Authentication and Authorisation for Research and Collaboration

Introducing the Federated Identity Management for Libraries Initiative (FIM4L)

Eko-Konnect Users Conference 2020, Lagos, Jan. 29, 2020

Peter Gietz
DAASI International

Parts of the slides co-authored by

Jos Westerbeke
Erasmus University

Erasmus University Library

Jiří Pavlík
Moravian Library

MORAVSKÁ ZEMSKÁ KNIHOVNA

Library as a safe place

The library must

- provide a private (virtual) place
- protect its users

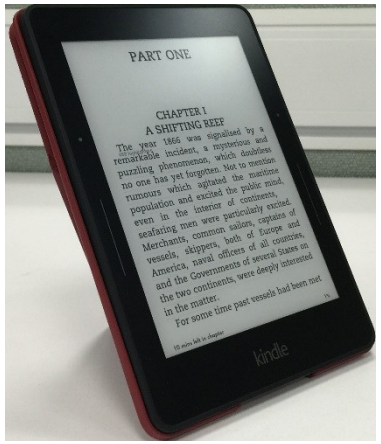
A trusted safe place with privacy

Library as digital content provider

late 20th century - early 21th century

Print -> Digital -> Remote

Remote = any time, any place, any device.

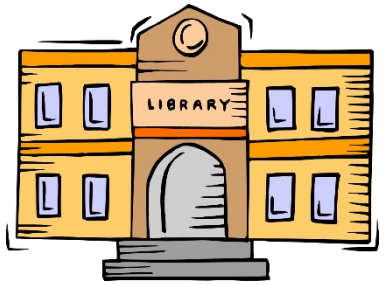


We want:

authenticated and authorized access

preserving privacy

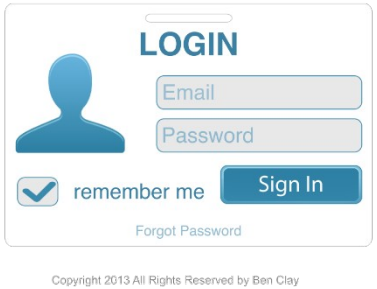
Access: from IP to SSO



IP based authentication:

location based

(Artificially adapted for any place by VPN and Proxy.)



SSO authentication:

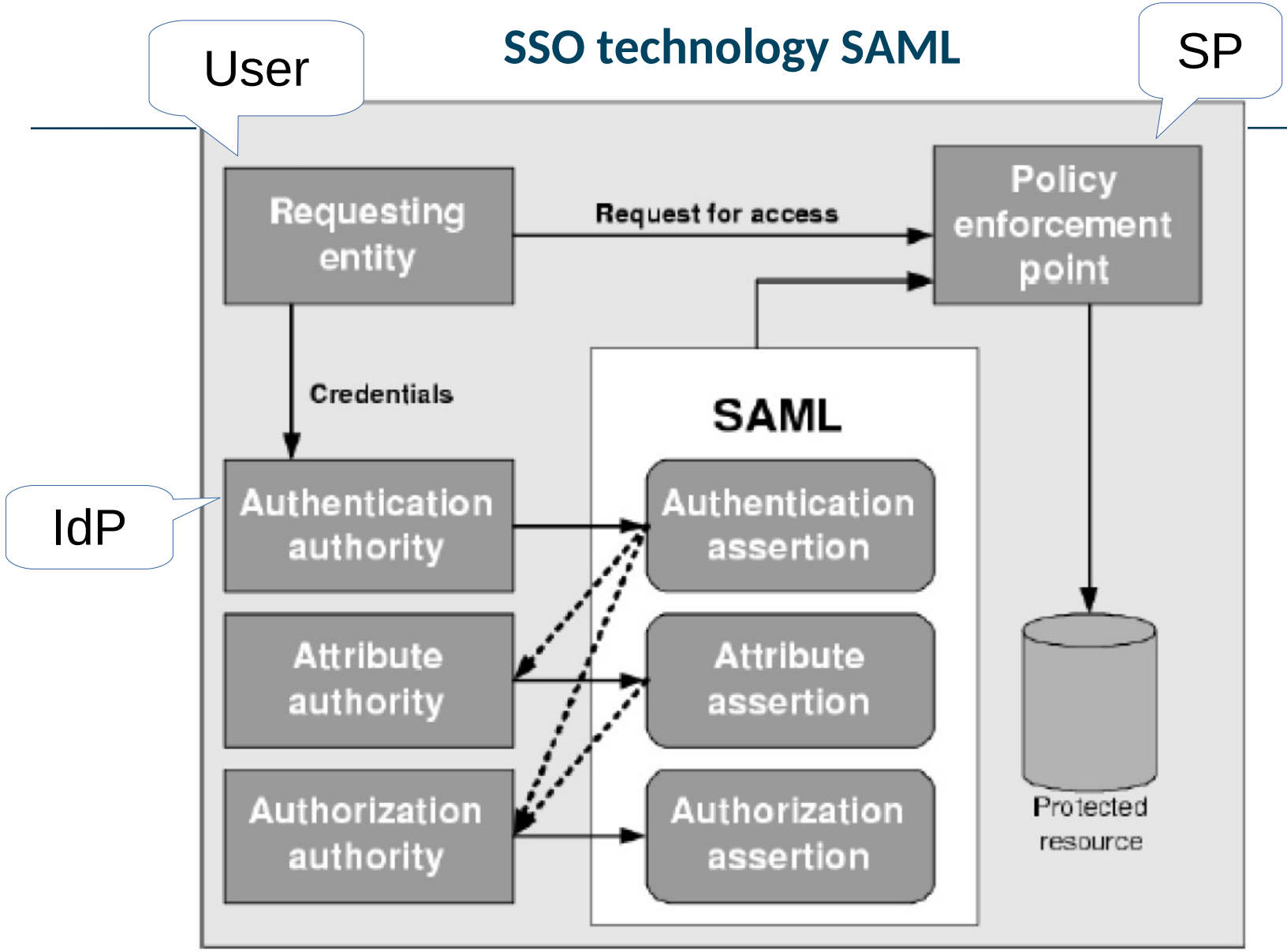
person(ID) based

(Any place, any time, any device)

SSO benefits

- tailored contracts: e.g. buy a subscription for one faculty.
- In case of abuse: just one user can be blocked, instead of all staff and students alike.
- better statistics

SSO technology SAML



(c) RUBENKING, NEIL J.: Securing web services

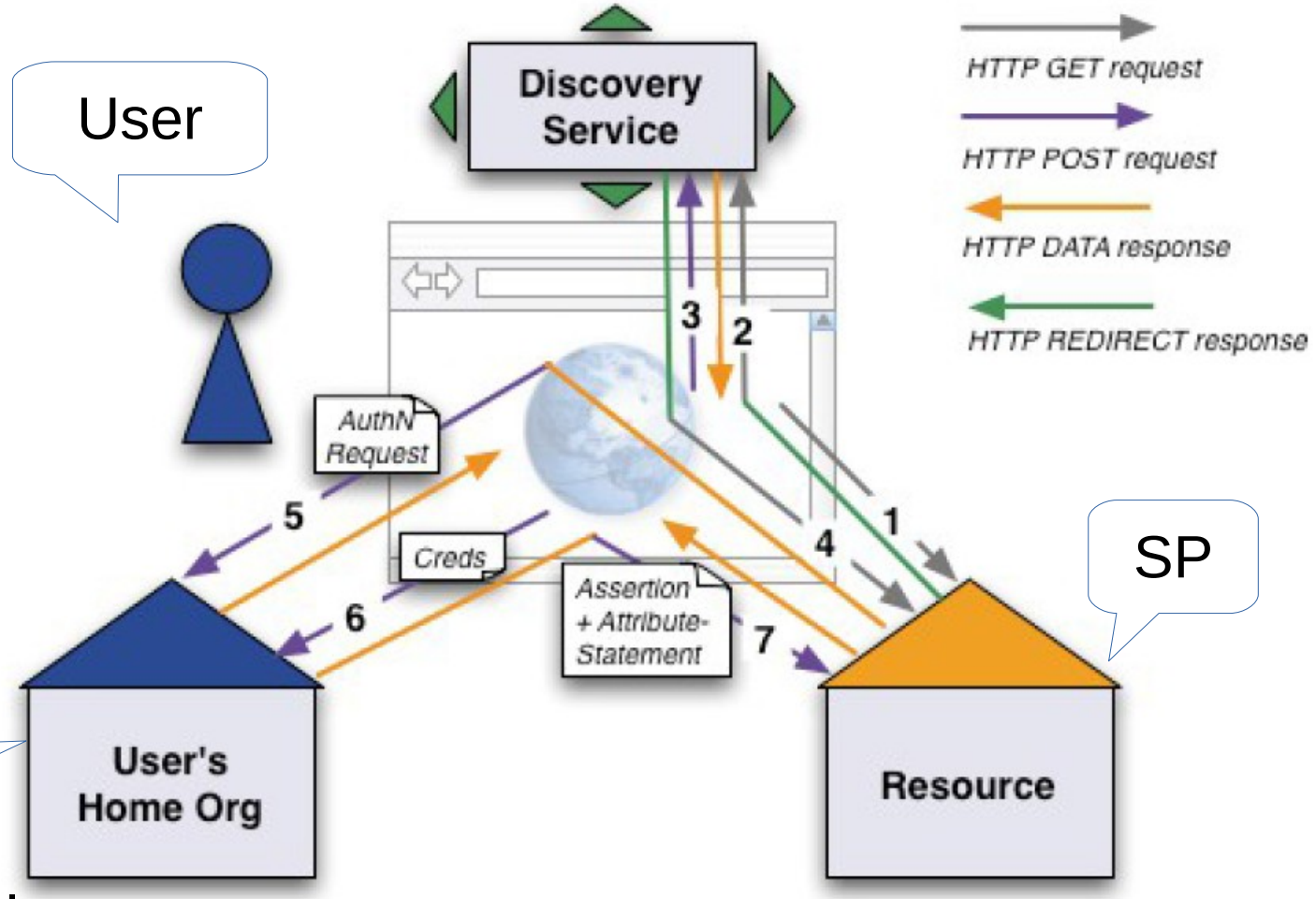
SAML: Security Assertion Markup Language

SP: Service Provider

IdP: Identity Provider

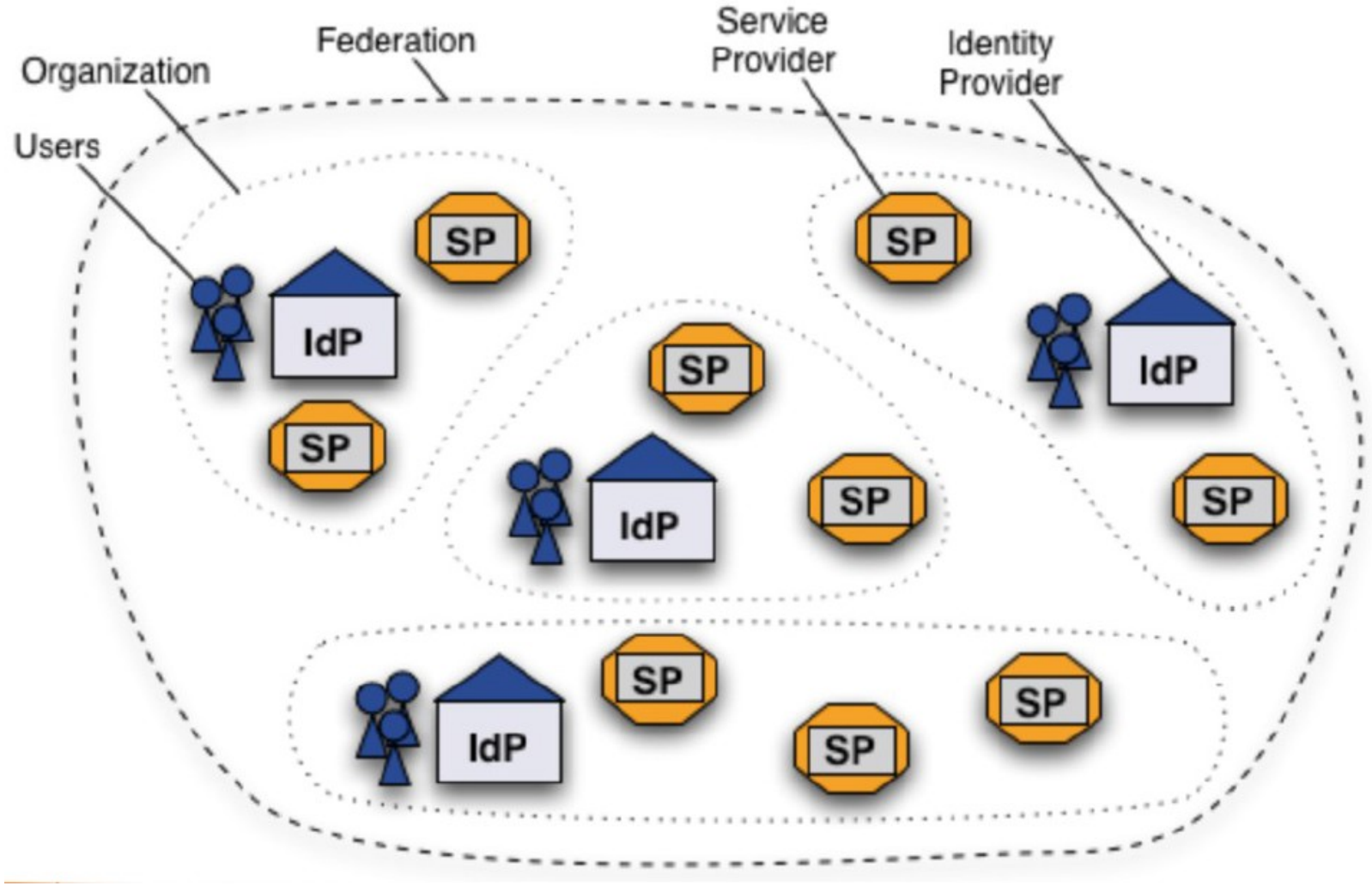
Attribute Assertions potentially contain privacy related information (name, email, etc.)

SAML implementation Shibboleth



SSO enabler

SAML based (national) Federation



What are relevant activities to look at?

The international federation of national research and education federations.

The research and education federations group. Which has a Data Protection Code of Conduct (CoCo)

An international working group on FIM for Research Infrastructures

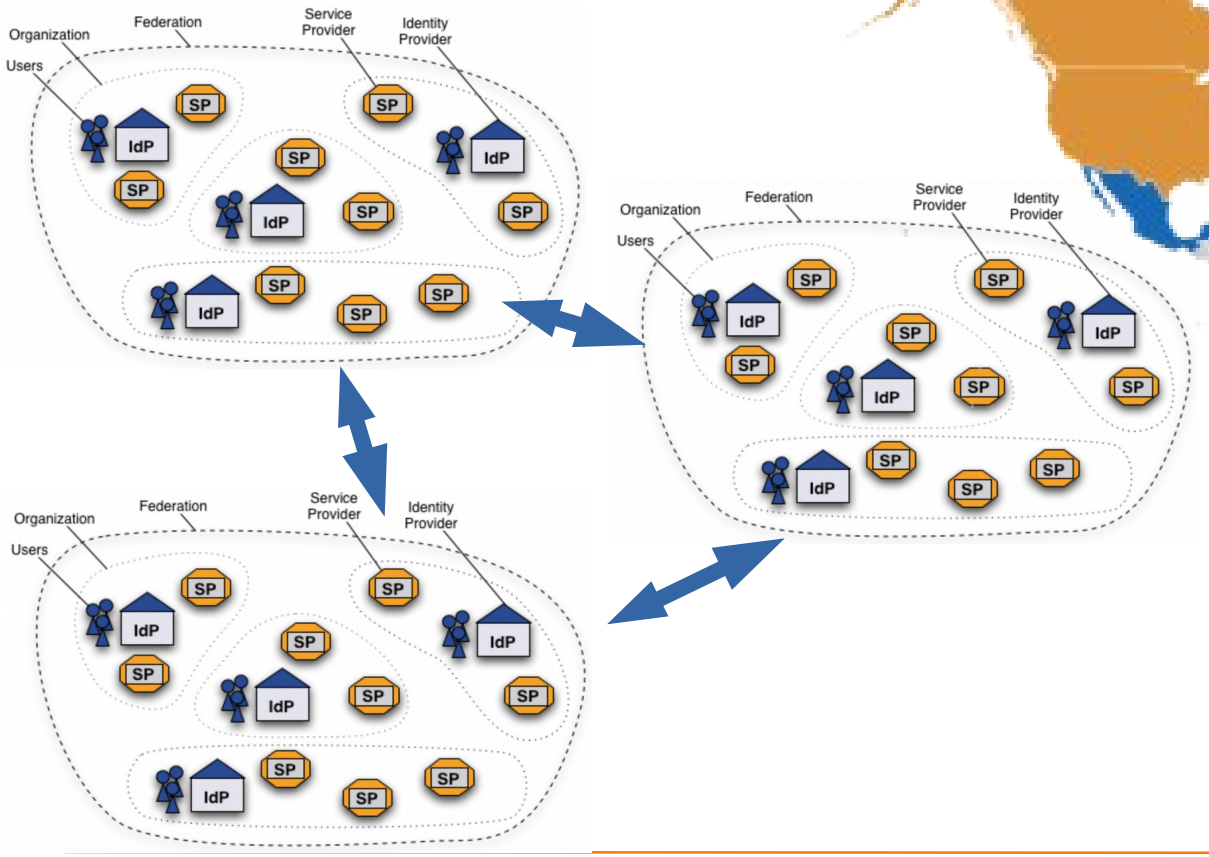
Former EU-funded project on FIM technologies. Providing Policies, trainings and the Blue Print Architecture

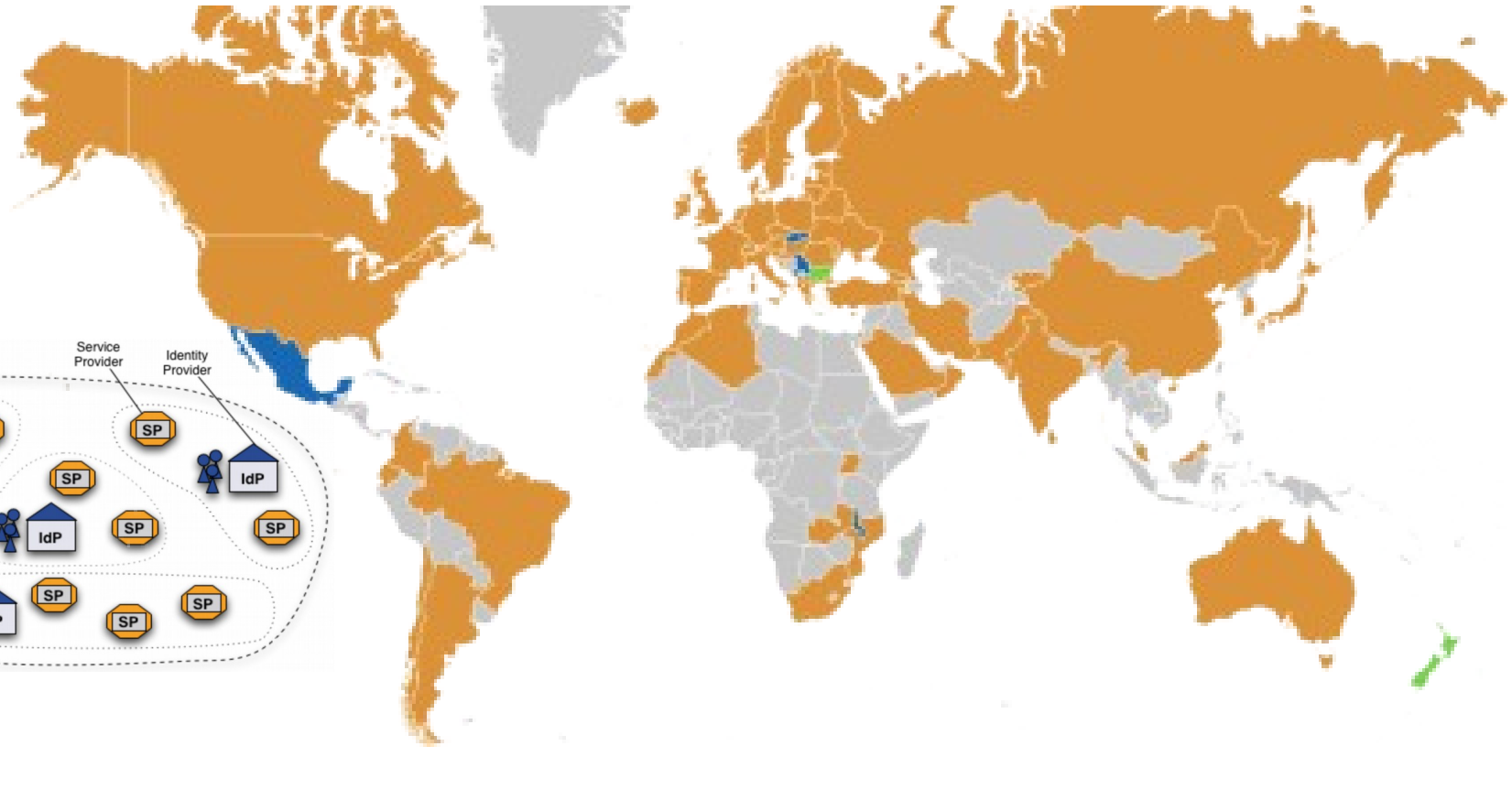
Activity on enhancing the user experience in federated identity management (FIM).

SeamlessAccess.org: Follow up from RA21 to further FIM based SSO

An international working group on FIM which was conceived to be library driven

SAML based Inter-Federation eduGAIN





AARC Project

- Authentication and Authorization for Research Collaboration
- EU funded Horizon 2020 project in two phases (2015-2017 and 2017-2019)
- Inspired by the work from FIM4R Group
- Objectives:
  - Deliver production-ready architectuaral building blocks and best practices to enable research and e-infrastructures to build interoperable AAI
  - Avoid a future in which new research collaborations develop independent AAIs
- One Work Package is dedicated to create pilot services
  - A number of AARC 1 pilots were dedicated to library use-cases
- See also https://aarc-project.eu/libraries/

What can libraries expect in the coming years?

- Increasing demand by publishers for SSO access
- Increasing denial by publishers for IP based access
- More awareness by users of personal data exchange
- Recommendations from initiatives such as RA21.org
- Recommendations from Library consortia

Federated IDentity management stake holder in library context

The library serves hundreds of publishers. A publisher like Elsevier serves thousands of institutions.

They work together through federations when it comes to Identity and Access Management (or IAM)

Libraries wants to negotiate contracts with the publishers while preserving privacy of their users

The publishers wants contracts and most probably also user data

Library representation

- Although AARC I had libraries as one user community there is no real representation of libraries interests in the FIM ecosystem
  - RA21 and SeamlessAccess.org can be seen as publisher driven
- Thus the Group FIM4L was established as an international working group which was conceived to be library driven
  - First without affiliation except that it can be seen as a spring-off of AARC
- FIM4L (Federated Identity Management for Libraries) is a library-led working group that aims to bring a seamless user experience and service set up, the latest standards and technologies and a focus on user privacy into federated authentication for library services.
  - FIM4L advocates minimal disclosure policy

FIM4L Member organisations

- Although FIM4L is library driven, since FIM is a process, in which many different stakeholders act, FIM4L has members out of the following stakeholder groups:
  - **Libraries**, institutions to which a library belongs and library associations
    - e.g. LIBER, Erasmus University Rotterdam, Moravian Library Brno, State Library Berlin
  - **Federation Infrastructure providers** (NRNs, NRN Associations, Computing centers)
    - e.g. GEANT, DFN, SURFnet, Eko-Konnect, UbuntuNet Alliance
  - **IT Consultancies** working on FIM technologies
    - e.g. OCLC, Spherical Cow Consulting, DAASI International
  - **Publishers** engaged in FIM technologies
    - Elsevier

FIM4L Member organisations from Europe, America and Africa

- Although FIM4L is library driven, since FIM is a process, in which many different stakeholders act, FIM4L has members out of the following stakeholder groups:
  - **Libraries**, institutions to which a library belongs and library associations
    - e.g. LIBER, Erasmus University Rotterdam, Moravian Library Brno, State Library Berlin
  - **Federation Infrastructure providers** (NRNs, NRN Associations, computing centers)
    - e.g. GEANT, DFN, SURFnet, Eko-Konnect, UbuntuNet Alliance
  - **IT Consultancies** working on FIM technologies
    - e.g. OCLC, Spherical Cow Consulting, DAASI International
  - **Publishers** engaged in FIM technologies
    - Elsevier

FIM4L LIBER Working Group

- The first library association that was responsive to the idea FIM was LIBER
- Thus we decided to set up a FIM4L LIBER Working group for representing FIM4L in Europe
  - The FIM4L Working Group as part of LIBER's Strategic Direction on Research Infrastructure, which in turn is one of the key pillars of our 2018-2022 Strategy.
  - See https://libereurope.eu/strategy/research-infrastructures/fim4l/
- The international group will exist and will be the „mother of all FIM4L activities"
  - We are seeking to create an IFLA WG for the international representation

FIM4L Charter: Problem statement

- There are many different situations.
  - Some countries or libraries are afraid that when using another technology, it means releasing more personal data.
  - 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.
- Managing access based on attribute release has two major pitfalls:
  - the provider of the attributes does not release the correct set or correct values
  - the provider releases more attributes than strictly necessary, violating the privacy of the user
- For publishers and libraries, it is complex and expensive to manage access to a certain resource, when libraries (in different countries) require different attributes.
- The way federated authentication is implemented by the various publishers differs a lot. This results in more confusion to our end user.

FIM4L Charter: 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

FIM4L Charter: 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
- **Promote the adoption** of state-of-the-art and privacy preserving Federated Authentication and Authorisation Infrastructure (AAI) and principles

FIM4L Charter: Planned activities

- Work together with REFEDS, RA21 [and SeamlessAccess.org] 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?

FIM4L Guidelines and recommendations: SSO implementation principles

- 1) The configuration and solution has to be in line with data protection regulations, in particular the General Data Protection Regulation (EU GDPR).
- 2) For access to services based on licensed content, next to the option of access based on IP addresses, it is recommended to use the SAML 2.0 protocol (or its follow-up technology OIDC/OAuth2 if the involved IdPs are able to handle it) to connect and control access.
- 3) eduGAIN has been established as a proper means to interfederate between identity federations, and thus enables service providers to greatly expand their user base. Thus scholarly libraries should prefer publishers who are connected to eduGAIN. Libraries should encourage publishers to make use of eduGAIN.

FIM4L Guidelines and recommendations: SSO implementation principles

- 4) The following lists the recommended options for authentication attributes, ordered by degree of privacy control, with a. being better privacy preserving than b. and so on:
  - 4.a) The publisher only requires a transient identifier - "privacy star"
    - During a session the user is identified by a transient identifier (NameID) containing a unique string with no semantic content
  - 4.b) The publisher requires a persistent but targeted identifier - "personalisation and subject tracking possible"
    - A persistent identifier (ID) contains a unique string, like the transient one, identifying the user for a specific SP, but persisting over multiple sessions: on every authentication, for the same user the same ID is used.

FIM4L Guidelines and recommendations: SSO implementation principles

- 4.c) In addition to 4.a or 4.b the SP can require extra ('non-identifiable') information
  - If more information is needed to allow for billing, access control etc. identity providers can supply one or more of the following attributes
    - eduPersonEntitlement, with the specific value urn:mace:dir:entitlement:common-lib-terms
    - eduPersonScopedAffiliation, with the kind of affiliation with the IdPs institution, e.g. student, staff, faculty, …
    - eduPersonEntitlement, with other values, representing group or role memberships in alignment with AARC Guidelines on expressing group membership and role information
    - Usage of schacLocalReportingCode attribute is recommended for statistics purposes once it is well defined

FIM4L Guidelines and recommendations: SSO implementation principles

- 5) SP software should be able to handle more attributes, but not require more attributes.
  - Some publishers state “I need an email address, as my software can’t function without it”.
  - Publishers with (older) systems that require more attributes for authentication to function should adapt their systems ASAP.
  - Libraries are recommended to stop or don’t start using services that require more personally identifiable information (PII) than a transient or persistent ID during authentication
- 6) Apart from generally working according to the GDPR, when requesting information from users, for instance in a profile page, publishers have to adhere to the most recent EU “Guidelines on Consent” to make sure that free consent is given in compliance with the GDPR.

FIM4L Guidelines and recommendations: SSO implementation principles

- 7) When providing PII to a SP, whether based on consent or not, a respective data processing agreement (DPA) may be needed.
- 8) Publishers are encouraged to declare compliance with the GÉANT Data Protection Code of Conduct.
- 9) Publishers are encouraged to declare compliance with the assertions of the REFEDS Sirtfi framework (Research and Education FEDerations group, Security Incident Response Trust Framework for Federated Identity).
- 10) Usage of RA21 WAYF/DS (Resource Access for the 21st Century, Where Are You From/Discovery Service) is recommended.

More Info at:

https://fim4l.org

http://aarc-project.eu/libraries/

https://libereurope.eu/strategy/research-infrastructu
fim4l/

Peter.gietz@daasi.de

Thank you

Any Questions?

AARC

https://aarc-project.eu

DAASI International

The work leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 730941 (AARC2).