence of this attribute as identifier attribute.

## Properties Overview

The following table provides a brief overview about the different identifier attributes and their characteristics as well as usage.

<table>
<thead>
<tr>
<th>Property/Identifier</th>
<th>ePPN</th>
<th>ePTID</th>
<th>ePUID</th>
<th>mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Non-reassignable</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opaque</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistent</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Targeted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transient</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability in eduGAIN</td>
<td>Very good</td>
<td>Good</td>
<td>Poor</td>
<td>Very good</td>
</tr>
</tbody>
</table>