Attributes and namespaces in SURFconext

Project	: Collaboration Infrastructure 2013
Project year	: 2013
Project manager	: Remco Poortinga-van Wijnen & François Kooman (SURFnet)
Author(s)	: Bob Hulsebosch & Martijn Oostdijk (Novay)
Due date	: September 2013
Version	: v1.0

Summary

The use of attributes in SURFconext, SURFnet's identity federation, is likely to increase in the near future due to developments like new license models for cloud providers, collaborative organisations, fine-grained authorisation, and authentication levels of assurance.

In order to prevent proliferation of local-defined attributes and namespaces that potentially could hamper consistent and efficient attribute exchange in SURFconext measures may need to be taken. This deliverable analyses the current use of attributes and attribute namespaces and provides recommendations for future activities in these areas.

Regarding new license models several approaches are possible. Either the identity provider or the license provider provides the attributes. Entitlement or isMemberOf attributes could be used. It is recommended to use a license ID attribute for scoping purposes. This requires the creation of an OID-based namespace for the registration of its values and the entitlements associated to them based on e.g. group membership.

Regarding collaborative organisations it is recommended to better support the alignment of their attributes with those of the service provider via meta-data. Furthermore, a namespace for collaborative organisations should be created under the flag of Mace.

Regarding fine-grained authorisation it is recommended to use entitlements and establish an OID-based namespace for their values. Service providers should register and publish (via e.g. SAML metadata) the entitlement values they expect to receive for authorisation purposes. Identity or third party attribute providers should provide the entitlement values in a scoped manner, e.g. urn-based.

Regarding the authentication levels of assurance it is recommended to register eventual SURFconext assurance profiles at the global IANA registry that is currently being setup.

Colophon

Program:	GigaPort3
Workpackage:	Collaboration Infrastructure 2013
Activity:	401 Maturing OpenConext/SURFconext
Deliverable:	2013-401-L
Access rights:	Public
External party:	Novay, www.novay.nl

This project was made possible by the support of SURF, the collaborative organisation for higher education institutes and research institutes aimed at breakthrough innovations in ICT. More information on SURF is available on the website www.surf.nl.

Contents

1 Introduction

Much of the power of identity federations such as SURFconext is derived from the economies of scale accomplished by large numbers of identity and service providers speaking a lingua franca. Attributes are the language in which access control and release policies are written and are the piece of the federated infrastructure for which avoiding unnecessary proliferation of names is most important. Standards bodies have traditionally defined common attribute names and semantics (e.g. SCHAC, eduPerson, etc.) that are used in identity federations.

However, the use of attributes in SURFconext is likely to increase in the near future due to a number of developments:

- There is an increasing need to abandon the current campus license model for service usage and to switch to a more flexible and customised license model for smaller user groups. Such a model is typically required for cloud services. Attributes will be used for the transmission of license and user group information between SURFmarket, SURFconext, and (cloud) service providers;
- The desire to support collaborative organisations via SURFconext results in the transmission of e.g. collaborative organisation specific attributes such as self-asserted and proprietary role information for authorisation purposes in the context of the collaboration;
- The transmission of service provider specific attributes required for fine-grained authorisation enforcement;
- The requirement of service provider to differentiate between the strength of various authentication solutions in terms of levels of assurance (LoA).

Some of these attributes are expected to be provided by the identity providers in SURFconext. For attributes regarding license or group information there is an emergence of so-called third party attribute providers. These attribute providers are authoritative for certain attributes that the identity provider typically cannot or will not provide.

Choosing future-proof and unambiguous names and value ranges for the attributes in these use cases is currently, as far as SURFnet can estimate, difficult if not impossible via the existing and standardized attribute schemas like eduPerson and SCHAC. Yet, to prevent potential proliferation of local defined attributes and their namespaces action may be needed to harmonize and control the attribute availability and semantics and to ensure their uniqueness.

1.1 Objectives and goals

The goal of this deliverable is to:

- Investigate the current use of attributes and their values and namespaces within services like SURFconext and parties like SURFmarket and SURFnet.
- Provide recommendations regarding their future use whilst taking into account the emerging use cases for attribute management and the capabilities identity and service providers have for processing or delivery of attributes.

1.2 Approach

The research approach consisted of conducting a number of interviews with SURFnet employees and an external expert from TERENA. During the interviews, challenges regarding the current and future use of attributes and their namespaces and solution directions were discussed. Desk research was used to further analyse the challenges and come to recommendations.

1.3 Reading guide

Section 2 starts with an overview of existing attribute schema, namespaces and protocol profiles. Section 3 describes the relevant use cases that require specific attributes that potentially do not fit within the current standardised definitions and values. For each use case, it describes current practice and proposes solution directions for future use. Section 4 concludes with the major findings and recommendations.

2 Overview of existing attribute schemas and namespaces

This section describes the current state of the art of attribute schemas and namespaces. Moreover, it also provides an overview of the attribute definitions of several common protocols that are used for attribute exchange in identity federations.

2.1 Attributes

Attribute is defined as a quality or characteristic of a person, place or thing. For instance, an identity attribute, as asserted by an Identity Provider (IdP) in a federation, consists of an attribute name (such as "eduPersonScopedAffiliation"), and a value (such as "student@example.edu"). Values are often simple numbers or strings but can also be complex, such as XML-structured data, if desired. Many attributes can be sent at one time. Attributes can express things specific to a particular subject, such as a unique identifier, or things shared among many subjects, such as group membership or affiliation. In many scenarios identity attributes are very useful to Service Providers (SPs) for access control, personalization, and other purposes.

2.2 Schemas

A schema describes the syntax and usage constraints (and, to some degree, the semantics[1]) for attributes exchanged, for example, between the IdPs and SPs in a federation. As such, a schema provides a common understanding of the attributes between the exchanging parties and facilitates interoperability.

A number of schemas exist that are relevant in this context. They are illustrated in Figure 1.

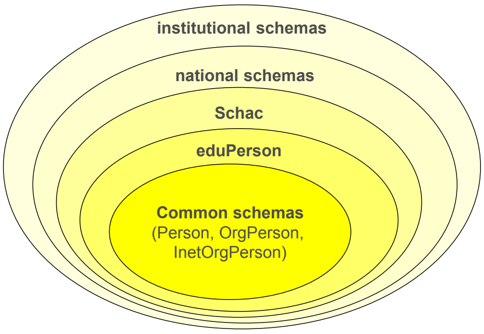


Figure 1: The schema onion [source SCHAC implementation and related issues presentation by Mikael Linden at EuroCAMP 18-19.10.2006].

[1] In the form of a Natural Language description and/or comments.

2.2.1 Common schemas

At the centre of Figure 1 RFC4519[2] and RFC4924 (and IETF documents referenced from these) define common identity attributes that can be used to describe users in general.

2.2.2 eduPerson

eduPerson is an LDAP schema designed to include widely-used person and organizational attributes in higher education and refines the above-mentioned RFCs. It was developed, and is maintained, by the Internet2 MACE-Directories Working Group (MACE-dir), a project of the Internet2 Middleware Initiative[3].

eduPerson defines 11 attributes that are used for authorisation or identification purposes[4]:

- eduPersonAffiliation
- eduPersonNickname
- eduPersonOrgDN
- eduPersonOrgUnitDN
- eduPersonPrimaryAffiliation
- eduPersonPrincipalName
- eduPersonEntitlement
- eduPersonPrimaryOrgUnitDN
- eduPersonScopedAffiliation
- eduPersonTargetedID
- eduPersonAssurance

Each attribute also has an OID. In eduPerson profile, only the eduPersonPrincipalName and eduPersonScopedAffiliation attributes are scoped, i.e. they are valid in a certain domain (e.g. bob@novay.nl vs staff@novay.nl). All other attributes are not scoped. It is always possible to define attributes which are scoped, but normally such scoping is a private agreement between the IdP and SP. Each value of 'scope' defines a namespace within which the assigned attribute values must be unique. For instance the value of the eduPersonEntitlement attribute could be scoped as follows: urn:mace:uva.nl:service-x:attribute-value. The SP in this case immediately knows that it involves a user of the University of Amsterdam. Moreover, SURFconext knows that the attribute is valid for service-x and can use this information for filtering purposes, i.e. it allows for easy removal of eduPersonEntitlement attribute values of other services.

Attributes from the person, OrgPerson and inetOrgPerson schemas are used as well. Common examples are:

- cn (commonName, included in person);
- displayName (defined in RFC2798, inetOrgPerson);
- givenName (defined in RFC4519, inetOrgPerson);
- mail (defined in RFC4524, inetOrgPerson);
- postalAddress (included in orgPerson);

[2] See http://tools.ietf.org/html/rfc4519 and http://tools.ietf.org/html/rfc4524.
[3] MACE-Dir working group, see http://middleware.internet2.edu/dir/.
[4] eduPerson schema, see http://middleware.internet2.edu/eduperson/docs/internet2-mace-dir-eduperson-201203.html.

- postalCode (included in orgPerson);
- sn (surname, RFC4519, included in person);
- uid (defined in RFC4519, inetOrgPerson);

National and institutional schemas supplement international schemas with national and local specialties. For example, funetEduPerson (Finland), norEduPerson (Norway), and swissEduPerson (Switserland). The funetEduPersonTargetDegree attribute for instance is used to describe the national codes that are maintained by the statistical center of Finland. In this case, a doctor of theology is represented by urn:mace:funet.fi:attribute-def:funetEduPersonTargetDegree:university:311.

2.2.3 Schemata for Academia: SCHAC

Another important schema is the SCHema for ACademia, SCHAC. SCHAC defines and promotes common schemas in the field of higher education to facilitate inter-institutional data exchange. The SCHAC directory schema is not oriented to any particular technology. It defines a set of attributes to describe individuals in academic and research institutes. These definitions assume that other attributes describing individuals are already available and properly coded in agreement with the current standards, such as eduPerson schema. SCHAC is the result of work in the area of attributes coordination carried out within the TERENA Task Force on European Middleware Coordination and Collaboration, TF-EMC2[5].

The SCHAC attributes have been distilled based on the synchronization of attributes used in national schemas of Croatia (hrEdu), Finland (funetEdu), France (supAnn), Norway/Sweden (norEdu), Poland (plEdu), Spain (iris), and Switzerland (swissEdu), and from contributions of other European NRENs and experts worldwide.

TERENA used to manage the SCHAC namespace which was delegated by MACE (i.e. urn:mace:terena.org:schac). In August 2011, the IETF approved the urn:schac namespace for SCHAC. The RFC 6338 describes the procedures and policies governing its use[6]. As of 1 January 2013, the urn:mace:terena.org has been deprecated for all SCHAC purposes. The new namespace urn:schac is the current namespace for SCHAC. The following SCHAC attributes are defined in this namespace[7]:

- urn:schac:homeOrganizationType
- urn:schac:personalUniqueId
- urn:schac:userStatus
- urn:schac:personalPosition
- urn:schac:projectSpecificRole
- urn:schac:personalUniqueCode

These and other SCHAC attributes are described in more detail in SCHAC IAD version 1.3.0[8].

[5] SCHAC, see http://www.terena.org/activities/tf-emc2/schac.html.
[6] RFC 6338, see http://tools.ietf.org/html/rfc6338.
[7] See http://www.terena.org/registry/terena.org/.
[8] SCHAC attributes, see http://www.terena.org/activities/tf-emc2/docs/schac/schac-schema-IAD-1.3.0.pdf.

The idea was for each NREN to maintain a national vocabulary for national extensions of HomeOrganizationType, UniqueCode and UniqueID and optionally to delegate namespaces for institutional vocabularies. TERENA was envisioned to gather links to the national vocabularies and publishes them in http://www.terena.nl/registry/terena.org/schac/. Looking at the corresponding namespaces there are very few national extension defined.

There was in the beginning of 2013 a discussion started on the SCHAC mailing list about the home organization type attribute. This multi-valued attribute can be set by organisations/IdPs to declare their type/category[9]. An organization can define local (federation dependent) values like urn:schac:homeOrganizationType:ch:university. Especially, in an interfederation context it also makes sense to set one of the international values like urn:schac:int:universityHospital.

However, it turned out that there is a need to register additional international values (and some guidelines to assign them) for these attributes as of today only very few values are registered. To determine which international values are needed in addition to the specified and registered, it is useful to first get an overview of what is in use today. Therefore, a REFEDS wiki page was created to gather the organisation types that are used already in federations as attributes[10]. The goal of this quick survey is to get a rough overview of what types/categories are in use already as attributes. The idea was that this could help to find reasonable values for the international values for SCHAC home organisation type. The survey shows that there are several values in use (such as library, hospital, nren, or vo). There is a wish list of about 8 new values including e.g. museum, supercomputing centre and research institution.

2.2.4 Schema and attributes in SURFConext

The table below lists the core attributes that have been defined for use within the SURFconext federation. This list was created in consultation with the connected institutions and is largely based on the inetOrgPerson, EduPerson and SCHAC tables. A number of attributes that are specific to the SURFconext federation and the Netherlands have been added.

Attribute name	Description	Example values
Uid	User ID / login name	Joebloggs, 434943938
Sn	Surname	Bloggs, Smith
givenName	Given name	Joe, Prof H.A.B.
Cn	Full name	Joseph Bloggs
displayName	Display name	Joey
Mail	e-mail address	j.bloggs@rug.nl
eduPersonAffiliation	Affiliation type	Student, employee
eduPersonEntitlement	Entitlement	Depends on SP

[9] homeOrganizationType attribute, see http://www.terena.org/registry/terena.org/schac/homeOrganizationType/index.html.

[10] See https://refeds.terena.org/index.php/SchacHomeOrgType_usage.

eduPersonPrincipleName	Unique name	joebloggs@rug.nl
preferredLanguage	Preferred language	nl, en
schacHomeOrganization	Domain name	tudelft.nl
schacHomeOrganizationType	Type or organisation	urn:mace:terena.org.schac:homeOrganizationType:eu:higherEducationInstitution
nlEduPersonHomeOrganisation (deprecated)	Name of institution	Delft University of Technology, Utrecht University of Applied Sciences
nlEduPersonStudyBranch	ROHO code	52734
nlEduPersonOrgUnit	Department name	Faculty of Humanities, Library
nlStudielinkNummer	Studielink number	S123456789
nlDigitalAuthorIdentifier	DAI number	070014345

2.3 Namespaces

In general, a namespace uniquely identifies a set of names so that there is no ambiguity when objects having different origins but the same names are mixed together. In order to ensure the global uniqueness of URN namespaces, their identifiers (NIDs) are required to be registered with the IANA[11]. Registered namespaces may be "formal" or "informal". An exception to the registration requirement is made for "experimental namespaces".

Formal namespaces are those where some Internet users are expected to benefit from their publication and are subject to several restrictions. Informal namespaces are registered with IANA and assigned a number sequence (chosen by IANA on a first-come-first-served basis) as an identifier in the format "urn-" <number>. Experimental namespaces take the form "X-"<NID>. Namespaces of this form are intended only for use within internal or limited experimental contexts, and are not required or expected to be globally unique.

The IANA URN Namespace Registry page lists all currently registered URN Namespaces[12].

MACE-Dir is the directories working group of the Middleware Architecture Committee for Education (MACE) and administers a Uniform Resource Name (URN) namespace. It supports the assignment of unique, global, persistent names to resources of various kinds by MACE and its delegates[13]. The Namespace Identifier (NID) of the namespace is "mace". RFC 3613 defines the "mace" namespace and describes the procedures and policies governing its use[14]. Other relevant URN RFCs are:

[11] IANA, see www.iana.org.
[12] IANA URN namespaces, see http://www.iana.org/assignments/urn-namespaces/urn-namespaces.xml.
[13] MACE URN namespace, see http://middleware.internet2.edu/urn-mace/.
[14] RFC3613, Definition of a Uniform Resource Name (URN) Namespace for the Middleware Architecture Committee for Education (MACE), see http://www.ietf.org/rfc/rfc3613.txt.

- RFC 3406, "Uniform Resource Names (URN) Namespace Definition Mechanisms", describes how URN namespaces are registered.
- RFC 2141, "URN Syntax", defines the syntax of URNs.

The “mace” namespace is a controlled namespace that is registered with the IETF and IANA for MACE working groups and organizations it works with. The namespace is intended to be delegated to individual organizations through registration with MACE. Once a subspace of urn:mace has been delegated to another organization(e.g. urn:mace:switch.ch) that organization is responsible for any naming and resolution within that subspace. However, it is not permissible to arbitrarily define new attributes within the urn:mace namespace, or in any subtree you have not been granted.

Current registered names in the urn:mace namespace are amongst others[15]:

- urn:mace:switch.ch
- urn:mace:terena.org
- urn:mace:feide.no
- urn:mace:shibboleth
- urn:mace:incommon
- urn:mace:uva.nl

SURFnet hasn’t registered a mace namespace (yet). The attributes in the SURFconext federation (see table above) are defined in three different namespaces: urn:mace:dir:attribute-def, urn:mace:terena.org:schac and urn:mace:surffederatie.nl:attribute-def.

MACE’s attribute-def registry defines URNs for "attribute definitions"[16]. In general these are direct mappings of "attribute types" as used in schema definitions for e.g. eduPerson.

The entitlement registry of MACE defines URNs for values of the eduPersonEntitlement attribute. These values are generally used to indicate the entitlement of a subject to access a resource. This registry is appropriate for defining "global" values that are intended to be useful to the MACE-Dir community. Since any URI can be an entitlement value, values for use by other communities or for specific purposes can be defined and documented by anyone (that is, they need not be registered here). Currently there is only one eduPersonEntitlement attribute registered in urn:mace:dir:entitlement: common-lib-terms[17].

MACE also administers an Object Identifier arc for identifying resources of various kinds for Internet2/MACE projects and working groups[18]. OIDs are used to identify protocol and schema objects in several different technical areas, including X.500/LDAP-based directories, and X.509-based public-key infrastructures. The currently assigned OIDs are:

[15] Overview of registered names in the urn:mace namespace: http://middleware.internet2.edu/urn-mace/urn-mace.html.
[16] Current registered names in urn:mace:dir:attribute-def registry, see http://middleware.internet2.edu/urn-mace/urn-mace-dir-attribute-def.html.
[17] See http://middleware.internet2.edu/urn-mace/urn-mace-dir-entitlement.html.
[18] Internet2/MACE OID registry, see http://middleware.internet2.edu/oid-mace/.

- 1.3.6.1.4.1.5923 for the Internet2 OID arc
- 1.3.6.1.4.1.5923.1 for MACE related work
- 1.3.6.1.4.1.5923.1.1 for eduPerson
- 1.3.6.1.4.1.5923.1.2 for eduOrg
- 1.3.6.1.4.1.5923.1.3 for Higher Education PKI
- 1.3.6.1.4.1.5923.1.3.1 for Higher Education Bridge Certificate Authority
- 1.3.6.1.4.1.5923.1.4 for InCommon™
- 1.3.6.1.4.1.5923.1.5 for eduMember attributes and object classes
- 1.3.6.1.4.1.5923.1.6 for eduCourse attributes and object classes
- 1.3.6.1.4.1.5923.1.7 for usPerson Working Group
- 1.3.6.1.4.1.5923.1.8 for eduPermission
- 1.3.6.1.4.1.5923.1.9 for eduPermissionGroup

Furthermore, MACE-Dir has developed guidelines for higher education institutions wanting to use groups[19]. This work includes attribute definitions and an auxiliary object class for use in LDAP directories. An eduMember schema file is available that can be used to define these in openLDAP-based directories. The schema contains attributes like isMemberOf and hasMember. The isMemberOf attribute associated with an entity is a collection of values each of which identifies a group to which that entity belongs. The hasMember attribute associated with a group is a collection of values each of which identifies an entity that belongs to the group.

2.4 Protocol profiles

Schemas are often embedded in protocols; sometimes protocols have their own schemas.

2.4.1 SAML

In identity federations like SURFconext, SAML is an important protocol for carrying attributes.

The recommended use of attribute definitions from the Internet2 MACE-Dir Working Group with the SAML2.0 specifications have been described[20]. These profiles enable SAML applications that wish to exchange MACE-Dir-specified and profiled attributes to interoperate.

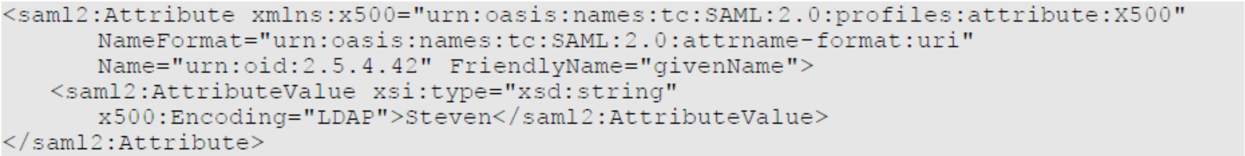
It is recommended that URIs be used for attribute naming in SAML 2.0 attribute statements because of the uniqueness and namespace control they provide. A popular way of naming attributes in SAML is through URLs. The creation and meaning of URLs is generally well understood by many people, and the DNS namespace is already extremely structured. New URLs should only be defined in namespaces that are under control of the creator. Attribute proliferation should be prevented.

[19] See http://middleware.internet2.edu/dir/docs/internet2-mace-dir-ldap-group-membership-200507.html for details.

[20] MACE-Dir SAML Attribute Profiles, April 2008, see http://middleware.internet2.edu/dir/docs/internet2-mace-dir-saml-attributes-200804a.pdf.

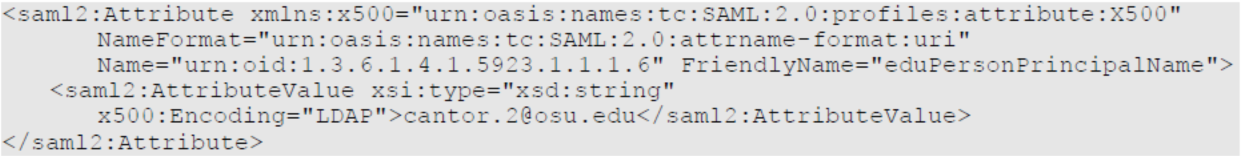
The X.500/LDAP attribute profile of SAML2.0 defines a common convention for the naming and representation of such attributes when expressed as SAML attributes[21]. Section 8.2 of the SAML 2.0 Profiles suggests that LDAP attributes name themselves by utilizing the urn:oid namespace. These names are simply constructed using urn:oid followed by a standard OID. This naming scheme ensures that the derived SAML attribute names are unambiguous. For purposes of human readability, there may also be a requirement for some applications to carry an optional string name together with the OID URN. The optional XML attribute FriendlyName (defined in the core specification of SAML) may be used for this purpose.

The following is an example of a mapping of the givenName directory attribute, representing the SAML assertion subject's first name[22].



```
<saml2:Attribute xmlns:x500="urn:oasis:names:tc:SAML:2.0:profiles:attribute:X500"
      NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri"
      Name="urn:oid:2.5.4.42" FriendlyName="givenName">
   <saml2:AttributeValue xsi:type="xsd:string"
      x500:Encoding="LDAP">Steven</saml2:AttributeValue>
</saml2:Attribute>
```

Another example is a mapping of an eduPersonPrincipalName directory attribute with the LDAP value of "cantor.2@osu.edu". It is a scoped attribute.



```
<saml2:Attribute xmlns:x500="urn:oasis:names:tc:SAML:2.0:profiles:attribute:X500"
      NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri"
      Name="urn:oid:1.3.6.1.4.1.5923.1.1.1.6" FriendlyName="eduPersonPrincipalName">
   <saml2:AttributeValue xsi:type="xsd:string"
      x500:Encoding="LDAP">cantor.2@osu.edu</saml2:AttributeValue>
</saml2:Attribute>
```

2.4.2 XACML

The eXtensible Access Control Markup Language (XACML) defines a declarative access control policy language and a processing model describing how to evaluate authorization requests according to the rules defined in policies. XACML is primarily an Attribute Based Access Control system (ABAC)[23], where any attributes associated with a user, action, resource, or environment of an access attempt can serve as inputs into the decision of whether a given user may access a given resource in a particular way.

Attributes within XACML messages are grouped in user, action, resource, and environment sections. Each of these sections is basically a bag for attributes, where each attribute has a name, a value, and a data-type. The attribute names used in typical XACML messages contain a mixture of namespaces:

- OASIS XACML specific: "urn:oasis:names:tc:xacml:1.0:subject:subject-id"
- Standard specific: "http://schemas.xmlsoap.org/claims/UPN"

[21] Profiles for the OASIS Security Assertion Markup Language (SAML) V2.0, OASIS Standard, 15 March 2005, see http://docs.oasis-open.org/security/saml/v2.0/saml-profiles-2.0-os.pdf.

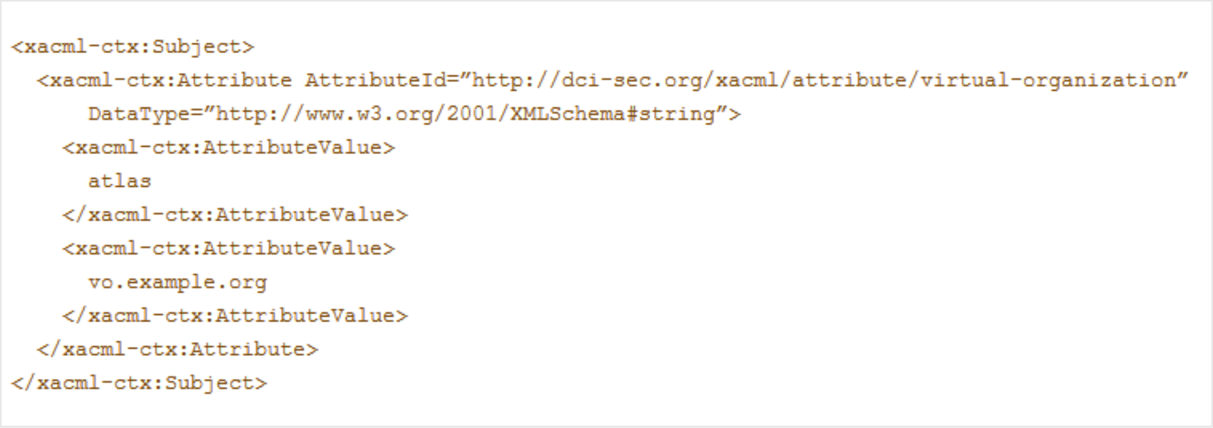
[22] From MACE-Dir SAML Attribute Profiles, April 2006, Document identifier: internet2-mace-dir-saml-attributes-200604, see http://middleware.internet2.edu/dir/docs/internet2-mace-dir-saml-attributes-200604.pdf.

[23] Note that Role-based access control (RBAC) can also be implemented in XACML as a specialization of ABAC.

- Use case specific: “http:// usarmy.org/security/clearance”

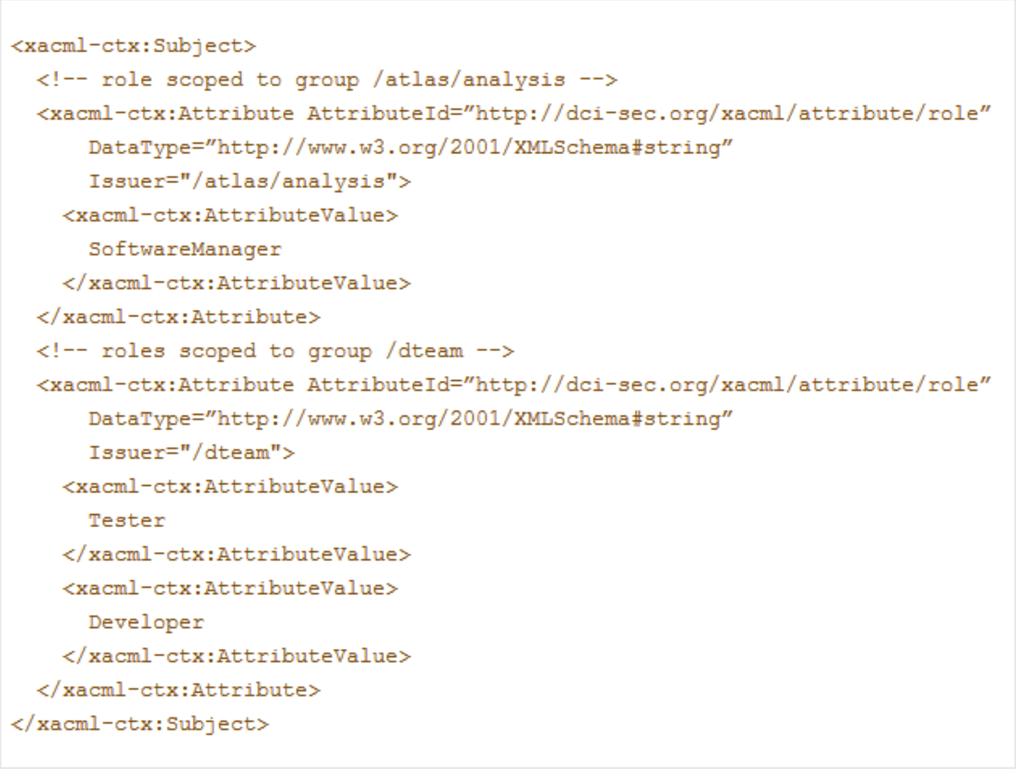
The subject *section* or the *environment* section of a XACML request, are good candidates for handing over virtual organisation (CO) or group related attributes in the same way that traditional SAML requests contain such attributes.

CO membership can be coded as follows in XACML:



```
<xacml-ctx:Subject>
  <xacml-ctx:Attribute AttributeId=”http://dci-sec.org/xacml/attribute/virtual-organization”
      DataType=”http://www.w3.org/2001/XMLSchema#string”>
    <xacml-ctx:AttributeValue>
      atlas
    </xacml-ctx:AttributeValue>
    <xacml-ctx:AttributeValue>
      vo.example.org
    </xacml-ctx:AttributeValue>
  </xacml-ctx:Attribute>
</xacml-ctx:Subject>
```

In this example ‘atlas’ is the CO name. Group membership is expressed in a similar manner. Attributes that are defined in the context of a CO (e.g. roles) and used for authorisation purposes are coded as follows:



```
<xacml-ctx:Subject>
  <!-- role scoped to group /atlas/analysis -->
  <xacml-ctx:Attribute AttributeId=”http://dci-sec.org/xacml/attribute/role”
      DataType=”http://www.w3.org/2001/XMLSchema#string”
      Issuer="/atlas/analysis">
    <xacml-ctx:AttributeValue>
      SoftwareManager
    </xacml-ctx:AttributeValue>
  </xacml-ctx:Attribute>
  <!-- roles scoped to group /dteam -->
  <xacml-ctx:Attribute AttributeId=”http://dci-sec.org/xacml/attribute/role”
      DataType=”http://www.w3.org/2001/XMLSchema#string”
      Issuer="/dteam">
    <xacml-ctx:AttributeValue>
      Tester
    </xacml-ctx:AttributeValue>
    <xacml-ctx:AttributeValue>
      Developer
    </xacml-ctx:AttributeValue>
  </xacml-ctx:Attribute>
</xacml-ctx:Subject>
```

In this example, in the context of the CO atlas the role SoftwareManager is defined. Roles scoped to the dteam group are Tester and Developer.

2.4.3 SCIM

The System for Cross-Domain Identity Management (SCIM) specification is designed to make managing user identity in cloud based applications and services easier. The SCIM specification includes several attribute schemas[24].

Each SCIM Resource (Users, Groups, etc.) includes the below common attributes:

- id – unique identifier for the SCIM Resource as defined by the Service Provider.
- externalId – an identifier for the Resource as defined by the Service Consumer.
- meta – a complex attribute containing resource metadata.

SCIM provides a schema for representing Users, identified using the following URI: 'urn:scim:schemas:core:1.0'. SCIM also defines a number of additional attributes. These attributes amongst others are a username, title, userType, nickName, timezone, password, email, addresses, groups, entitlements, and roles.

Group resources are meant to enable expression of common Group or role based access control models, although no explicit authorization model is defined. It is intended that the semantics of group membership and any behaviour or authorization granted as a result of membership are defined by the Service Provider, and are considered out of scope for SCIM.

The enterprise user extension of SCIM is identified using the following URI: 'urn:scim:schemas:extension:enterprise:1.0'. The following attributes are defined:

- employeeNumber - numeric or alphanumeric identifier assigned to a person, typically based on order of hire or association with an organization.
- costCenter – identifies the name of a cost center.
- organization – identifies the name of an organization.
- division – identifies the name of a division.
- department – identifies the name of a department.
- manager – the User's manager.

2.4.4 VOOT

The VOOT (Virtual Organization Orthogonal Technology) protocol is a subset of OpenSocial used to manage group membership[25]. The primary motivation for VOOT is as a simple tool for managing virtual organization in R&E federations.

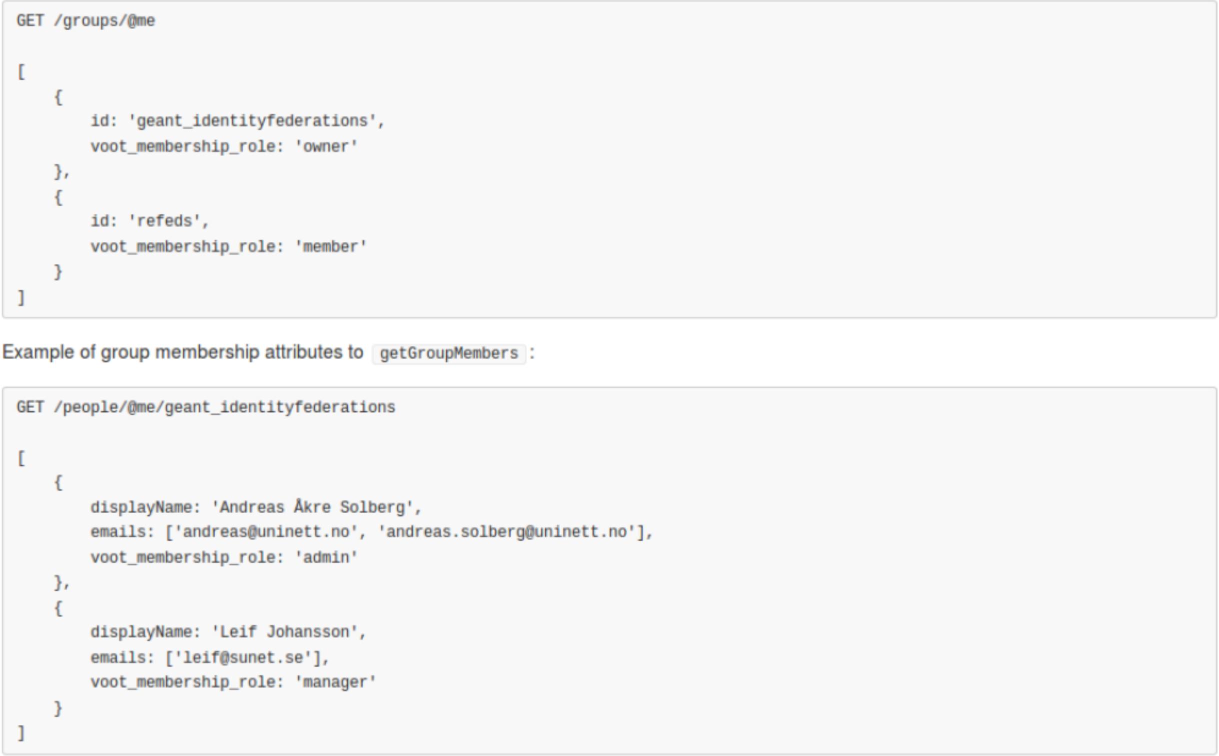
There may be VOOT attributes for both Person and Group objects; but also for group membership. Attributes for group membership may be merged into both Person and Group objects depending on the protocol method used.

[24] SCIM2.0 core schema, see http://tools.ietf.org/html/draft-ietf-scim-core-schema-00.

[25] VOOT 2.0 specification, see http://openvoot.org/voot-2.0.html#anchor13.

There are two calls in VOOT whose answers contain attributes: /groups/@me and /people/@me/<groupId>. The first call returns the attributes that represent the current person’s membership to each of the returned groups. The second call returns the attributes that represent the membership of the relevant group for each of the persons returned. All VOOT specific membership attributes is prefixed with voot_membership_. Currently there is only one attribute: role.

Below is an example of how group attributes are communicated via VOOT.



```
GET /groups/@me

[
    {
        id: 'geant_identityfederations',
        voot_membership_role: 'owner'
    },
    {
        id: 'refeds',
        voot_membership_role: 'member'
    }
]
```

Example of group membership attributes to `getGroupMembers` :

```
GET /people/@me/geant_identityfederations

[
    {
        displayName: 'Andreas Åkre Solberg',
        emails: ['andreas@uninett.no', 'andreas.solberg@uninett.no'],
        voot_membership_role: 'admin'
    },
    {
        displayName: 'Leif Johansson',
        emails: ['leif@sunet.se'],
        voot_membership_role: 'manager'
    }
]
```

2.4.5 OpenID Connect

OpenID Connect is an identity protocol based on OAuth2.0. While the primary use case for OpenID Connect is to relay authentication requests to an IdP, there is also the possibility to convey attributes from IdP to SP. Because of its reliance on OAuth2.0, messages are encoded in JSON format and attributes are expressed as JSON Web Tokens (JWT). The specification calls user attributes “claims”. A very limited set of claims can be part of the authentication response (from the so-called “Token endpoint”). An additional service can be accessed right after authentication to get more user attributes (from the so-called “UserInfo endpoint”). A pre-defined set of attributes is defined by OpenID that includes amongst others name, picture, website, email, birthdate, and address.

IdPs are free to include other attributes as part of a UserInfo response. On the subject of namespaces of these additional attributes the OpenID specification (Section 2.5.4) refers to the JWT specification:

> "While this specification defines only small set of Claims as standard Claims, other Claims MAY be used in conjunction with the standard Claims. When using such Claims, it is RECOMMENDED that collision resistant names be used for the Claim Names, as described in Section 4.2 (Public Claim Names) of the JSON Web Token (JWT) [JWT] specification. Alternatively, Private Claim Names can be safely used when naming conflicts are unlikely to arise, as described in 4.3 of the JWT specification. Or, if specific additional Claims will have broad and general applicability, they can be registered with Reserved Claim Names, per Sections 4.1 and 9.1 of the JWT specification."

The JWT specification appears to exclude the above list of attribute names, and recommends to register additional attributes in the IANA JWT registry[26].

2.5 Summary

A fair amount of standardized attributes is available. The same holds for their namespaces. MACE is an important administrator of namespaces for higher education and research. These namespaces are registered at IANA to ensure uniqueness of the attributes.

Not all protocols that carry the attributes are able to deal with URN- or OID-based namespaces. They have their own set of defined attributes. This implies that attribute mapping is necessary if multiple protocols are used, as is the case in SURFconext that uses SAML and VOOT, and experiments with OpenID Connect.

[26] OAuth Working Group, Internet-Draft, draft-ietf-oauth-json-web-token-10, expires January 15 2014, see http://tools.ietf.org/html/draft-ietf-oauth-json-web-token-10.

3 Attribute usage by the SURF-community

This section describes a number of motivating use cases that have been identified and require specific attributes that potentially do not fit within the current standardised definitions and values or registered namespaces. For each use case the current practice to deal with the attributes and namespaces is described first and is followed by proposed solutions to improve the situation.

3.1 License information

3.1.1 Description

In the coming years, ICT will be utilised more and more via the cloud. Cloud services can be seen as a new method of delivery, the natural successor to the CD, DVD, and downloading. The cloud adds a new dimension to providing access to the services and utilising them and consequently adds new demands on the way licenses should be provided. Licensing rights should be made available with a high degree of flexibility so that applications can be used "any time, any place, and on any device".

Another increasing trend is for students and employees of educational institutions to collaborate nationally and internationally in teams that go beyond the level of particular institutions and disciplines – in fact in "collaborative organisations". This means a broader target group. Today's campus-based license structure will not suffice anymore to support this trend. Licenses for a more targeted and flexible user base will be needed to meet the emerging demands of eResearch, e.g. a license for a webconferencing tool for a limited group of 20 users.

To support both trends, somehow, appropriate license information should be communicated to the SP. SURFmarket has the ambition to cater for this by offering group-based licenses for cloud and other services via their portal to institutional users. The SURFmarket portal could in this case become an attribute provider for SURFconext. This way, SURFmarket can provide the necessary licensing attributes during a federated authentication session via SURFconext. Moreover, the SURFmarket portal will need to offer group-functionality to allow the provisioning of users that are entitled to use the service. It is up to SURFmarket or the service provider to assure that indeed only 20 users have access to the webconferencing tool.

The problem, however, is how to provide the group-based licensing attributes in such a way that they can be understood by the SPs. The current attribute schemas do not cater for such attribute information.

3.1.2 Current practice

The current solution doesn't involve any specific license-based attributes and is focussed on authorisation. Most licenses are campus licenses and they are enforced by SURFnet via an access control list. This ACL is based on the information SURFnet obtains from SURFmarket regarding licenses bought by an institution for a specific application. The ACL determines which institution has access to which application. So access is based on institutional level. Also billing is based on an institutional level. For 'smaller', group-based licenses this is not going to work as it requires a more fine-grained approach.

There is one exception. Elsevier (Scopus) implements the "common-libs-terms" entitlement (part of the eduPerson attribute scheme) to grant license holders access[27]. Elsevier has configured their service in such a way that it looks for the eduPerson entitlement attribute with a value of urn:mace:dir:entitlement:common-lib-terms in order to authorize access. The identity provider of the institution has to provision the attribute and Elsevier verifies against its own administration during authentication if the attribute is valid.

3.1.3 Proposed solution

With an increasing variety of license models for smaller user groups within the institutions or even across institution, the maintenance of such an access control list becomes a large burden for SURFnet. Access cannot be enforced on institutional level anymore. For instance, two faculties of the same university can have bought a license at the same service provider. A more fine-grained access control based on additional attributes is required in this case. This can be realised via a combination of existing attributes or via a new attribute: the license ID, a unique number that is generated for each bought license. This attribute isn't part of the existing schemes at the moment. Moreover, the license ID attribute values should be scoped to the context of SURFmarket to ensure its uniqueness.

For this purpose, SURFmarket could register its own (local) namespace that includes the license ID attribute. For the time being, this could be sufficient as licensing in other NRENs isn't as advanced as in the Netherlands. In particular the role of SURFmarket as procuring organisation of licenses for the Dutch education and research community is unique. However, with the popularity of cloud providers it is expected that similar license issues will appear in other NRENs too. In that case an international approach is required, one that focuses on standardisation of the license ID attribute and its values on an international level.

The license ID attribute can be either provided by the IdP or by SURFmarket as third party attribute provider.

[27] See http://www.info.sciverse.com/scopus/scopus-training/faqs.

An interesting use case is group-based licenses. In that case there is a group ID associated to the license ID. In case the group is relatively generic, such as faculty members or chemistry students, the IdP should provide the attributes (e.g. the EduPersonOrgUnit attribute in combination with schacHomeOrganization or license ID attribute). If the group is relatively small and diverse (e.g. a collaborative organisation), the IdP cannot provide such information. In that case SURFmarket could provide either the license ID or the associated group ID attribute. The eduPerson entitlement attribute could be used for that purpose. Moreover, the federation trust fabric should be sufficient to guarantee the trustworthiness of the provided attribute values, i.e. the SP trusts SURFmarket to provide valid license IDs. Section 3.2 describes the group aspect in relation to attributes and namespaces in more detail.

3.2 Collaborative Organisation or group information

3.2.1 Description

CO/group management is essential in any eResearch environment. It provides the flexibility for managing access and offering collaborative services. CO/groups can be generated based on an attribute about a party, or they can be formed ad hoc, when parties need to collaborate and share resources. Identifying the groups or COs that a person belongs to can be essential to defining which resources a party can access. This allows better granularity for service authorisations and helps with group management for access to e.g. wikis, blogs and other collaborative software.

A typical use case that illustrates the need for efficient authorisation and group management is a PhD student who needs to prove that:

- StudieLink says she is a student;
- A collaboration management platform that says she is a member of a specific CO;
- A particular faculty says she is part of their team;
- The university says she is part of their faculty.

Each of these authorities could make a claim about the identity of the student by asserting some identity-related attributes.

CO or group related attribute characteristics are:

- There may be many of them and a user may participate in any number of them.
- They contain meta information: the group probably has an owner and there may be subgroups.
- The members may come from other institutions, federations and/or countries that potentially use other attribute naming, syntax and semantics.
- Group or CO membership is privacy sensitive.
- They may have their own role definitions for the members (owner, admin, member).

Note that the CO does not have its own IdP, it builds on top of identities from participating institutions. Nevertheless, CO or group attributes describe users by their memberships, roles, and capabilities they have within a CO or group. These attributes are typically managed and issued by a dedicated collaboration management system such a COmanage or SURFteams. These systems can be considered as attribute providers that are authorative for CO-specific attributes.

These aspects should be taken into account for efficient CO or group related attribute exchange with SPs.

3.2.2 Current practice

There are 2 possible ways of describing group or CO membership with SAML attributes as they are currently available: SchacHomeOrganization or isMemberOf.

isMemberOf is used by SURFconext (and other federations such as Switch) and seems to work, as long as the URI is scoped accordingly (examples follow below).

SchacHomeOrganization, combined with SchacHomeOrganizationType is an alternative approach. For a CO the type is either nationally or internationally:
urn:schac:homeOrganizationType:<country-code>:<string>

Where country code can also be "int" to indicate international. The actual value of <string> is still open for debate. SCHAC states it is a <string> from a nationally controlled vocabulary, which is in contradiction with the fact that an "int" country code would not fall under a "nationally controlled vocabulary".

The current use of SchacHomeOrganization in federations is that it contains the institution (only). When asserted by an institutional IdP this makes sense. However, in the context of COs, an institutional IdP is not always authorative for CO-specific attribute assertions. In this case the CO or collaboration platform asserts the attribute. An additional challenge here is that SchacHomeOrganization itself is single valued. This implies that it is not allowed to have both the institution and the CO to be present as attribute values. A solution here would be to create/define a multivalued schacVirtualOrganization attribute.

Given the complexity the use of SchacHomeOrganization brings, using isMemberOf is currently the most pragmatic solution. As there are many COs/groups, scoping is of utmost important here. Other specific attributes may be defined or used in the context of a particular CO (e.g. for authorisation purposes, see section 3.3).

For the actual value of the CO/group attribute several approaches are possible:
isMemberOf=https://vo.surfconext.nl/3TU or
isMemberOf=https://vo.surfconext.nl/3TU/lecturers. Where the first part provides the authoritative source ("vo.surfconext.nl"), and all following parts signify group ("3TU") or subgroup ("lecturers" in "3TU") memberships.

In this case, the provided attribute values mean something (e.g. "lecturers" being a subgroup of "3TU"). It could be tempting for SPs to 'interpret' the values and this may come at the cost of privacy. Another approach would be to use meaningless attribute values.

The use of a URL representation also allows for discovery of (additional) group attributes, if the URL actually points to something useful. It would preferably require a standardised protocol or API to provide these additional attributes. VOOT for example is used by SURFconext for this purpose (e.g.: https://vo.surfconext.nl/group/[userid]/[groupid]).

3.2.3 Proposed solution

The isMemberOf attribute is suitable for communicating group membership information. The attribute values are typically defined and set by the CO manager via the CO management platform. Being part of the federation, the CO management platform operates as a trusted authority issuing the attributes.

COs/groups must be created within namespaces. These namespaces should be arranged hierarchically to support subgroups:

faculties	group
faculties:arts	sub-group
faculties:arts:all_staff	sub-sub-group

Namespaces scope the authority to create and name groups. The CO management platform is responsible for the namespaces:

surfteams:faculties:arts:all_staff

CO management platforms should provide functionality to register CO attributes and values (e.g. Grouper offers such functionality).

As multiple CO management platforms (COmanage, COMS) or group providers (SURFteams) could be connected to SURFconext, a higher-level namespace is required to facilitate scoping:

surfconext:surfteams:faculties:arts
surfconext:comanage:co

In case of internationally oriented groups or COs a MACE-operated namespace is useful:

mace:groups:surfconext:surfteams:faculties:arts

A users can be member of multiple groups/COs. In case the same SP serves these groups, the must be asked to select a group/CO (some sort of Which Group Are You From). This means that the group name must be defined in a for the user understandable manner (i.e. a FriendlyName in SAML terms).

3.3 Authorisation

3.3.1 Description

In most identity federations, attribute-based access control (ABAC) is the preferred authorisation policy approach. In ABAC permissions are defined in terms of privilege-giving attributes, also known as entittlements. With the increasing number of service providers in federations, the variety of attributes that is used for authorisation purposes increases as well. The following use cases exemplify this trend:

- Sometimes an SP requires the IdP to provide a specific attribute for service access, e.g. a license-related attribute for downloading educational content. IdPs are then obliged to store these attributes in their LDAP databases.
- A PhD researcher shares her experimental data with several researchers in other institutions in the federation. She defines a group of collaborating researchers and grants them access to her data and two additional web-based collaboration tools. Access to the data is based on group membership. When a student is going to support the PhD researcher, he is added to the group as well and immediately gains access to the experimental data and tools of that group.
- A CO-specific SP defines two specific roles with distinguished rights: operator and observer. The CO members need these roles to be provisioned by the CO manager somehow.
- SURFnet's Authorisation Management service (SAB) allows institutional administrators to control access to SURFnet specific services.

As is the case for many applications, the complex authorisation hierarchy that consists of several roles is implemented in the application itself. Particularly in group contexts where there is a need for reusing group-specific attribute across multiple group services, this means that the appropriate attributes need to be provided by the IdP or group or collaboration attribute provider in order to get the right authorisations at the SP.

In all cases, it is the SP that determines which attributes values are required for authorisation enforcement. The federation operator determines the attributes that are used in the federation.

3.3.2 Current practice

Three approaches to ABAC are possible:

1. Give the precise attributes and values of the user so that the SP can immediately enforce access. SURFconext knows exactly which attributes the SP needs and provides them.
2. Give the available attributes and values of the user and leave it to the SP to interpret them and draw authorisation conclusions. SURFconext doesn't know which attributes the SP needs and just passes the available ones.
3. Do attribute mapping between provided and required attributes at an intermediary hub.

At the moment, the first and third approaches are adopted by SURFconext. The first approach is implemented in the form of attribute release policies. These policies specify the attributes that an IdP releases towards an SP and are set during the configuration of the federated connection between the two parties. Sometimes the IdP cannot provide the right attributes. In that case SURFconext does attribute mapping, i.e. it converts IdP provided attributes into ones that are understandable by the SP (e.g. 'urn:mace:dir:attribute-def:givenName' => 'gn'). Such mapping is used for instance in the case of WordPress[28]. This approach, however, doesn't scale well in the case of numerous COs and SPs that each have their own specific attribute requirements.

Entitlements are often used for authorisation purposes. A specific service that makes use of entitlements is the TCS service for eScience. This service requires that an identity has been properly validated by setting a specific value in the eduPersonEntitlement[29]. The eduPersonEntitlement = urn:mace:terena.org:tcs:escience-user means that the user is authorised to use the TCS service. Another example involves the use of entitlements for the Rationale cloud service. In this example two institutions, UvA and HU, provide the eduPersonEntitlement attribute. The value of this attribute differs for each institution: urn:mace:uva.nl:rationale:rationale-user and urn:mace:hu.nl:rationale:rationale-user. The use of such urn-based values allows SURFconext to filter efficiently and allows Rationale to distinguish users from UvA and HU.

SAB, as mentioned above is an example of a solution. SAB recognises about 10 roles and is utilized via a SURFnet owned LDAP that is structured with an OID-based namespace. SAB is used for SURFnet services only.

3.3.3 Proposed solutions

In case the SP requires attributes that are not part of the standard schemes, it should create its own attributes and values including a namespace to guarantee uniqueness. This allows IdPs or COs to provide the required attributes or to find a mapping between their own attributes and the ones the SP requires (the actual mapping can e.g. be done by the federation operator).

Several improvements are possible to better match the SP required attributes with those that can be provided by the CO management platform or IdP in order to reduce the mapping overhead:

[28] See also SURFfederatie - simpleSAMLphp for Service and Identity Providers, François Kooman, 2010.

[29] See section 3.2.3 of Certificate Practice Statement (CPS) for the TERENA (TERENA Trans-European Research and Education Networking Association) Certificate Service (TCS) eScience Personal Certificate Authority, http://www.terena.org/activities/tcs/repository/cps-personal-escience.pdf.

- Harmonisation of ABAC attributes by standardising a relatively small set of well-defined attributes that could be used for federated authorisation purposes. On a national level this is implemented by SWAMI, the Swedish Alliance for Middleware Infrastructure. SWAMI has defined and adopted a General Model for Authorization Information (GMAI). The model suggests that most authorization decisions can be based on a tuple with two or more elements of authorization information. The two first elements contain information about Application/Application Area, and Role/User Type respectively. If applicable there may, in addition, be one or more elements defining restrictions on the Scope of Authority. These tuples can be explicitly stored in for example an LDAP directory or generated as requests for authorization information are received. The model stimulates the use of the following roles in the Swedish higher education sector: Self Reporter; Handling Officer; Reviewer; Certifier; Controller; Reader. The drawback of this harmonisation approach is that it only works on a national level. It will not work for internationally oriented COs or for inter-federation scenarios.
- Better alignment of CO- and SP-attributes. COs should get the means to administrate their own attributes (e.g. roles and rules) and map them onto the SP required attributes and their values. This approach tackles several CO specific challenges:
  - There is a CO management platform that registers the attributes and their values and is authoritative for issuing them to SPs.
  - The issued attributes are scoped to a specific CO.
  - There is a global understanding as COs may be internationally oriented.

  In terms of namespaces this could look as follows:
    - InstituteX_for_serviceY_for_collab_Z_admin
    - Urn:a:b:c:instX:serviceY:collabZ:admin

  Since the authorisation attributes are scoped (in the context of a CO and a particular SP), they require a namespace. This namespace can be operated by the CO management platform and delegated to the CO-manager. Most of the existing CO management platforms lack functionality to automatically create and manage their own namespaces. It could be helpful to provide guidelines on how COs should define their namespaces and handle authority/releasing of attributes.

The proposed alignment of attributes between SPs on the one hand and IdPs/CO management platforms on the other hand can easily be automated and done via self-service. SAML metadata can be used for this purpose. SP attribute requirements may be called out in SAML metadata by using the <md:RequestedAttribute> element. The <md:RequestedAttribute> element (based on type saml:AttributeType) has a boolean attribute (isRequired) that "indicates if the service requires the corresponding SAML attribute in order to function at all (as opposed to merely finding an attribute useful or desirable)."[30] Standard SAML metadata supports zero or more <md:AttributeConsumingService> elements each containing one or more <md:RequestedAttribute> elements in SP metadata. These static elements are used to communicate SP attribute requirements to IdPs or CO management platforms. The <saml:AttributeValue> can be used to express attribute values in the metadata.

An attribute management interface is required for the CO manager to define CO specific attributes and to map them onto SP required attributes. Specific provisioning solutions depend on the type of application and are discussed by Oostdijk et al.[31]

The use of SP metadata may take away the need for COs to define proprietary attributes. Should this be needed then it is recommended that they only define new attributes as a last resort when no suitable definition exists elsewhere. It is strongly recommended that any new attribute definitions follow the existing SAML attribute, SCIM or OpenID Connect naming conventions[32]. COs should carefully consider the privacy and data protection implications of any newly invented attribute. The following steps should be followed when naming a new attribute:

- Is this attribute standardized or defined by any organization which has already assigned it a unique identifier? If so, it should be used if at all possible.
- If the attribute is defined through an LDAP object class, there is probably already an OID assigned. When possible, leverage the existing urn:oid namespace.
- If no suitable name yet exists for this attribute, consider creating one preferably through constructing a proper URL, or if necessary using a delegated urn:mace namespace.

Another approach is to better utilize the entitlement attributes in the standard schemas for authorisation purposes. An entitlement is an indication of something that the person is allowed to do once they have been authenticated; an attribute is some property of the individual - eye colour, age, sex and staff category being examples. Note: in practice, both attributes and entitlements (as used here) are carried as SAML attributes - the difference lies only in their semantics.

[30] Metadata for the OASIS Security Assertion Markup Language (SAML) V2.0 http://docs.oasis-open.org/security/saml/v2.0/saml-metadata-2.0-os.pdf.

[31] M. Oostdijk et al., Provisioning scenarios in identity federations, GigaPort3 EDS-4 deliverable, 2010, see http://www.surfnet.nl/nl/Innovatieprogramma%27s/gigaport3/Documents/EDS-4%20Provisioning%20Scenarios%20in%20Federations%20Final.pdf.

[32] The use of multiple conventions requires mapping conventions between them.

In most use-cases it is possible to use either generic attributes or entitlement attributes to achieve a particular task. For example, individuals with a staff category of 'librarian' (an attribute) may be inferred by the SP to be allowed to order new books within a library management system - anyone with that attribute is allowed to do so. Alternatively, a 'bookOrdering' entitlement may be used - only people with that entitlement are allowed to order new books, irrespective of whether they are a librarian or not.

Entitlements are a major way for a service provider with many federation members to simplify its application work by receiving authorisation information from those members. A shared set of business rules, among the IdPs or with SURFconext, about access control permissions could translate into a consistent assertion of appropriate entitlements. Note that entitlements may reflect a community of interest consensus that may span multiple federations. On the other hand, consensus may not even apply to all members of a single federation.

It makes sense to work on a standard for expressing entitlements and on how to pass these to SPs. The use of meta-data, again, could be very useful for the communication of entitlement values between SPs and IdPs. The SAML Metadata specification allows for publication of entitlement values by the SP. Moreover, establishing an entitlement takes a namespace to anchor values and business logic to help IdPs or federation gateways like SURFconext properly calculate the entitlement value to send. This anchor is already available: "urn:mace:dir:attribute-def:eduPersonEntitlement". The value will need to be defined by the SP, after all they have to enforce access based on the entitlement value. The SP could for instance register a namespace for a service and define the value: "urn:mace:sp.nl:service:entitlement-value". IdPs could use this entitlement-value and scope it by providing it in a urn-style: "urn:mace:idp.nl:service:entitlement-value. The use of such scoped entitlement values allows for efficient filtering by SURFconext and for determining the origin of the attribute by the SP.

3.4 LoA

3.4.1 Description

Increasingly service providers require stronger user authentication solutions than a username and password combination. The strength of the authentication solution is usually expressed in terms of levels of assurance (LoA). Four more or less standardised LoAs exist:

Level 1 – Little or no confidence in the asserted identity

Level 2 – Some confidence in the asserted identity

Level 3 – High confidence in the asserted identity

Level 4 – Very high confidence in the asserted identity

These levels have been specified by multiple independent parties – such as NIST, STORK and eHerkenning – and they largely overlap.

3.4.2 Current practice

In SAML the LoA can be communicated as an attribute or via the SAML Authentication Context ability[33]. The attribute way is not described in the SAML specification. A model to overcome this has been proposed[34]. The proposed extensions integrate nicely into a standard saml:AttributeStatement and convey the metadata about individual attributes to an SP that can make a more nuanced access control decision. An advantage of the Authentication Context approach is that it is well-described in the SAML specification. A drawback is that it cannot be easily (re)used in other protocols like SCIM or VOOT that lack the concept of authentication context. If support for multiple protocols is required, the attribute approach is preferred as it allows for easier reuse of LoA attributes. Another approach is possible but implies mapping of SAML authentication context to LoA attributes.

Another aspect to take into account is authorization. If SPs wish to grant access based on a user's identity attributes and the level of identity assurance, then it makes sense to include the LoA as a subject attribute. Whilst this does not require the IdP to insert the LoA attribute into the SAML attribute assertion, as the Policy Enforcement Point (PEP) can always pull it out of the SAML authentication context statement and put it into the authorization request to the Policy Decision Point (PDP), it would make the SP's task easier if the LoA attribute was already in the attribute assertion.

The US Incommon federation is one of the few federations that makes use of LoAs. It has the notion of 'Silver' and 'Bronze' members in their Identity Assurance program[35]. These LoA profiles align with those of NIST and are communicated via the SAML Authentication Context construct. This is done as follows: when the user requests access to a service, the SP sends an Authentication Request message containing a <RequestedAuthnContext> element to the IdP. The IdP determines whether it can satisfy the request and sends a response back to the SP.

3.4.3 Proposed solution

Several solutions are possible:

[33] R.L. Morgan et al., SAML V2.0 Identity Assurance Profiles, Version 1.0, OASIS Committee Draft 01, 22 September 2009, see http://docs.oasis-open.org/security/saml/Post2.0/sstc-saml-assurance-profile-cd-01.pdf.

[34] Ivonne Thomas & Christoph Meinel, An Identity Provider to manage Reliable Digital Identities for SOA and the Web, IDtrust '10, April 13-15, 2010, Gaithersburg, MD.

[35] Incommon Identity Assurance program, see http://www.incommon.org/docs/assurance/Assurance_InfoSheet.pdf.

- To claim the existing eduPersonAssurance attribute for this purpose. This multi-valued attribute represents (i.e. URIs pointing to) standard identity assurance profiles. The MACE-Dir group decided to leave eduPersonAssurance attribute fairly open so that adoption would not be discouraged by overly stringent or bureaucratic controls. The drawback of this decision is that a common controlled vocabulary, universally known and accepted, for values to assign to this attribute is lacking. It could be considered to use the four LoA values of the US NIST or EU STORK or Dutch eHerkenning for this purpose. However, Incommon is already using 'Silver' and 'Bronze', so that would require mapping of attributes values to make it work. Important is that the values are internationally known and accepted; the values should mean the same for every IdP in the world. Work is underway to setup a global IANA registry for assurance profiles. This will allow every community to define their assurance profiles and register it according to RFC 6711[36]. The eduPersonAssurance attribute can then point to those registered LoA profiles.
- Definition of a new attribute, e.g. schacIdentityAssurance with a LoA value. This schac attribute might be redundant to the eduPersonAssurance attribute and might cause confusion.

Obviously, the first solution direction seems the way to go.

Another, less technical approach, regarding attribute assurance is to professionalise IdPs in the federation. Setting and governing high-quality levels regarding identity and attribute management by the IdPs may help improving the overall trustworthiness of the attributes that are asserted by the IdPs. Auditing is an excellent tool to enforce and guarantee high-quality levels.

Furthermore, if the SP cares about the reliability of one or more attributes it can deal with that in the contract that it engages in with the IdP/Attribute Provider/Federation Operator. In other words such a SP does not treat all IdPs equal.

3.5 Generic challenges

A number of generic challenges related to attributes need to be addressed.

3.5.1 Attribute filtering and scoping

CO/group membership and role attributes are privacy sensitive and should only be communicated to the SP when relevant. At the moment SURFnet filters attributes on the basis of ARPs and ACLs. This filtering goes beyond attributes and in some cases also includes filtering at the level of attribute values (one attribute may have multiple values depending on the context/scope of the attribute request). With the emergency of a plethora of attributes the scalability and maintainability of the filtering become an issue. Automated ARP specification via metadata exchange and self-service, standardisation and scoping become important to tackle these issues.

[36] Global LoA registry, see http://tools.ietf.org/html/rfc6711 and the accompanying website that tries to explain the process: http://levelofassurance.org/process.html.

Scoping can be improved by using urn-based attribute values. The urn should be constructed such that it allows determination of the origin of the attribute-value. For instance:

- urn:mace:uva.nl:rationale:rationale-user
- urn:mace:surfmarket.nl:institution:license-ID

3.5.2 URL or URN style?

URIs can be classified as locators (URLs), as names (URNs), or as both. A uniform resource name (URN) functions like a person's name, while a uniform resource locator (URL) resembles that person's street address. In other words: the URN defines an item's identity, while the URL provides a method for finding it.

Attributes profiles can use either the URL or the URN formats. Both styles present advantages and disadvantages.

The URL style namespace is preferred because it does not require the registration of a namespace with any standardization body. The uniqueness of the namespace is derived by the uniqueness of the domain name. Moreover, additional (discovery) services for XML schema resolution and location can be established at the registered domain. URL attribute names may even be resolvable into documentation, providing helpful information for unwitting relying parties.

The URN style namespace is instead desirable for reasons of compatibility with standards bodies like OASIS and IETF. However, using a URN requires the formal registration of the namespace with bodies like IANA. To obviate this problem, a profile can define a URN starting with the "x-" prefix, for "experimental namespace", that doesn't require registration with IANA.

A specific example of a URN style namespace is based on Object Identifiers (OIDs). OIDs enjoy the benefits of flexibility and structure but suffer from readability. URN style namespaces are most commonly used.

Currently in most federations attributes are communicated in the urn:oid as well as the urn:mace:dir format. Typically urn:mace:dir is used for SAML1.1 and urn:oid for SAML2.0.

3.5.3 Attribute harmonisation

A returning issue is the difference in naming, syntax and semantics of attributes used. With difference in naming, other attributes are used to denote the same (type of) information. With difference in syntax, the same attributes are encoded differently (e.g. using a numeric (O)ID instead of a string based ID). With semantic differences, the same value can mean different things in a different context (e.g. 'affiliation=staff' has a different meaning in the US than in the UK or mainland Europe).

Within a federation these issues normally should not appear, as attribute naming, syntax and semantics are usually well defined within a federation boundary. Possible problems arise when federations connect in one form or another, or when information — as a basis for attributes — is retrieved from multiple sources.

With a Collaborative Organisation (CO) both situations can occur, since members can come from multiple federations (either directly or via eduGAIN) and a CO management platform may retrieve additional information about users (e.g. for group membership) before releasing attributes to an SP.

Although these issues may not appear for small (national) setups, for typical 'CO' scenarios such as CLARIN and LifeWatch they most likely will. These are therefore not issues that can be ignored.

Note that in the end all these issues are essentially mismatches between attributes provided by (virtual) IdPs and consumed by SPs, both syntactically and semantically.

Note that before a solution can be applied, the occurrence of the problem should be identifiable. In other words: how does one know that the attributes supplied by an IdP do not match with the attributes expected (in syntax or semantics) by an SP? For a workable solution, identifying the problem should — as much as possible — be an automated process. This implies the use of metadata to describe the attributes released and consumed. At least for SPs this is already possible using the SAML specifications (see also the eduGAIN metadata at http://mds.edugain.org for example).

Three different solution directions can be used to tackle attribute interoperability issues[37]:

[37] R.J. Hulsebosch et al., Virtual collaboration attribute management, Gigaport3 2011 deliverable, see http://www.surfnet.nl/nl/Innovatieprogramma%27s/gigaport3/Documents/EDS%2011-06%20Attribute%20Management%20v1.0.pdf.

1. *Standardisation.* Interoperability can be guaranteed by standardising attribute naming, syntax and semantics. Although, reality seems to show that global harmonisation of attributes is not a feasible goal in the short term, there have been successful attempt to harmonise individual attributes. See for instance the proposal made by Andrew Cormack on eduPersonScopedAffiliation (ePSA[38]). Another rather successful attempt to harmonise attributes for inter-operability use-cases was provided by Schac. However, the interoperability issues mentioned above exist *despite* standardisation, the issue is that standardisation is limited to a specific context (national or international). In order to solve these issues, standardisation has to be done on a global scale, which is nearly impossible. However, standardisation can be used to solve (or lessen) this problem by specifying profiles and template sets of attributes (in other words: standardising and identifying differences). For example: within eduGAIN a finite set of (both semantically and syntactically) well-defined attributes are used, which can be identified as the 'eduGAIN profile'. Similar profiles can be identified for different circumstances or collaborations. Specific SPs can identify in their metadata the profile expected and the specific attributes from within that profile; for ease of reference specific combinations within a profile can be grouped together in templates.

2. *Semantic (web) based approaches.* The idea behind the semantic web approach is that by describing the attributes used and their semantics using ontologies, tooling can be used to provide automatic translation between the different representations both in syntax and semantics. Although the idea of having automated translation sounds attractive, the precondition is that the attributes at the two sides of the translation have to be described in the first place, which in practice has to be done (again) manually, only to be able to provide 'automatic' translation in — most likely — fairly simple scenarios. Adding to this the relative complexity of semantic tools and languages (RDF) makes this solution a heavyweight approach to a lightweight problem.

[38] See http://www.terena.org/activities/refeds/docs/ePSAcomparison_0_13.pdf.

3. *Attribute translation support.* For the interfederation between the SURFconext federation and eduGAIN some interoperability issues are present. The approach taken by the SURFconext federation is to do specific translation and filtering for eduGAIN. This translation is configured statically; this is possible because the hub-&-spoke model of the SURFconext federation means that the central hub plays a (proxying) role in the attribute exchange between IdP and SP, and because the attributes, values and semantics are standardised on both sides with no 'overlapping ambiguity'. The setup and limited set of standardised eduGAIN attributes also means that the translation can be setup statically.
   The CO approach shares similar characteristics, although there are some differences. The similarity is that the required mapping will most likely be fairly simple in the majority of the cases, consisting of simple one-to-one translations of the attribute names or values. The difference is that the required translations depend on the links between multiple IdPs and SPs, or, in other words, on the precise setup of a specific CO. This dependency implies that static translations cannot be built into the CO management platform beforehand. Translation support in the CO management platform case should therefore consist of a configurable translation engine, although due to the simplicity of the majority of the cases the type of translation supported can be limited to simple one-to-one mappings. If more complex support is deemed necessary, other approaches can be used, such as a scriptable or pluggable translation engine.

Given the pros and cons described in the solution directions above, combined with the majority of the use cases being fairly simple, attribute translation support is the appropriate choice for solving this issue. If providing a solution for more elaborate scenarios is needed, then a scriptable or pluggable setup — possibly realised in steps — is called for.

The Kantara Initiative has started a new working group called "Attributes in motion" with the aim to create a set of best practice documents around[39]:

- The handling of attributes by Identity Providers, Relying Parties, and Service Providers
- The definition and proposed use for contexts
- The definition, best use, requirements and criteria of an Attribute Broker

It may be worth monitoring the progresses in this space.

[39] Kantara Attributes in Motion, see http://kantarainitiative.org/confluence/display/AIM/CHARTER+-+Attributes+in+Motion+Work+Group.

4 Conclusions

SURFconext relies on consistent attribute naming to deliver information about users in a mutually understood way between the IdPs, APs and SPs. Every attribute which has a definition and semantics that differs from others must have its own unique representation in the protocols to be used for attribute exchange (e.g. SAML or VOOT) to ensure that there are no misinterpretations or communication failures. This name must be expected and handled by relying parties. Moreover, the attribute values, vocabularies, and their meaning should be clear as well. This, at least, limits proliferation of attributes and overcomes a potential attribute ambiguity problem in the federation.

Standards bodies have traditionally defined common attribute schemes (e.g. schac, eduPerson, etc.) that are used in identity federation. Moreover, these attribute schemes - consisting of attribute names, values and semantics - are published to public URN/URL namespace repositories.

However, use cases are emerging that require attributes that are not catered for by these standards. Examples are the provisioning of attributes regarding license information concerning the use of cloud services, for CO or group membership, for expressing identity assurances (i.e. LoAs) and for specific authorisation purposes. In most of these cases, the attributes are not provided by the institutional IdP but come from an authoritative third party attribute provider (e.g. a collaboration platform such as COmanage, a license provider such as SURFmarket, or an external authentication provider for higher LoAs).

For these scenarios a common, accepted list of attributes and associated definitions is currently not (yet) achievable in its entirety. Moreover, there is a tendency to create local attributes and namespaces, i.e. on a federation or national scale. Such local definition of attributes may create problems when international exchange of attributes becomes relevant (e.g. in collaborative organization settings).

In most cases the service provider determines the attributes that need to be exchanged for authorization or personalization purposes. To prevent miscommunication or non-interoperability with IdPs or third party attribute providers the service providers should consider the following steps:

Step 1: Check whether an existing and standardized attribute and its values suffice. If so, use that, but only if the intended semantics correspond with the intended use of the attribute.
Step 2: Make this clear to prospective federation operators, COs and/or IdPs.
Step 3: If not, define an own attribute and/or value, register it in an existing, internationally accepted delegated URN-based namespace.

For each of the scenarios the following can be recommended regarding the use of attributes and namespaces.

Regarding the communication of license information:
An emerging type of attributes is related to licenses for in particular cloud application usage. These licenses are often short-lived and associated to relatively small user groups. Access to licensed service can be based on:

- A specific entitlement attribute value that is assigned by the IdP to selected users. These entitlement values should be urn-based and scoped such that the institutional origin is traceable. An alternative approach to scoping is to make use of a license ID attribute that is created and provided by the organisation that sells licenses (e.g. SURFmarket). The provisioning of this attribute helps the service provider to determine the usage conditions of the service (e.g. institution, CO, and amount of users that are allowed to make use of the service). A namespace for the license IDs is required and should be managed by the license ID provider. This will help to scope the license ID attribute values in order to determine the origin of the attribute (i.e. to be able to prove to the service provider that the attribute was provisioned by SURFmarket). Or,
- A specific license ID attribute that is coupled to a group membership attribute. Being member of a certain group allows the user to make use of the service. In this case the isMemberOf attribute should be used. Values should be urn-based and scoped as well since third party attribute providers are likely to provide them. Linking the group ID with the license ID attribute eases the scoping effort.

In case multiple user groups within a single institution have bought a license for the same service provider access cannot be enforced anymore on an institutional level as is the current practice. Scoping of the attribute values becomes even more important, i.e. the urn should include group characteristics. A user that participates in multiple groups for the same service provider should be asked to select a particular group prior to accessing the service.

The use of the team membership attribute is preferred in case of inter-institutional collaborations (see below).

Regarding CO membership attributes:
The isMemberOf attribute is suitable for communicating group membership information. The attribute values are typically defined and set by the CO manager via the CO management platform. Being part of the federation, the CO management platform operates as a trusted authority issuing the attributes.

COs/groups must be created within namespaces. These namespaces should be arranged hierarchically to support subgroups. Namespaces scope the authority to create and name groups. The CO management platform is responsible for the namespaces. CO management platforms should provide functionality to register CO attributes and values. The MACE owned namespace register could be a suitable place for these CO platforms to claim their own namespace.

Regarding attribute-based authorisation:
For authorisation purposes entitlement attributes should be used to express specific rights of a user in an application. The SP determines the entitlement attributes and their values that are required for accessing a service. Typically the eduPersonEntitlement attribute will be used for this purpose. The required values should be listed in an SP managed namespace and published via e.g. SAML metadata. Based upon this information, the IdP or third party attribute providers can take measures to provide them. Scoping of the provided entitlement attribute values is required. This could be achieved by using urn-based values.

So far, entitlements have barely been defined in existing schema. Further work on a standard for expressing entitlements and their values and on how to pass these to SPs is required.

Besides standardisation of attributes, attribute mapping is a straightforward manner for overcoming attribute interoperability between SPs, IdPs and third party attribute providers like CO management platforms. The function of attribute mapping is to map external (unknown) attributes into internally known ones. With this approach the unique combination of issuer, attribute type and attribute value are mapped into a locally understood attribute value or role, and the remote users with these credentials inherit the permissions that are granted to the local attribute value or role. Of course, this requires the mapping function to be configured such that it can map new types of attributes, but this is a tractable problem that can be tackled by using SAML metadata for expressing attribute requirements.

Regarding LoA attributes:
In SAML the LoA can be communicated as an attribute or via the SAML Authentication Context ability. Since multiple identity assurance frameworks exist, they should be conveyed in conformance with one of them. Work is underway to setup a global IANA registry for assurance profiles (LoAs). This will allow every community to define their assurance profiles and register it according to RFC 6711. SURFnet could register their assurance profile(s) at this registry or it can decide to make use of assurance profiles registered by others.