Federated Identity Management for Libraries (FIM4L)

ELAG lighting talk - fall 2020



Jos Westerbeke Library IT specialist Erasmus University Rotterdam, Netherlands jos.westerbeke@eur.nl www.linkedin.com/in/jwesterbeke twitter.com/JosWesterbeke

Erasmus University Rotterdam

zafino





Introduction to FIM4L

- FIM4L (Federated Identity Management for Libraries) is initiated and headed by libraries. With the ending of the <u>AARC</u> project in which LIBER was <u>involved</u>, FIM4L started. (<u>http://fim4l.org</u>)
- FIM4L has involved people from related organizations including, besides a number of research related libraries, AARC, <u>GÉANT</u>, <u>REFEDS</u> and <u>NREN</u>s. FIM4L supports the <u>Stanford Statement on Patron Privacy and Database Access</u>.
- The group currently has about 50 members globally, including participants from Europe, North America and Africa.
- FIM4L cooperates with the coalition for <u>Seamless Access</u>.
- FIM4L is now a <u>LIBER working group</u> since October 2019.



The library as a trusted place

The library has to

- Provide a private place, even online
- Protect its patrons

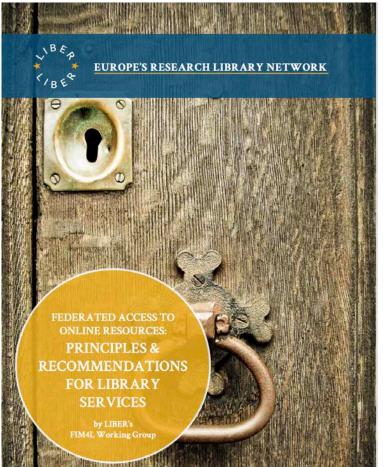


A trusted safe place with privacy



What has FIM4L done?

- Drafted a document for libraries to come to consensus about best recommendation for establishing SSO (Federated authentication)
- (almost) Published this document as a LIBER publication. (Backed by 400+ EU libraries)
- Talks with publishers
- Work together with Seamless Access



Find the PDF at: www.fim4l.org



Principle 4: Authentication

There are two recommended options for authentication attributes. If the purpose of the service is to recognize returning users, so it can present personalized features such as saved searches, profile-based recommendations for reading articles, etc, then Personalized Access is recommended for providing these options to users.

TRANSITORY ACCESS (4.A) Holds the highest level of privacy.

The publisher only requires a transient identifier: "privacy star". During a session the user is identified by a transient identifier (NameID) containing a unique alphanumeric string, for a certain Service Provider (SP). If the user logs in again, a new transient identifier will be generated.

This allows for maximum privacy. It doesn't allow the publisher to recognize a returning customer, which makes it impossible to know what resource is downloaded by the same user.

In exceptional cases, for example where misconduct is suspected, users could be identified if libraries (IdPs) have configured their systems to allow for a thorough investigation of log files, and if libraries are willing to carry out this investigation.

PERSONALIZED ACCESS (4.B)

Maintains a high level of privacy based on a pseudonym, and more user information and tracking can be added.

The publisher requires a persistent but targeted identifier: "personalisation and subject tracking possible". A persistent identifier (ID) contains a unique alphanumeric string, such as the transient one, identifying the user for a specific SP, but persisting over multiple sessions. The same ID is then used for the same user on every authentication.

This is an option for services that have a need to recognize returning customers.

It will give the user options for personalized features such as saved searches, profile-based recommendations for reading articles, etc

~

For both privacy preferences, the service provider can require extra non-identifiable information.

Two options:

4.A. Transitory Access - This access holds the highest level of privacy.

4.B. Personalized Access - Maintains a high level of privacy based on a pseudonym, and more user information and tracking can be added.



TRANSITORY ACCESS (4.A)

Holds the highest level of privacy.

The publisher only requires a transient identifier: "privacy star". During a session the user is identified by a transient identifier (NameID) containing a unique alphanumeric string, for a certain Service Provider (SP). If the user logs in again, a new transient identifier will be generated.

This allows for maximum privacy. It doesn't allow the publisher to recognize a returning customer, which makes it impossible to know what resource is downloaded by the same user.

In exceptional cases, for example where misconduct is suspected, users could be identified if libraries (IdPs) have configured their systems to allow for a thorough investigation of log files, and if libraries are willing to carry out this investigation.

PERSONALIZED ACCESS (4.B)

Maintains a high level of privacy based on a pseudonym, and more user information and tracking can be added.

The publisher requires a persistent but targeted identifier: "personalisation and subject tracking possible". A persistent identifier (ID) contains a unique alphanumeric string, such as the transient one, identifying the user for a specific SP, but persisting over multiple sessions. The same ID is then used for the same user on every authentication.

This is an option for services that have a need to recognize returning customers.

It will give the user options for personalized features such as saved searches, profile-based recommendations for reading articles, etc What attributes are exchanged during authentication?

Transient identifier (4.A)

-or-Persistent identifier (4.B)

And what comes along with it?



What comes along with it? Non-identifiable information. If required by publisher, the library can disclose e.g.:

eduPersonEntitlement (with value 'common-lib-terms') eduPersonScopedAffiliation (value e.g.: 'student')

Transient identifier (4.A) -or- Persistent identifier (4.B)

For both privacy preferences, the service provider can require extra non-identifiable information.



FIM4L.org

FIM4L LIBER working group: <u>https://libereurope.eu/strategy/research-infrastructures/fim4l/</u> FIM4L email list: <u>https://lists.daasi.de/listinfo/fim4l</u> FIM4L website: www.fim4l.org

