why-is-email-not-an-appropriate-user-identifier

Jump to:

Reason 1: Email is not a guaranteed unique identifier | Reason 2: Email address may be reassigned | Reason 3: Life events and changes in affiliation lead to email address change | Reason 4: Email address is not always assigned by the institution | Reason 5: Email address may not be validated

Email address is a popular way to identify a user and her organizational affiliation when processing federated authentication in consumer-oriented and business-facing applications. It is easy. Everyone has at least one email address from a consumer ISP or medial platform. Companies always issue an email address to their employees. One can always deduce which company a person works for from the domain in their email address.

Right?

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

Reason 1: Email is not a guaranteed unique identifier

Email is a means of contacting its owner/recipient. It is no different than a telephone number. Institutions may not enforce that multiple individuals not use the same email address when communicating with the institutions (domestic partners, siblings, etc. attending the same school may register the same email address to communicate with the university) and therefore the email address may not be strongly bound to the individual.

Reason 2: Email address may be reassigned

Institutions may 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*.

Reason 3: Life events and changes in affiliation 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.

Reason 4: Email address is not always assigned by the institution

Some institutions allow some parts of their user community to supply their own preferred email address (bring-your-own-email) instead of requiring the use of an institutionally assigned email address. *Services in the HE 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 own preferred email address. Over 60% of the students do so, therefore, do not have a @university email on record.

Reason 5: Email address may not be validated

The intent of an email address is as a means of contact. Depending on organizational practices, there can be cases where the email address is not strongly validated. Unless the organization performs some type of proof-of-control confirmation for the email mailbox, it is entirely possible for a person to enter someone else’s email address as a contact leading to the threat of an attack where one individual is able to impersonate another. Service providers that rely on the email attribute as a primary identifier must carefully consider the operating practices of the identity provider. Where it is not practical to evaluate the operating practices (for example, evaluating an entire trust federation), or the risk associated with a spoofing attack is unacceptable, the service provider should not use the email address as a user identifier.

Choosing the right user identifier

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