19 August 2020 Notes

Attendees:

- Scott Cantor
- Heather Flanagan
- Alan Buxey
- Benjamin Oshrin
- Jiri Pavlik
- Meshna Koren
- Chris Shillum

Notes

Recap from last call

- On the last call, we came to consensus that what we're looking for here is not an entitlement. It is not something that should be used for authorization purposes.
- The primary use case that kicked this off: n the publisher use case, when a user comes to access content, they will verify that the user authenticate and that they have an entitlement to the data. That happens in real-time. Separate to that, there is a different set of people (e.g., the people who have purchased the access to the report) and they want to understand the use of that resource and classify that into buckets. That report is offered outside the authentication/authorization transaction and at a later date.
- A potential alternate use case is for HPC, where there may be a need to report back on HPC usage

Discussion

- Is this also a billing use case? This could be usefulin multi-tenant services where the SP wants to know what contract the usage should be charged against.
    - Current assumption is that the code provides no semantic information to the SP. That means the SP cannot have billing information on its side that it would map to.
    - This code still could be used as an internal charge-back code. The IdP can pass whatever code it wants; the SP will bucket or segment their reporting data based on those values.
    - If this has to be meaningful to both the IdP and the SP, there has to be an out of band agreement between the two as to what the codes mean. This is outside the scope of this spec.
    - The spec should say: The SP should be able to deal with any value that conforms to the syntax without prior knowledge of what those values are.
    - Should specify that IdPs do not use confidential values.
- Discuss the semantics of what should be stored on the IdP and what should be sent to the SP
    - From a federated protocol perspective, we're trying to define an opaque bucket for data to go into with a name that is interoperable. It is not appropriate, therefore, to get into defining structure fo rate data to make it easier for the IdP do to the job. That's an internal issue for the IdP.
    - eduPerson is primarily, but not exclusively, an LDAP schema. Any time you store data with structure, LDAP doesn't really deal with this very well. If it's not a simple name-value pair, there's complexity which causes problem here. We may instead need to assume a computation is happening to generate this value.
    - If we do this in terms of an eduPerson attribute, would advise we take the approach of not considering this an LDAP attribute. We can offer guidance, but it isn't our main target. We should avoid rules as to what goes on the wire beyond length and character set.
    - The attribute in its form is an opaque thing with rules; if you choose to store it in LDAP here's a recommended way on how to do that. LDAP interop is not our use case. IdPs should be free to do this however best for their situation.
    - This value will be constructed from a set of attributes in LDAP; in some use cases it might correspond to a sub organizational identifier, or a demographic attribute, or something else.
    - Are these attributes supposed to be correctable or not? IdP organizations may want to aggregate this across SPs, so they have freedom to use the same value set across multiple SPs. The spec can be silent beyond stating this should not compromise individual user privacy.
    - There is a field in SAML that has the semantic we're looking for, which is SessionIndex.
    - We could follow what voPerson is doing, which is to create a companion document beyond the core spec that would define the implementations and/or considerations for LDAP
    - Note on whether we should follow case exact matching - suggest we should not require this as apps tend to not do this properly. Original thought is that case sensitivity should have been easier, but reality has proved this not to be true.
- Does this even belong in eduPerson, given the LDAP predilection?
    - The schema board is already considering how to move to a voPerson model and abstract out the implementation details from the core spec
- Consider if there are any other use cases where this code might be used
    - This approach might be better for demographic reporting than we see with affiliation right now, and allow more freedom in the report (no out of band agreement, no limit of vocabulary, breakdown of guest information). The IdP could do a lot more analysis without releasing identifiable info to the SP.
    - Could have relevance to the error handling call back
- Figure out a better name

    - Segment - makes sense if you’re an analytics person, but not if you’re a computer science person
    - Analytics code

Next Steps

- Scott Cantor (with whatever support Heather can offer) will work on a revised draft proposal that abstracts the core spec from LDAP requirements

Old Draft text:
https://docs.google.com/document/d/1HGmz39bVMOq5VU74bhCd1Uu0nV_Tq9JaPBU98Zm-Fe0/edit#heading=h.j6288fhwkrg0