why-is-eptid-deprecated

eduPersonTargetedID was originally conceived as a privacy-preserving identifier to correlate users between an identity provider (IdP) and a specific service provider (SP) in a SAML SSO transaction. Over time, the variations among IdP implementations of eduPersonTargetedID have made reliable use of eduPersonTargetedID problematic:

Case-sensitive vs case-insensitive string comparisons

eduPersonTargetedID derives it's original syntax from the SAML V2.0 Name Identifier format of "urn:oasis:names:tc:SAML:2.0:nameid-format:persistent" (see http://www.oasis-open.org/committees/download.php/35711). One of the traits it inherits is that an eduPersonTargetedID value is case-sensitive. However, implementation gaps and data entry lapses frequently make accurate matching on eduPersonTargetedID problematic.

Its intended replacement, the SAML Pairwise Subject Identifier, addresses this gap by making the value explicitly case-insensitive. It also further clarifies the value's format/syntax to eliminate matching problems.

eduPersonTargetedID is often not stored in LDAP

Unlike most of the user attributes in use in InCommon, an eduPersonTargetedID is frequently generated on the fly and not stored in an LDAP directory. Since it is generally only used during a SAML SSO transaction, it makes sense to define the identifier in a SAML profile instead of as a part of the LDAP object class.

We need a more widely accepted identifier definition

An correlating identifier is fundamental to any SAML SSO transaction, not only during research and educational collaborations. We need an identifier that is adopted across industries: a SAML pairwise subject identifier defined in a SAML profile.

Transitioning to SAML Subject Identifiers

Case-folding argument

The work to define the SAML v2.0 was informed by experience with how identifiers have been handled as case-insensitively when defined as case-sensitive (See Section 2.1 of https://docs.oasis-open.org/security/saml-subject-id-attr/v1.0/cs01/saml-subject-id-attr-v1.0-cs01.html). The ambiguity [*e.g.*, whether XyZ@example.org is "the same" as xyz@example.org] requires applications to perform equality matches in contravention of the specification for broad interoperability and opens the possibility that a system assigning identifiers per the specification, taking into account the case sensitivity, could create identifiers that collide.

FROM OASIS SPEC (ka Scott C)-- In addition, it has come to light that many, if not most, applications have a predisposition to handle identifiers case-insensitively, partly due to a long-standing, though factually untrue, assumption that e-mail address mailbox names are case-insensitive data. SAML's "persistent" NameID definition explicitly requires case-sensitive handling, making them impossible to use safely with such applications without resorting to additional layers of profiling. Note that any other specification promulgating [case-sensitive] identifiers is potentially unsafe in combination with such applications and should be used with caution.

Choosing the right user identifier

- SAML General Purpose Subject Identifier (subject-id)
- SAML Pairwise Subject Identifier (pairwise-id)
- eduPersonTargetedID (eptid)
- Why is email address not an appropriate user identifier?
- eduPersonUniqueID
- Why is eduPersonTargetedID deprecated?
- eduPersonPrincipalName (eppn)
- Understanding Federated User Identifiers

Related content

- An Introduction to User Data
- R&S Explained in Plain English
- Understanding Federated User Identifiers
- Why is email address not an appropriate user identifier?
- Why is eduPersonTargetedID deprecated?

Get help

Can't find what you are looking for?

help Ask the community