NameIdentifiers

A name identifier, represented by the `<NameIdentifier>` element in SAML1 and the `<NameID>` element in SAML2, is a direct way to name the subject of a SAML assertion. Name identifiers can be anything: an email address, a Kerberos principal name, a certificate subject, an employee ID, a username, or literally anything else. SAML 2.0 also defines more specialized identifier types with particular properties that were presumed useful in federated applications.

Unfortunately this generality makes interoperability much more complex than one would prefer.

Strictly speaking, SAML assertions don't have to contain a name identifier. The subject may be implicitly identified as the bearer of the token or anybody able to demonstrate possession of a key. In SSO use cases, one reason for including an identifier is to enable the relying party to refer to the subject later, such as in a query, or a logout request. So-called "transient" identifiers that are generated uniquely for each assertion are often used to support those use cases and are a common pattern in Shibboleth deployments.

Every name identifier is associated with a *format*. Formats label the identifier at runtime to help applications process them appropriately. They're conceptually similar to an Attribute Name and in fact one conventional way to express a SAML Attribute as a name identifier is to encode its Name as a Format (assuming the Attribute Name is a URI).

Though this is a retroactive view of the design, Name identifiers can be described by the following characteristics:

- **persistent** - whether a given name identifier is intended to be used across multiple sessions. An identifier intended to be used for a single session only is called a *transient identifier*.
- **revocable** - whether a given name identifier can be revoked. An identifier that persists over the entire lifetime of a subject's relationship with an IdP is called a *permanent identifier*.
- **reassignable** - whether a given name identifier, once revoked, may be reassigned to a different subject
- **opaque** - whether a relying party can positively identify the subject from a given name identifier. (A UUID is an example of an *opaque* identifier.) An identifier that can be used to positively identify the subject is called a *transparent identifier*. Many email addresses and network login IDs (such as `eduPersonPrincipalName`) are transparent when derived from a subject's name.
- **targeted** - whether a given name identifier is intended for a specific relying party (or parties) and not for anyone else. An identifier that is not targeted is a *shared identifier*. An identifier targeted at a specific affiliation of relying parties is also a shared identifier. An identifier targeted at a single relying party is not shared.
- **portable** - whether a given name identifier is usable across security domains.
- **global** - whether a given name identifier value is globally unique. However, a name identifier may be "qualified" to ensure global uniqueness. Typically, the qualifier is the identifier of the issuer or a DNS domain associated with the issuer.

A special type of globally unique identifier is a *scoped attribute*, which has the form *`userid@scope`*. In practice, the scope value is a DNS domain, which ensures global uniqueness.

Here are some examples (not all of these are actually encoded as SAML name identifiers, some are defined solely as Attributes):

Identifier / Attribute	Persistent	Revocable	Reassignable	Opaque	Targeted	Portable	Global	Qualifier
SAML2 Transient NameID	No	N/A	N/A	Yes	N/A	N/A	Yes	N/A
SAML2 Persistent NameID	Yes	Yes	No	Yes	Yes	Yes	No	Issuer ID
SAML2 Subject ID Attribute	Yes	Yes	No	Yes	No	Yes	Yes	Scoped
SAML2 Pairwise ID Attribute	Yes	Yes	No	Yes	Yes	Yes	Yes	Scoped
eduPersonTargetedID	Yes	Yes	No	Yes	Yes	Yes	No	Issuer ID
eduPersonPrincipalName	Yes	Yes	Yes	No	No	No	Yes	Scoped
eduPersonUniqueId	Yes	Yes	No	Yes	No	No	Yes	Scoped
Social Security Number	Yes	No	N/A	No	No	Yes	No	US Citizens
Phone Number	Yes	Yes	Yes	No	No	No	Yes	N/A
OIDC public `sub` claim	Yes	Yes	No	N/A	No	No	No	Issuer ID
OIDC pairwise `sub` claim	Yes	Yes	No	N/A	Yes	No	No	Issuer ID
ORCID	Yes	Yes	No	Yes	No	Yes	Yes	N/A

Notes:

1. The SAML2 Persistent name identifier and the eduPersonTargetedID attribute are functionally equivalent. Indeed, the value of the latter is precisely a SAML2 Persistent `<NameID>` element.
2. The SAML2 Persistent name identifier (and hence eduPersonTargetedID) are portable in the sense that any issuer can assert a known SAML2 Persistent `<NameID>` element. For example, a SAML2 Persistent `<NameID>` can transit a SAML IdP Proxy as-is, without modification. The same applies to the newer SAML2 Subject ID and Pairwise ID Attributes.
3. The SAML2 Persistent name identifier and the OIDC pairwise `sub` claim differ with respect to the portability characteristic only. In particular, the `sub` claim can not transit a gateway since the `iss` claim is required for global uniqueness.
4. A Phone Number is not universally portable but within the US, Phone Number is indeed a portable identifier. In fact, it is one of the few portable identifiers with no qualifier.

Attributes vs. Identifiers

In SAML, subjects are also commonly described with Attributes. In contrast to name identifiers, SAML Attributes can have multiple values and aren't necessarily usable as identifiers, but any name identifier can usually be expressed as an Attribute.

Shibboleth deployments traditionally have focused on the use of Attributes to describe subjects, and default to the use of transient name identifiers (or omitting them). Commercial SAML deployments less commonly make use of Attributes and tend to use loosely or improperly specified name identifiers.

The properties above used to describe name identifiers also apply to attributes when those attributes are themselves unique identifiers for a subject. Of course, many attributes are not identifiers at all, merely data of various kinds.