Federated Identity Management for Libraries (FIM4L)

Charter for a Working Group/Task Force

Version 1.0 (June 2019)

Content:

Introduction

Libraries for long have been providing access to online publisher resources based on IP-based access: if the person wanting access was in the library, he/she was granted access. The user experience with IP-based access when the subject is on-campus/on-premise is unsurpassed. For off-campus/off-site access, IP-based access has many drawbacks.

Several developments have made it so that this access model is no longer found to be sufficient:

1. a shift by users from accessing resources only from within the library, to more and more requiring access from ‘any location’.
2. publishers want more control over access to control licensing and abuse (some people ‘downloading everything’, for instance to publish it on websites like Sci-Hub).
3. libraries want more control to allow for smarter (cheaper) contracts that for instance allow access for only certain groups or faculties, so as to agree on a lower price (since not ‘everyone’ can access the resource).
4. publishers want to be able to offer a more personalized experience to users.
5. libraries and users find using VPN-technology a lot of work and complex.

Looking for a way to solve this, most parties have agreed using federated authentication, from a library standpoint next to the option of IP-based access while on campus, is the way forward: a user uses an identity that allows an institution to testify whether he/she has a certain relation with them. This is expressed in the form of *attributes*, which can be used to check whether someone should be allowed access. From a technological point, federated and Single Sign-On (SSO) enabling technologies like SAML2 (like used by Shibboleth and other products) and OIDC/OAuth2 is currently used that all allow for attribute based access control.

While seeing the benefits, libraries want to protect users' privacy. To set up federated connections the right way, libraries and publishers should have a clear policy about how to deal with privacy related matters, and it would be very beneficial to have a best practice in place.

Problem statement

Problems perceived when working with attribute based access instead or in addition to IP-based access control:

1. There are many different situations.
    a. Some countries or libraries are afraid that when using another technology, it means releasing more personal data.
    b. Some publishers don't want to receive (many) attributes, while in some cases parties agree that their specific scenario requires some extra attributes to be released.
2. Managing access based on attribute release has two major pitfalls:
    a. the provider of the attributes does not release the correct set or correct values
    b. the provider releases more attributes than strictly necessary, violating the privacy of the user
3. For publishers and libraries, it is complex and expensive to manage access to a certain resource, when libraries (in different countries) require different attributes.
4. The way federated authentication is implemented by the various publishers differs a lot. This results in more confusion to our end user.

This causes libraries headaches about what their policy should be, and causes many discussions and delays in the contractual phase.

Workgroup aims

To improve the situation, the FIM4L workgroup aims:

- to come to a consensus on library policy for federated authentication that protects users identities. This policy should help libraries and publishers and needs to be clear for account managers, license managers, etc. (those who make the deals), while also including enough technical information for IT staff.
- to seek broad support for the policy amongst libraries and publishers.
- to promote the use of uniform implementations of authentication procedures by service providers

This workgroup is supposed to be a library-led initiative. We state this, as this could easily become dominated by technical specialists. We explicitly need this to stay library-led.

Planned activities:

- Collect libraries' requirements for Federated Identity Management (FIM) and input them to relevant groups like FIM4R, REFEDS, the eduGAIN community, STM and RA21.
- Draft guidelines and recommendations for attributes release that respect users privacy and allow single sign-on for personalisation with users consent
- Bring together all relevant stakeholders with an interest in progressing FIM-based authenticated access to e-resources instead of IP-address-based authentication in libraries (research and non-research)
- Increase the awareness of the existence of federated authentication amongst people responsible for purchasing electronic resources; advocate for making federated authentication a requirement during the negotiation phase (if/when possible)
- Engage with libraries in the discussion of the suitability of the RA21 recommendations regarding:
    - Central IdP discovery service (WAYF/DS) for seamless off-campus usage
    - user-friendly/technology agnostic terminology such as "Login via your Home Institution" instead of "Login via Shibboleth",or "via OpenAthens".
- Promote the adoption of state-of-the-art and privacy preserving Federated Authentication and Authorisation Infrastructure (AAI) and principles (including attribute release) by libraries from research and educational institutions in particular but also all other libraries in general, as they would benefit as well.
- Work together with REFEDS and RA21 on recommending a best practice on exchanging attributes between libraries and publishers
    - Which attributes? (E.g. a persistent identifier, affiliation, etc.)
    - What to do with additional (demanded) Personally Identifiable Information (PII) sign-ins for service personalization?

Scope

This document and initiative is about providing access to online licensed resources by entitled users of academic institutions and public libraries including walk-in users. It is concerned about the privacy aspect of access via federated identity and authentication.

Parties involved in this initiative

Although this initiative is library driven, there are participants of altogether 4 stakeholder groups:

- Libraries, institutions to which a library belongs and library associations
- Infrastructure providers (GEANT, NRNs and Computing centers)
- IT Consultancies
- Publishers

Related initiatives

The workgroup has identified the following related initiatives:

- **RA21**. That initiative focuses on a better login user experience. As they might also touch on the aspect of attribute release, we should keep in contact with them
- **REFEDS Research and Scholarship Entity Category (‘R&S’)** defines an attribute release scheme for research. See whether we can adopt that same scheme or use it as a basis for a scheme for libraries. Currently the R&S v1.3 ‘definition’ includes this text that excludes using it for access to publisher resources: “This Entity Category should not be used for access to licensed content such as e-journals”
- **The GÉANT Data Protection Code of Conduct (‘CoCo’)**. Information about attribute release for authentication in a scholarly and research context.
- **eduGAIN.** Federation of educational and research identity federations.