RA21 Problem Statement

• Access to STM content and resources is traditionally managed via IP address recognition.
• For the past 20 years, this has provided seamless access for users when on campus.
• However, with modern expectations of the consumer web, this approach is increasingly problematic:
  — Users want seamless access from any device, from any location.
  — Users increasingly start their searches on 3rd party sites (e.g. Google, PubMed) rather than publisher platforms or library portals and run into access barriers.
  — A patchwork of solutions exist to provide off-campus access: proxy servers, VPNs, Shibboleth, however the user experience is inconsistent and confusing.
  — Publishers are facing an increasing volume of illegal downloads and piracy, and fraud is difficult to track and trace because of insufficient information about the end user.
  — The lack of user data also impedes the development of more user-focused, personalized services by publishers.
  — The increase in privacy and fraud also poses a significant risk to campus information security.
Fundamental Expectations of the Community

• Researchers
  — **Seamless access** to subscribed resources, from any device, from any location, from any starting point
  — A **consistent, intuitive user experience** across resources
  — Increased **privacy of personal data**
  — Streamlined text and data mining

• Resource Providers
  — Ability to provide **individualized and differentiated** access for better reporting to governing bodies and customers
  — Ability to offer **personalized services** to accelerate insight and discovery
  — Ability to **ensure the integrity** of content on both institutional and commercial platforms

• Customers
  — Minimization of **administrative burden** of providing access to authorized user communities
  — Maximization the **use of the resources purchased**
  — **Protection of the privacy** of user communities and advocacy for their security
RA21 Guiding Principles

1. The user experience for researchers will be as seamless as possible, intuitive and consistent across varied systems, and meet evolving expectations.
2. The solution will work effectively regardless of the researcher’s starting point, physical location, and preferred device.
3. The solution will be consistent with emerging privacy regulations, will avoid requiring researchers to create yet another ID, and will achieve an optimal balance between security and usability.
4. The system will achieve end-to-end traceability, providing a robust, widely adopted mechanism for detecting fraud that occurs at institutions, vendor systems, and publishing platforms.
5. The customer will not be burdened with administrative work or expenses related to implementation and maintenance. The implementation plan should allow for gradual transition and account for different levels of technical and organizational maturity in participating institutions.
Pilot program

- Pilot program through Q3 2017
  - Broad spectrum of stakeholders
  - Address a variety of use cases
  - Includes both academic and corporate efforts
- Self organized, registered and tracked under the larger umbrella of RA21
- Feedback and results shared with the community
- Ultimate goals
  - Move away from IP authentication – lack of scale
  - Balance with the concept of privacy (General Data Protection Regulation 2018)
  - Create a set of best practice recommendations for identity discovery

Important to have multiple pilots so we can address the problem from multiple angles

The RA21 task force will not build a specific technical solution or an industry-wide authentication platform
RA21 Pilots

• Corporate Pilot
• Three Academic Pilots
  • The Academic (Shared ‘Where are you from’ (WAYF)) Pilot
  • Privacy Preserving Persistent WAYF Pilot
  • Client-based WAYF Pilot
• All seek to address the User Experience for off-campus access

*By the end of today, we are hoping to have more participants involved in each of the pilots*
Corporate Pilot

• Corporate pilot participants
  – Pharma Documentation Ring (P-D-R) member companies
    • Roche, GSK, Novartis, BASF, Abbvie
  – Scholarly publishers

• Pilot goals
  – Validation of SAML-based federated authentication in lieu of IP-based authentication for access to scholarly resources.
  – Potentially customized identity attributes to facilitate granular usage reporting.
  – Demonstration of a consistent and streamlined user experience for user authentication across multiple STM publisher sites; regardless of the user’s location and device used.
Corporate Pilot – Identity Landscape

Identity

Individual Identity
“Who are you?”

Institutional Identity
“Where are you from?”

These two things are independent
Corporate Pilot – Identity Landscape

Identity

Individual Identity

“Who are you?”

Institutional Identity

“What are you from?”

What can you do?
(What are you entitled to?)
Corporate Pilot – Identity Landscape

ALL VISITORS

Individual Identity
(we know who you are)

Institutional Identity
(we know where you’re from)

Known User
Access to Free Content
+ personalized services

Known & Entitled User

RA21
Pseudonymous & Entitled User

Anonymous
Entitled Access

Anonymous
Access to Free Content

Today’s IP Address Recognition
Corporate Pilot – Progress To Date

• Pilot officially formed in late 2016
• Survey sent to all P-D-R companies to understand identity management capabilities and readiness
• Part-time paid facilitator with support from CCC, GSK, and participating publishers
• Face-to-face meeting in March 2017
  – Whiteboarded the user experience flow
• Now developing clickable prototype
• Will test with P-D-R users in May/June

• Exploring the possibility of forming a federation among all P-D-R companies
Privacy Preserving Persistent WAYF (P3W) Pilot

Pilot goals
– To improve current Shibboleth Identity Provider discovery process
  • Incorporate additional “WAYF hints” such as email domain and IP address into federation metadata
  • Improve sign-in flow using those WAYF hints
  • Enable cross-provider persistence of WAYF choice using browser local storage

Pilot participants (confirmed so far)
– Academic Institutions
  • MIT
– Vendors/Service Providers
  • Ping Identity
  • Proquest
  • Eduserve
– Scholarly Publishers
  • Elsevier
Current Typical Sign In Flow – Step 1
Current Typical Sign In Flow – Step 2
Current Typical Sign In Flow – Step 3

ScienceDirect

Login via your institution

You may be able to login to ScienceDirect using your institution's login credentials. Learn more
We will remember your login preference the next time you access ScienceDirect from this computer.

OpenAthens login  Massachusetts Institute of Technology

Search for your institution and click the name to login.

Massachusetts Institute of Technology

Or choose your institution's region or group and click the name from the results below to login.

US Higher Education (InCommon)

Please choose one of the institutions listed below. If your institution is not listed, try using the search feature above.

US Higher Education (InCommon)

Dartmouth College
Current Typical Sign In Flow – Step 4

Welcome, please identify yourself to access MIT services.

Have an MIT certificate?

- **Use Certificate - Go**
  - Always login with this

No certificate? Use Kerberos username

- **Username:** [Enter username] @mit.edu
- **Password:** [Enter password]  
  - Login

Have Kerberos tickets?

- **Use existing tickets - Go**
  - Always login with this

Certificates are your key to secure web services at MIT, (such as Benefits, Request Tracker, SAPweb, and WebSIS) and the preferred way to access MIT servers.

If you don’t have a certificate installed on this machine, you may login using your MIT username (the 8 characters before your @mit.edu email address) and your Kerberos password.

Using your Kerberos tickets to authenticate to MIT Touchstone requires correct browser configuration. This option is useful for Athena and WIN.MIT.EDU users.
Improved First-Time Flow – Step 1

Enter your institutional email or domain to check access

Email address*

a.professor@mit.edu

Remember this account

*Your email address will not be stored

Continue
Improved First-Time Flow – Step 2
Improved First-Time Flow – Step 3
 Improved Next-Time Flow – Step 1

Chose an account to check access

+ Add Account

mit.edu >
Preserving Privacy

<table>
<thead>
<tr>
<th>Technique</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only domain part of email address needs to be transmitted from browser to publisher platform to select IDP</td>
<td>Need to define and test a standardized UI that makes this clear to users</td>
</tr>
<tr>
<td>IdP preference is stored locally in the browser, retrieved using centrally served javascript, not on a central server</td>
<td>Need to adapt Account Choose mechanism to support SAML IdPs vs OpenID Connect Authorization Servers</td>
</tr>
</tbody>
</table>

*Email address*

| a.professor@mit.edu |

+ Add Account
## Achieving the vision

<table>
<thead>
<tr>
<th>Objective</th>
<th>Work for Core Team*</th>
<th>Work for Participating Libraries</th>
<th>Work for Participating Publishers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporate additional “WAYF hints” federation metadata</td>
<td>• Agree schema for WAYF hints</td>
<td>• Add new attributes to IdP metadata</td>
<td>• Read new attributes from IdP metadata</td>
</tr>
<tr>
<td>Improve sign-in flow using those WAYF hints</td>
<td>• Design new UI flow (can borrow and adapt from URA pilot)</td>
<td>• Test new UI flow with end users and collected feedback</td>
<td>• Implement new UI flow in platform (at least as a prototype)</td>
</tr>
<tr>
<td>Enable cross-provider persistence of WAYF choice using browser local storage</td>
<td>• Adapt Account Chooser javascript</td>
<td>• Educate users</td>
<td>• Incorporate Account Chooser into UI flow (at least as a prototype)</td>
</tr>
</tbody>
</table>

*Core team* = Vendors, Publishers and Libraries interested in defining the UX and building the technology for this pilot.
Deliverables

• A set of recommendations for WAYF hints to be incorporated into federation metadata
• A recommendation on a streamlined WAYF UI flow
• A working adaptation of the Account Chooser software to meet the needs of the pilot
• A report on experience learned during the pilot and the pros/cons of taking it forward into production
There will be two independent work streams that can proceed in parallel:

- Incorporation of additional WAYF hints such as email domain and IP address ranges into participating federation’s IdP metadata, and utilization of this metadata in streamlined IdP discovery workflows by participating Service Providers
- Deployment of a shared WAYF service which allows an end-user’s preferred IdP account to be selected and stored securely in their browser, and for this choice to be securely accessed by participating SPs thus allowing the user’s WAYF choice to be persisted across sites.
Schedule

• The pilot will commence in Q2 2017 and aim to provide final recommendations by the end of 2017
The core group will need to include individuals with the following skills/experience:

• Software developers with experience of SAML, OpenID Connect and/or web application development

• UI/UX experts

• Individuals with expertise in SAML metadata schemes and standards

• Project management
Shared WAYF
Organizational Login

wayf?

Publisher
Organizational Login

Log in

Email or Username:

Password:

Remember me

Log in  Clear

Forgotten your password?
Shibboleth
OpenAthens

Shibboleth Login

If you are a member of one of the institutions displayed below you can log in using your institution username and password.

To log in using your institution’s credentials, select a geographic region.

Geographic Region:  UK Higher Education

Select one of the institutions listed below. If your institution is not listed, please contact your librarian.

UK Higher Education

- AYRSHIRE & ARRAN PRIMARY CARE NHS TRUST
- Aberystwyth University
- Academy of Sciences Library
- Accrington & Rossendale College
- Alexander Technological Educational Institute of Thessaloniki
- Anglia Ruskin University
- Aristotle University of Thessaloniki
- Arts University Bournemouth
- Ashridge
- Aston University OpenAthens
- Athens – Research and Innovation Center
Organizational Login

Log in

Email or Username:
Password:

☐ Remember me

Log in  Clear

Forgotten your password?
Shibboleth
OpenAthens

Shibboleth Login

If you are a member of one of the institutions displayed below you can log in using your institution username and password.

To log in using your institution's credentials, select a geographic region.

Geographic Region:  UK Higher Education

Select one of the institutions listed below. If your institution is not listed, please contact your librarian.

UK Higher Education
- AYRSHIRE & ARRAN PRIMARY CARE NHS TRUST
- Aberystwyth University
- Aberystwyth University
- Academy of Sciences Library
- Accrington & Rossendale College
- Alexander Technological Educational Institute of Thessaloniki
- Anglia Ruskin University
- Aristotle University of Thessaloniki
- Arts University Bournemouth
- Ashridge
- Aston University OpenAthens
- Athens - Research and Innovation Center

Not a great user experience … but it can be improved
Organizational Login... across publishers
User Perspective
User Perspective

➡ Poor Experience
Shared WAYF Pilot

why don’t publishers ask each other

instead of

all asking the user
wayf-cloud
User Experience Possibilities*

*even for first time visitors
Shared WHAT?

- publisher id
- publisher specific device id
- wayf-cloud device id
- entity ID
- IdP ID
User Control
User Control

<Browser Name> version <Browser Version number>

Operating System

- last seen Today at Awesome Publisher

Overview

Activity

- Tuesday, 24-Jan-17 13:41:45 UTC
  - Awesome Publisher
  - SAVE
  - view Details

- Tuesday, 24-Jan-17 13:41:45 UTC
  - Awesome Publisher
  - GET
  - view Details

- Tuesday, 24-Jan-17 13:41:45 UTC
  - Awesome Publisher
  - GET
  - view Details

- Tuesday, 24-Jan-17 13:41:45 UTC
  - Awesome Publisher
  - DELETE
  - view Details

- Tuesday, 24-Jan-17 13:41:45 UTC
  - Awesome Publisher
  - SAVE
  - view Details

- Tuesday, 24-Jan-17 13:41:45 UTC
  - Awesome Publisher
  - SAVE
  - view Details
Privacy

• non-personal user data
• user control
• opt-in option
Open Development

• WAYF-cloud source code is available in a public repository on Github
  – available with an Opensource Software License (Apache 2.0)
  – no code yet promoted from the development branch to the master branch

• Why?
  – Transparency / Trust / Adoption
  – No vendor lock
  – Community driven evolution
Pilot Objectives

• wayf-cloud development
  – embrace Opensource Software development model

• Easy integration to publisher platforms

• Seamless user access across publishers
  – leverage organisational authentication systems
  – Utilize shared wayf data
Pilot Participation

• Publishers with the ability to
  - integrate their publisher platforms using the wyaf-cloud API
  - adapt organizational login UI from shared WAYF data

• Organizations & Institutions with organisational authentication systems - already integrated with publisher participants (or willing to integrate)
UX examples
Sign in

Please login to access your account

Choose your account
Select an account to sign-in.

More options

Remove account NEXT

GlaxoSmithKline

Remove account NEXT

Pfizer
Choose your account

Select an account to sign-in.

- Select your organization
- Sign in as an individual

GlaxoSmithKline
Remove account
Next

Pfizer
Remove account
Next
Sign in
Please login to access your account

Select your organization
Search your organization's name.

Search

☐ Remember my selection across publishers

View list
Back
Sign in
Please login to access your account

Select your organization
Search your organization's name.

University
University of Athens
Stanford University
Universitat de Barcelona
Université de Genève

See all 132
Choose your account
Select an account to sign-in.

Select your organization
Sign in as an individual

GlaxoSmithKline
Remove account
NEXT

Pfizer
Remove account
NEXT
Sign in
Please login to access your account

Enter your username

Enter your email or phone number

Sign in with: in  tw  go  fb

Back  Next
Enter your password

Remember me

Forgot password?

Back

Sign in
Sign in
Please login to access your account

Choose your account
Select an account to sign-in.

More options

GlaxoSmithKline
Remove account NEXT

John Doe
john.doe@atypon.com
Remove account NEXT
Architecture
Architecture (Elements)

client component

WayF Widget

Server components

Publisher Platform

WayF Cloud

Institutional IdP
WAYF Widget

• Included in content sites
URA Widget - in action

GET https://www.awesomepublisher.com

200 OK index.html

set-cookie: localID=xxxx
URA Widget - in action

1. The web-browser makes a GET request to `https://berkeley.edu/shibboleth`.
   - Status code: 200 OK
   - Set-cookie: `localID=xxxx`

2. The web-server serves the `index.html` file.

3. The web-browser makes a GET request to `https://wayf-cloud.org/ura-widget.js`.

The widget is loaded.
URA Widget - in action

- **web-browser**
- **web-server**
- **WAYF cloud**

**GET** https://berkeley.edu/shibboleth

200 OK  `index.html`

set-cookie: `localID=xxxx`

**GET** https://wayf-cloud.org/ura-widget.js

**POST** https://wayf-cloud.org/ura/session

body: {
    localID: `xxxx`
}
WYAF Cloud

- **Cloud Service** operated by and agreed upon entity
- Creates device specific global IDs - stored in a cookie in the domain name of the WYAF cloud server
- Maintains device data per global device ID
- Interfaces with
  - WAYF Widget
  - Publisher Platforms
WAYF Cloud - in action

web-browser

GET https://www.awesomepublisher.com

200 OK  index.html

web-server

POST https://wayf-cloud.org/ura/session

body: {
  localID: xxxxx
}

200 OK

WAYF cloud

set-cookie: ura/gid=R1.0
WAYF Cloud - in action

Single **device** that visited multiple **web-servers**
(where the WAYF widget is installed)
Publisher Platform

• Consumes WYAF cloud services:
  – GET device data
  – PUT device data

• Directs visitor users to IdPs they have used in the past
Publisher Platforms - in action

```json
{
    host: pub1,
    timestamp: DD-MM-YYYY-HH-MM-SS,
    sso: {
        protocol: SAML,
        entityID: https://xyz.com/shibboleth
    }
}
```
Publisher Platforms - in action

platform 1
   - localID: B131

platform 2
   - localID: A-123

{  
  host: pub1,
  timestamp: DD-MM-YYYY-HH-MM-SS,
  sso: {  
    protocol: SAML,
    entityId: https://xyz.com/shibboleth
  }
}
samlbits WAYF

• samlbits discovery objectives
  – Improve Identity Provider (IdP) discovery processes
    • Use a shared discovery service that uses both browser information and shared metadata hints to narrow down IdP options for the user without tracking the user
  – Determine the best way to populate the metadata registry with hints from the Service Providers regarding what IdPs are likely to work in an authorization scenario
Process – user perspective

• Step one: discovery service checks the browser’s local store and displays the last IdP (or set of IdPs) used by the user.

• Step two: if the local browser store is empty, or if the user chooses not to use any of the IdPs offered, the user will be presented with a search interface or a list that is built based on the database of IdPs that will be known to work with that SP (the samlbits component)
SWAMID Test-SP

Logga in via SWAMID
You are connecting to...

**SWAMID Test SP**

This is an application for testing federated authentication.

This service requests that you **identify yourself**. Select your preferred identity service below...

Start typing to search for your identity service...

**Account System University of Gävle**

Account System for University of Gävle

**Accounting and payroll services for empl..**

Statens Servicecenter levererar tjänster inom löneadministration och ekonomiadministration.

**Accounting and payroll services for empl..**

Statens Servicecenter levererar tjänster inom löneadministration och ekonomiadministration.

**Blekinge Institute of Technology**
You are connecting to...

**SWAMID Test SP**

This is an application for testing federated authentication.

This service requests that you **identify yourself**. Select your preferred identity service below...

**United ID**

United ID is a secure, public and privacy-respecting identity provider
- sign up for free at https://unitedid.org!

If you always use one and the same login provider for everything it might be a good idea to remember your choice by pressing *Login and use Forever*. Press *Start Over* to redo the search and pick another login provider.

**Proceed to Login**  **Login and use Forever**

**Start Over**
Process – Publisher Perspective

• Good discovery experience relies on two things:
  – Accurately predicting user needs
    • don’t present more UI than necessary
    • understand user context
    • integrate with the web platform
    • do mobile
  – Correctly representing the publisher-customer link
    • make search count
    • don’t disappoint the user
Process – Library Perspective

• Libraries might be the IdP
  – then they don’t have to do anything other than be an IdP
• Libraries are going to be critical for the UX guidance
  – the UI needs to be better at displaying IdPs that are NOT known to work with an SP
More information

Next Steps

• Follow up with the pilot coordinators and pilot leads
  – need developers, testers, UI feedback