A Note about Interpreting this Document

March 18, 2024:* *This is a community consultation draft document.

*The InCommon Technical Advisory Committee's SAM2Int/Entity Category Deployment Guidance Working Group has produced a series of deployment guidance to help InCommon Federation adopt/deploy support for the REFEDS Anonymous Access, Pseudonymous Access, and Personalised Access Entity Categories (we refer to them together as the * Access Entity Categories).*

This is a Three-in-One Document

These guidance materials are organized in three loosely connected volumes: 1. Understanding the Access Entity Categories, 2. Deployment Guidance for InCommon Participants, and 3. Working with Attributes required by these categories. They are joined together in a single document to facilitate community review. In their final published format, the topics will be parsed into a series of web articles cross-linked among each other.

More are Coming

We are aware that the InCommon community will likely need additional detailed guidance, for example, around migration strategies. A new TAC working group is forming to develop these additional materials. We welcome your input and participation.

Content

About the REFEDS Access Entity Categories

In 2023, REFEDS published the latest revisions of three attribute release entity categories designed to facilitate privacy-preserving, standard, and streamlined user information release in federated transactions. These are Anonymous Access, Pseudonymous Access, and Personalized Access categories. See Understanding the REFEDS Access Entity Categories.

The InCommon Federation (InCommon) endorses and strongly encourages the widespread adoption of these categories when requesting and releasing user information in federated transactions. Specifically, InCommono recommends two ways to use these categories:

Adopt the categories as intended - These entity categories are designed to facilitate streamlined access to resources by allowing an identity provider (IdP) to configure automatic attribute release to any qualifying service provider (SP) in the federation. We recommend all InCommon IdP's to support these categories. We also recommend that whenever possible, all InCommon service providers declare their attribute requirements using one of these 3 categories.

Using these categories as default attribute bundles - Where automatic attribute release isn't feasible, we recommend that IdPs use the attribute bundles defined in these categories as default attribute bundle templates in their IAM integration process. An SP in the federation should always support attributes defined in these bundles when integrating with InCommon identity providers.

Volume I: Understanding the REFEDS Access Entity Categories

InCommon’s Attribute Release Recommendations

User Attribute	Personalized	Pseudonymous	Anonymous
user identifier (subject-id)	✅	🚫	🚫
pseudonymous pairwise user identifier (pairwise-id)	🚫	✅	🚫
person name (displayName, givenName, sn)	✅	🚫	🚫
email address (mail)	✅	🚫	🚫
organization (schacHomeOrganization)	✅	✅	✅
affiliation (eduPersonScopedAffiliation)	✅	✅	✅
assurance (eduPersonAssurance)	✅	✅	🚫

Legend

✅ Required by category

🚫 Not allowed in category

What about eduPersonEntitlement?

While not a required attribute in these categories, eduPersonEntitlement is also discussed in the context of releasing authorization support information. See Authorization for additional information.

The Personalized Access Category

The REFEDS Personalized Entity Category registers Service Providers that have a proven need to receive a small set of personally identifiable information to effectively provide their service to the user or to enable the user to signal their identity to other users within the service. The Service Provider must be able to effectively demonstrate this need to their federation registrar (normally the Service Provider's home federation) and demonstrate their compliance with regulatory requirements concerning personal data through a published Privacy Notice.

See: REFEDS Personalized Access entity category

In the InCommon Federation, a Service Provider must qualify as a REFEDS Research & Scholarship Category Service Provider to qualify as a Personalized Access category Service Provider.

The Pseudonymous Access Category

The REFEDS Pseudonymous Access entity category enables authenticated, privacy-preserving federated access where a Service Provider requires proof of successful authentication, and offers personalized user experience, but does not require any additional personal information that would identify the individual accessing the resource. The Pseudonymous Access category achieves this via the use of a pseudonymous user identifier (pairwise-id).

See: REFEDS Pseudonymous Access entity category

Common uses of this category include anonymous access to licensed content where the service wishes to allow the user to save settings.

In the InCommon Federation, any Service Provider (SP) may register as a Pseudonymous Access Category SP.

The Anonymous Access Category

The REFEDS Anonymous Access entity category enables anonymous access to a restricted resource in a way that adheres to privacy and data protection regulations. It enables a Service Provider to require proof of successful authentication, and receive information about the individual's relationship to the identity provider organization, but not receive any personal information that would identify the individual accessing the resource.

See: REFEDS Anonymous Access entity category

Common uses of this category include anonymous access to licensed content (library, online journals, etc).

In the InCommon Federation, any Service Provider (SP) may register as an Anonymous Access Category SP.

Volume II: Deployment Guidance

for Identity Providers

When developing an adoption plan, InCommon IdP operators should adopt the following two-part deployment strategy:

Part I: Implement the basics - all InCommon IdP should support the required attributes named in the categories

Whether your IdP can automatically release attributes based on an SP's entity category, your IAM operation should be ready to support every attribute named in each of the three categories. Doing so establishes a common vocabulary to communicate user information among InCommon registered services. Further, use the guidance provided in Working with Required Attributes to make sure your interpretation of these attributes is consistent with the InCommon community's expectations.

As you implement support for these attributes, consider using the three categories as basic attribute bundle templates in your IdP configuration. Whether you support the automatic release mechanism required by the REFEDS entity categories or not, you can at least use these templates to standardize attribute release to individual SPs.

Part II: Scaling support

In parallel, work with your organizational data stewards to support the entity categories, i.e., enable automatic attribute release using the entity category syntax to qualified service providers.

for Service Providers

Each InCommon Service Provider operator should implement processes to determine its services' user information needs. Based on that assessment, determine the privacy characteristics that apply to your SP; if applicable, declare your SP as one of the three Anonymous Access, Pseudonymous Access; or Personalized Access. Where applicable, plan appropriate migrations.

Within the InCommon Federation, an SP needs to qualify as a Research & Scholarship SP to register as a Personalized Access category SP; conversely, a current R&S SP should register as a Personalized Access SP and plan appropriate migrations from R&S to Personalized.

> **My SP has varying user information needs…**

> If your platform represents multiple resources with different data needs, it’s a strong indicator that you should register multiple SAML SP entities in the federation.

When requesting basic user information, an SP should use the attributes mentioned in these categories. Some of the attributes are more complex to work with than might be expected. Make sure to follow the guidance provided in Working with Required Attributes to ensure your interpretation of these attributes is consistent with the InCommon community’s expectations.

for Federation Operator

- Update tooling, documentation, and processes to drive the adoption described above.
- Engage international R&E federation to iron out EC-based release governance and mechanics
- https://wiki.refeds.org/display/ENT/Requirements+for+Federations+Operators+Assessing+Access-Related+Entity+Categories

Volume III: Working with Required Attributes

user identifier (subject-id)

The **subject-id** attribute, or SAML General Purpose Subject Identifier, is a single-valued, unique value used to identify an individual user. A subject-id is intended to be both globally unique and correlatable across system domains.

A subject-id consists of a left-hand side (a case-insensitive identifier value with a Very Constrained character set) and a right-hand side (a domain, or scope), separated by the '@' character.

> **There is a technical definition for "Very Constrained"**
>
> "VERY CONSTRAINED" is
>
> ```
> <uniqueID> = (ALPHA / DIGIT) 0*126(ALPHA / DIGIT / "=" / "-")
> ```
>
> where "=" is the padding in the base 32 alphabet,
> and "-" is to support UUIDs;
> thus, base 32 encoding of another value could be suitable.
>
> More on Base32 Encoding: https://en.wikipedia.org/wiki/Base32

See: SAML V2.0 Subject Identifier Attributes Profile Version 1.0

Guidance for Identity Provider

Longevity and Uniqueness

A subject-id is designed to be a unique identifier representing a person in systems across potentially many organizations. Once issued and shared, it becomes very difficult to change. Therefore, the most crucial property of a subject-id is its stability; avoid populating it with values that are likely to change in the course of normal business processes.

Remember: anytime you change a person's subject-id, you are taking on a substantial change coordination effort to update all service providers you integrate with to update their records as well. Failing to do so will likely cause access problems for that person.

Reuse existing identifiers when appropriate

Start by carefully reviewing the subject-id's definition. Do you have an existing identifier that meets the subject-id's requirements?

If so, consider reusing that identifier by configuring your IdP attribute release mechanisms to send that value as a subject-id as well as its original intended attribute. This approach allows you to support subject-id in your IdP quickly.

A commonly used identifier in InCommon is eduPersonPrincipalName (ePPN). The following checklist may help you determine whether your ePPN (or any identifier) is a suitable identifier to reuse as subject-id:

- Our ePPN is case sensitive, i.e., JOHN@domain and john@domain represent two different people.
- We allow the user to petition to change (parts) of their ePPN, e.g., our ePPN is <net-id>@<domain>, and we allow a user to change their <net-id>
- We re-assign ePPNs, i.e., we re-assign net-id, so two different people might have the same ePPN over time.
- We know our institution is about to change its name, and the domain we currently use will no longer be valid.

If you answered “Yes” to any of the questions above, your ePPN is a poor candidate as a subject-id. Do you have another identifier that would allow you to answer “No” to all of those questions?

Start Now

Introducing a new identifier in an IAM ecosystem is challenging. It is much more so to introduce a new identifier across a large community. We need everyone to start now.

If you have an existing identifier you can reuse, configure your IdP to release subject-id now. You are well ahead of the curve and are well-positioned to help the community widen support for these new attribute release categories.

If you don’t have an existing identifier, start devising plans to introduce one in your IAM system. Engage the InCommon community in conversation. Share your ideas and challenges. Make the community work for you.

Lending / Getting Help with subject-id Migration

We understand that introducing and migrating to new identifiers can be a complex and time-consuming challenge. To achieve widespread adoption of these categories, we believe that we must introduce a cohesive and comprehensive identifier migration plan in 2024. We need your input and help to make that happen. Stay tuned for a call for participation in 2024.

Guidance for Service Provider

Compared to other unique identifiers (eduPersonPrincipalName, eduPersonUniqueID, etc.) in use today, subject-id’s definition clears up syntax ambiguities, improves uniqueness, and generally facilitates its use by an SP. In particular, it is designed for case-insensitive comparison, has a defined size, has a limited character set, and is expressed in a form that is easy to store and display, but still globally unique.

subject-id is Atomic

When processing a subject-id, an SP must ensure that the entire subject-id string is treated as an atomic unit. While parts of a subject-id value have meaning, a subject-id should never be split into separate parts (left of @ and right of @) when stored. This is similar in concept in the treatment of a social security number (SSN). While parts of an SSN have meaning (area, group, serial number), an SSN is always stored as an atomic value.

Verify the Issuer

The domain (aka scope) part of a subject-id indicates the identifier's issuing organization. Before accepting a subject-id, an SP must verify that the IdP issuing a subject-id is authorized to issue identifiers using that scope by verifying that the identifier's domain appears in a `<shibmd:Scope>` extension in the IdP's SAML metadata.

Lending / Getting Help with subject-id Migration

We understand that introducing and migrating to new identifiers can be a complex and time-consuming challenge. To achieve widespread adoption of these categories, we believe that we must introduce a cohesive and comprehensive identifier migration plan in 2024. We need your input and help to make that happen. Stay tuned for a call for participation in 2024.

pseudonymous pairwise user identifier (pairwise-id)

The "pairwise-id" attribute is a SAML-defined “identifier” (that is, a single-valued, unique value used to identify an individual user) used to establish a consistent and privacy-preserving relationship between an identity provider (IdP) and a service provider (SP) for a specific user.

The pairwise-id value is generated by the IdP and is unique to the combination of the user and the SP. It prevents different SPs from correlating and linking a user's activities across multiple service providers. This helps protect user privacy and prevents the creation of comprehensive user profiles by aggregating data from different SPs.

By assigning a distinct and unique identifier to each user and SP combination, the IdP can provide a consistent user experience while minimizing the sharing of personal information between SPs.

When a user authenticates with an IdP and requests access to a specific SP, the IdP produces a pairwise-id for that specific user-SP relationship. The SP can use this identifier to recognize and provide personalized services to the user without being able to identify the user across different SPs. Of course, the same identifier must be produced for subsequent exchanges between that IdP and SP for a given user.

> **See: SAMLV2.0 Subject Identifier Attributes Profile Version 1.0**
>
> https://docs.oasis-open.org/security/saml-subject-id-attr/v1.0/saml-subject-id-attr-v1.0.odt

The format of this attribute is very precisely constrained. It is scoped (see also eduPersonScopedAffiliation), consisting of a left-hand side (a case-insensitive identifier value with a very constrained character set) and a right-hand side (a domain), separated by the '@' character.

Guidance for Identity Provider

Lending / Getting Help with pairwise-id Migration

We understand that introducing and migrating to new identifiers can be a complex and time-consuming challenge. To achieve widespread adoption of these categories, we believe that we must introduce a cohesive and comprehensive identifier migration plan in 2024. We need your input and help to make that happen. Stay tuned for a call for participation in 2024.

Guidance for Service Provider

Implementation Strategy

In contrast to older approaches to solving this problem, the "pairwise-id" attribute has several important properties to facilitate its use by SPs. In particular, it is designed for case-insensitive comparison, has a defined size, has a limited character set, and is expressed in a form that is easy to store and display, but still globally unique.

However, it is crucial for SPs handling this attribute to ensure that the value and scope are manipulated and stored as a unit, never split into separate parts. It is also crucial to ensure that identifiers are only accepted if they are asserted by an IdP authorized by some form of policy to assert a particular scope. Failure to do so may result in impersonation risks.

Lending / Getting Help with pairwise-id Migration

We understand that introducing and migrating to new identifiers can be a complex and time-consuming challenge. To achieve widespread adoption of these categories, we believe that we must introduce a cohesive and comprehensive identifier migration plan in 2024. We need your input and help to make that happen. Stay tuned for a call for participation in 2024.

person name (displayName, givenName, sn)

There are three common LDAP attributes historically mapped into SAML to express a person's name (legal or otherwise).

The “givenName” and “sn” attributes are used to express the traditionally Western concepts of “given” and “family” names, respectively. The primary value of separating the fields is to allow applications to control the sorting of name information.

One disadvantage is that not all cultures treat names the same way, and people may not always have a first or last name to populate. The “displayName” attribute is traditionally a way to allow a full name to be expressed without artificial constraints placed on the formatting, but it lacks standardization around the ordering of individual portions of the name. Leading with a family name is better for sorting, but looks more awkward when used in other contexts.

Lacking any perfect solution to this problem, providing all three of these attributes as a group is the best option we have.

Guidance for Identity Provider

Implementation Strategy

While there are few absolute constraints on these attributes, one notable difference in LDAP is that “givenName” and “sn” are multi-valued and “displayName” is not. This stems from the historical purpose of LDAP, which was a search. Many SPs are not likely to handle multiple values for these attributes well, and it is best to limit them to a single value when possible.

Notably, there is no constraint on whether these attributes should carry legal or so-called “preferred” name values, but experience has shown that very few applications need a legal name, and the most common purpose for these attributes tends to be greeting people or presenting lists of users, and preferred names tend to work better for these use cases. Having said this, it is obviously not ideal for users to have full control over the values of these attributes with no oversight, since that creates opportunities for mischief. Most organizations leverage the data sufficiently that minimal oversight is sufficient to prevent egregious problems.

With respect to ordering, it is suggested that “displayName” be used to carry names in “speaking order”. In other words, for Westernized names, the given name is followed by the family name. Other cultures may have different conventions.

It is inadvisable to populate these attributes (externally at least) with “fake” values to signal their absence. It may be common in source systems to find whitespace or a single period or other conventions used to satisfy the constraints of badly implemented applications when users do not have a particular name value. Do not expose these conventions in SAML; simply omit any attributes that would not have a value.

Of course, the release of these attributes should always be limited to services for which the real identity of the user is important and relevant (or, if the default, by acknowledging clearly that the IdP is not operated as a privacy-preserving service).

Guidance for Service Provider

Implementation Strategy

As noted above, applications should be aware that the ordering of “displayName” is not standardized. They should also be aware that “givenName” and “sn” may contain multiple values or none at all. While this makes building user interfaces difficult, assuming anything contrary to the definitions of these attributes is not a solution to that problem. Forgoing the use of the information outside of very limited contexts (e.g., greeting a user directly) may be the best course.

Of course, support for Unicode in these attributes is quite important, more so perhaps than with most of the other attributes one handles. Consult your software’s documentation for details on any special steps needed in this regard.

mail

The “mail” attribute is a user attribute defined in RFC4524 to carry a user’s email address. From RFC4524: "The `mail` (rfc822mailbox) attribute type holds Internet mail addresses in Mailbox RFC5321 form (e.g., `user@example.com`)."

Guidance for Identity Provider

While this attribute is formally multi-valued and does not specifically connote “officialness”, it is suggested for interoperability to limit this attribute to a single value, generally the user’s official email of record at the home organization. Including multiple values, or including self-asserted, external email addresses, while permissible, is likely to lead to interoperability challenges with a variety of SPs.

Guidance for Service Provider

When working with InCommon Participants, an email address should only be used as a means of contact. The “mail” attribute is not a suitable user identifier, and in particular, lacks stability at many organizations due to name changes and other vagaries of email system management.

Why is an email address not an appropriate user identifier?

Email address is a popular way to identify a user and their organizational affiliation in consumer-oriented federated access use cases. It is easy. Everyone has at least one email address from a consumer ISP or social media platform. Companies always issue an email address to their employees. One can often deduce which company a person works for from the domain in her email address.

Right?

As it turns out, those assumptions don’t always hold in the research and educational space. There are several reasons why you should not rely on an email address as a unique user identifier when handling federated access in InCommon:

1. **Life events and changes in affiliation/role lead to email address change** - A person’s interaction in the higher education community often spans a long time. During that period, the person’s relationship with the community evolves. For example, a person may be a learner, a teacher, a researcher, an employee, a donor, and/or a parent to a learner. Further, a name change due to life events can also trigger an email address change. Email address is not a reliable persistent identifier when correlating identities across federated systems. Changing email addresses doesn’t scale. Many systems consume it and it isn’t feasible to identify what systems need to be notified.

2. **Email address may be reassigned** - Institutions frequently reassign an email address when a person leaves the institution. In federated systems that rely on an email address as a user identifier, this can lead to the wrong person accessing resources owned by/assigned to another.

3. **Email address is not always assigned by the institution** - Some institutions allow parts of their user community to supply their preferred email address (bring-your-own-email) instead of requiring the use of an institutionally assigned email address. Services deployed in the higher education community should not assume the @domain portion of a person’s email address is a reliable indicator of a person’s affiliation with an institution. For example, one of the largest universities on the West Coast allows its students to supply their preferred email address. Over 60% of the students chose that option. Those who do so will not have a @university email on record.

4. **Email is not a guaranteed unique identifier** - Email is a means of contacting its owner/recipient. It is no different than a telephone number. Just as people share telephone numbers, email addresses can be shared. For example, a university’s policy may allow family members studying at the same university to use the same email address when communicating with that university. An email address is not guaranteed to be unique to an individual.

5. **Email address may not be validated** - An email address is a form of contact, not a user identifier. Depending on organizational practices around contact information validation, an individual’s email address may not be strongly validated. Unless the organization performs some type of proof-of-control confirmation for the email mailbox, a person can enter someone else’s email address as a contact. A Service Provider relying on the email attribute as a primary identifier is vulnerable to impersonation attacks. Since a higher education identity provider does not process an email address as a unique identifier, A service provider working with a higher education institution should not depend on the email address as a user identifier.

organization (schacHomeOrganization)

schacHomeOrgnization specifies a person’s home organization using the domain name of the organization.

> **See: Official Definition of schacHomeOrganization**
>
> https://wiki.refeds.org/display/STAN/SCHAC+Releases

Guidance for Identity Provider

Which domain do I use?

schacHomeOrgnization’s definition does not provide detailed information on how to interpret “a person’s home organization”. There are two basic interpretations:

Home Organization is a person’s primary “real-life” association - a person’s home organization is the organization they are primarily associated with.

Home Organization is the IdP operator issuing the user’s credentials - a person’s home organization is the organization operating the IdP issuing the user’s credentials.

This distinction may be important when an IdP is a shared service representing multiple organizations, e.g., a university system-wide IdP representing member universities in a system.

The decision on what home organization to display will likely be influenced by technical and nontechnical factors within your organization.

Domain must be registered in Scope

When sending a domain value in schacHomeOrganization, the domain must be registered in the `<shibmd:Scope>` element of the IdP’s SAML metadata.

When to use schacHomeOrganization

Because shacHomeOrganization can only be a single value, it will have limited use for shared IdP representing multiple organizations, especially if people consider themselves to be members of more than one of the organizations served by the IdP.

For all * Access Categories, InCommon IdP operators should release a value that is present in their scope(s) registered with InCommon, and is explainable within the organization.

What is the SCHAC schema?

SCHAC, or SCHema for ACademia, is a common person data schema designed to facilitate higher education inter-institutional data exchange. This schema was originally produced by the European TERENA Task Force on Middleware. It was transferred to REFEDS Schema Editorial Board for ongoing maintenance.

Guidance for Service Provider

Implementation tips and strategies

Verify against Scope - On receiving a schacHomeOrganization value, an SP must ensure the value is present in the `<shibmd:Scope>` element of the Issuer's published SAML metadata. Any non-matching value is considered an invalid claim and should be discarded.

Be mindful of schacHomeOrganization's limits - The schacHomeOrganization attribute is a single value attribute, capable of indicating only one organization to which a person is affiliated. In scenarios where an Identity Provider (IdP) operates as a shared service in a multi-institutional environment, an individual might have associations with multiple organizations in that environment. The specific interpretation of these values is at the discretion of the IdP operator.

affiliation (eduPersonScopedAffiliation)

eduPersonScopedAffiliation conveys an individual's affiliations within a specific domain within an organization. In federated access, the Identity Provider (IdP) operator transmits one or more values to a Service Provider (SP), communicating broad categories that signify a person's association with the organization. An eduPersonScopedAffiliation value consists of a left and right component, separated by an "@" sign.

The left component, representing affiliation, is one of the 8 defined values from the eduPersonAffiliation attribute. The right-hand side component (scope) in eduPersonScopedAffiliation designates the domain associated with the person's affiliation. The scope presented in an eduPersonAffiliation value should match the right-hand side (scope) of the person's eduPersonPrincipalName identifier in the same assertion. Nevertheless, IdP operators have the flexibility to employ additional scopes to denote a person's connection with a sub-unit (e.g., campus, college, academic medical center) within a larger organization.

For instance, when a university system's IdP serves multiple campuses within the system, the right-hand side component may indicate the specific campus or campuses with which the person holds defined affiliations. A person studying at campus A while employed at campus B in the same system would simultaneously have affiliations of student@campusA.edu, member@campusA.edu, employee@campusB.edu, and member@campusB.edu.

> **See: Official Definition of eduPersonScopedAffiliation**
>
> https://wiki.refeds.org/display/STAN/eduPerson+2021-11#eduPerson202111-eduPersonScopedAffiliation

Basic Implementation tips and strategies

Know your people - Have the ability to identify who is a faculty, who is a student, etc in your organization; Grouper is a great tool for managing these relationships.

Multiple affiliations - Within higher education, a person can, and often have multiple affiliations with an institution; a law professor (faculty, employee) may be pursuing an MBA degree (student); an administrator may have split appointments with two schools (e.g., staff@dentistry.acme.edu, staff@nursing.acme.edu). Make sure your IAM system can support multiple affiliations for a person.

Affiliation != Authorization - More precisely, there is no need to assume that these affiliations must directly translate to authorization to access any service. As an IdP, focus on conveying how a person is related to your organization. It is the SP's responsibility to build authorization decisions based on these relationships. If you do need to convey explicit authorization to a service or feature, eduPersonEntitlement is the attribute to use.

eduPersonScopedAffiliation is useful beyond these Access categories. Regardless of your support status for the three REFEDS access entity categories, support eduPersonScopedAffiliation so that when needed, you are ready to send that information to any SP you interoperate within individual SP attribute release policies.

How do I plan the "right-hand side" values?

The right-hand side of any scoped attribute value is a claim of scope/domain. It is an IdP's way of conveying that the value holder has a relationship with the organization represented by that scope/domain.

To make such claims, an IdP must have the authority to do so (i.e., an IdP from the University of Texas cannot make claims on behalf of England's Oxford University). To ensure such authority within the InCommon Federation, an IdP must register any scope/domain it uses in attribute assertions in the "Scope" element in its IdP metadata.

An IdP operator may determine at its discretion any number of scopes to use to represent a person's relationship with units within its organization. To keep things manageable, we recommend keeping the division at a fairly high level, e.g., school/college within a university, etc.

What are the valid "left-hand side" values and which of them do I need to implement?

eduPersonAffiliaion, therefore eduPersonScopedAffiliation, defines 8 types of affiliations: `faculty`, `student`, `staff`, `alum`, `member`, `affiliate`, `employee`, `library-walk-in`.

As a Service Provider, how do I interpret eduPersonScopedAffiliation values received from an IdP?

eduPersonScopedAffilation conveys a person's relationships to an organization. It is not meant to convey authorization to access specific services. While there are finite valid values defined in this attribute, A person's home organization ultimately determines the precise interpretation of those values (e.g., not all institutions define "student" the same way).

As an SP, if your access policy is compatible, (e.g., any member of an organization, as determined by that organization, can access your service), eduPersonScopedAffiliation is a simple and scalable way to enable access.

When you need more information to determine access or authorization…

The Access Entity Categories likely do not fit your situation. The more tailored eduPersonEntitlement is likely a good attribute for individualized service needs.

Question for consultation reviewer: how much more do we say here?

Configuring eduPersonScopedAffiliation for Anonymous and Pseudonymous Access

As Anonymous and Pseudonymous Access categories are designed for privacy-preserving access, always consult your local/regional policies before releasing an individual’s specific affiliation values. When policies allow, all applicable values should be released, but in particular, an IdP should always assert `member` or `affiliate` for any applicable individuals.

Configuring eduPersonScopedAffiliation for Personalized Access

When working with the Personalized Access category, an IdP should assert all applicable defined affiliation values of an individual.

About “member” and “affiliate”

> **Are you using “member” and “affiliate” correctly?**
>
> *from the eduPerson specification:*
>
> *“... "Member" is intended to include faculty, staff, student, and other persons with a full set of basic privileges that go with membership in the university community (e.g., they are given institutional calendar privileges, library privileges, and/or VPN accounts)... “*
>
> *“... The "affiliate" value … indicates that the holder has some definable affiliation to the university NOT captured by any of faculty, staff, student, employee, alum and/or member. Typical examples might include event volunteers, parents of students, guests, and external auditors…”*

The `member` value is meant to represent a person who has a close and active relationship with the organization. Specifically, `faculty, staff, employee,` and `student` are `member` of an organization. The IdP’s operator’s home organization policies determine who is a faculty, student, employee, or student and any ambiguity in those policies will also be present in the `member` value.

Note: A holder of the affiliation `alum` is not typically `member` since they are not eligible for the full set of basic institutional privileges enjoyed by faculty, staff, and students.

The `affiliate` value for eduPersonAffiliation indicates that the holder has some definable affiliation to the university NOT captured by any `faculty, staff, employee, student,` `alum`, and/or `member`. Typical examples might include event volunteers, parents of students, guests, and external auditors. An IdP organization determines who is an affiliate within its institutions.

Comparison with eduPersonAffiliation

eduPersonAffiliation should contain the same list of unique values as the “left-hand side” values present in eduPersonScopedAffiliation. As noted above, the left-hand side values are of limited use in the entity categories and are of even less use if the IdP represents multiple sub-organizations.

assurance (eduPersonAssurance)

The eduPersonAssurance attribute provides information about the level of assurance or confidence that can be placed in the identity of an individual. It helps determine the extent to which an individual's identity has been verified, authenticated, or authorized within an educational environment.

> **See: Official Definition of eduPersonAssurance**
>
> https://wiki.refeds.org/display/STAN/eduPerson+2021-11#eduPerson202111-eduPersonAssurance

The InCommon Federation uses eduPersonAssurance to convey the level of an IdP’s confidence in the subject’s real-world identity, as defined by the REFEDS Assurance Framework. There are a variety of assurance frameworks defined, usually by the government or industry bodies; the REFEDS framework was defined by the worldwide higher education community.

Guidance for Identity Provider

How do I use eduPersonAssurance?

The REFEDS Assurance Framework defines signals allowing an IdP to convey two sets of information:

- The IdP meets the conformance criteria outlined in the REFEDS Assurance Framework
- The extent to which the identity of the individual accessing a resource (therefore referenced in an authentication assertion) has been vetted

Conveying an IdP’s conformance with REFEDS Assurance Framework

The InCommon Baseline Expectations for Trust in Federation requires all IdPs registered in the InCommon Federation to meet requirements comparable to the conformance criteria in the REFEDS Assurance Framework.

An InCommon-registered IdP should always send the REFEDS Assurance Framework conformance identifier (https://refeds.org/assurance) when eduPersonAssurance is a part of an assertion, regardless of the individual's identity assurance level. This simply allows the SP to make the relevant inferences based on the other values supplied (or based on their absence).

Expressing an individual's identity assurance level

See REFEDS Assurance Framework Implementation Guidance for InCommon Participants

Guidance for Service Provider

This section is left blank pending InCommon's updated identity assurance guidance based on REFEDS Assurance Framework 2.0

Additional Discussion: Authorization

The Anonymous category and, to a lesser extent, the other two categories, all lack an effective and appropriate means of handling authorization as a use case, as noted in the various category specifications. The most suitable attribute for this purpose, eduPersonEntitlement [Ref] is "outside" the formal attribute bundles because it is generally not automatable, and the bundles are at their core meant to lead to a more automated release of attributes.

That said, there are scenarios where authorization can reasonably be automated without compromising privacy, and the commonly encountered "site-licensed access" contracts common to many library subscriptions and some other cloud services are one such example. Such contracts typically apply to "everyone affiliated with the organization", and there is a standard entitlement value defined for this purpose, "urn:mace:dir:entitlement:common-lib-terms" [Ref].

IdPs are therefore encouraged to support this entitlement value and to make it available when it applies along with the other required attributes, for all three bundles.

SPs with authorization use cases are encouraged to support eduPersonEntitlement for this purpose, and those with a compatible licensing model are encouraged to support the standard value noted above when applicable.