What is RA21?

• RA21: Resource Access for the 21st Century
• Joint initiative of the International Association of STM Publishers (STM) and the National Information Standards Organization (NISO)
• Aimed at optimizing access protocols across key stakeholder groups
  • Corporate and university subscribers, libraries, software vendors, publishers, identity federation operators, etc.
• Purpose: To facilitate seamless user experience beyond IP address recognition, supporting network security and user privacy
Late 20th Century: from print to digital

- Imitate print experience for libraries and users
- Optimized for ease of use and removal of barriers to encourage migration from print to digital
- IP address recognition became the de facto industry standard for site access
Early 21st Century: digital and remote

- Technology evolved
- Growing diversified scholarly eco-system
- Multiple entry points, e.g. mobile and remote access
- Changing user expectations and behavior
- Significant growth of usage outside of corporate/campus networks
RA21 Problem Statement

• IP-based access management increasingly problematic
• No seamless access from any device, location, or search engine
• Inconsistent and confusing patchwork of access solutions while off of the corporate/campus network (e.g. VPN servers, Proxy servers, Shibboleth)
• Increasing volume of illegal downloads and piracy
• Lack of user data to develop user-focused, personalized services
How a user experiences access to resources on campus
How a user experiences access to resources on campus
How a user experiences access to resources on campus
How a user experiences access to resources off campus
How a user experiences access to resources off campus
How a user experiences access to resources off campus
How a user experiences access to resources off campus
How a user experiences access to resources off campus
How a user experiences access to resources off campus
How a user experiences access to resources off campus
RA21 Principles: It must be open

- The solution cannot be proprietary
- The solution should be (reasonably) easy to implement
- The solution must be vendor neutral
- Should not create tremendous amounts of new work, implementation cost, or ongoing maintenance.
- Should allow for gradual implementation

- RA21 will develop Best Practice recommendations
- RA21 will not develop a specific technical solution or one industry-wide authentication platform
Three Pilots

Corporate Pilot

Academic Pilots:

- **Privacy Preserving Persistent WAYF (P3W)**
  - A shared discovery service based on storing information in the browser

- **WAYF Cloud**
  - A shared discovery service based on centralized information sharing

Pilots working together on:

- User experience and a reference UI
- Privacy and security issues
RA21 Timeline

• Q3 16 – approval STM Board, taskforce, use cases, guiding principles
• Q4 16 – first public presentations on RA21, first workshop
• Q1 17 – staff hiring, project adoption by NISO
• Q2-Q4 17 – workshops and outreach
• Q1 18 – round-up pilots
• Q2 18 – 1st draft best practices
• Q3 18 – publication of project results

Anticipated Long-Term Outputs arising from RA21:
Operational User Communities
Who’s Involved

Steering committee

- Chris Shillum, Elsevier (Co-chair)
- Meltem Dincer, Wiley (Co-chair)
- Gerry Grenier, IEEE
- Laird Barrett, Springer Nature
- Ralph Youngen, ACS
- Dan Ayala, Proquest
- Don Hamparian, OCLC
- Leif Johansson, SUNet
- Ann West, InCommon
- Andy Sanford, Ebsco
- Josh Howlett, Jisc
- Rich Wenger, MIT
- Peter Brantley, UC Davis
- Helen Malone, GSK
- Todd Carpenter, NISO
- Eefke Smit, STM
- Ann Gabriel, Elsevier (RA21 Outreach Committee)

Outreach & Communications committee

- Michelle Brewer, Wolters Kluwer
- Sam Bruinsma, Brill
- Angela Cochran, ASCE
- Ann Gabriel, Elsevier (Chair)
- Don Hamparian, OCLC
- Robert Kelshian, American University
- Tim Lloyd, LibLynx
- Judy Luther, Informed Strategies
- Matt McKay, STM
- Jonathan Morgan, ACS
- Jean Shipman, Elsevier
- Lauren Tulloch, CCC
- Keith Webster, Carnegie Mellon University

Staff

- Julia Wallace, Project Director
- Heather Flanagan, Coordinator
- Academic Pilots
- Jenny Walker, Coordinator Corporate Pilot

Combined with our Multi-stakeholder Advisory Group & Pilot Participants:

- Over 65 organisations from key stakeholder communities are represented within RA21
Resource Access for the 21st Century

Position Papers

Heather Flanagan, RA21 Academic Pilot Coordinator
RA21 Position Papers

• What are position papers?
  – short, targeted documents describing agreed upon best practices that can be implemented today

• Who is the target audience?
  – IT managers and leaders
Recommendations to Identity Providers and Federation Operators

• need for more complete metadata records from Identity Providers (IdPs) in SAML-based federations
  – this will allow Service Providers (SPs) to offer end users a better user experience

• review session management configuration to (potentially) support logins once per business day and offer a seamless experience for all SPs
Recommendations to Content Providers

- normalize the language used on user-facing authentication pages
- basic presentation of login options to the user
- making use of the MDUI hints (esp. logos) offered by the IdP
(Possible) Future Papers

- Additional papers may be developed
  - looking for early wins and consensus on specific items
- These papers, along with the other outputs of the project, will wrap up in a final package that will be fed through the NISO standards process
RA21 Corporate Pilot:

A Customer Perspective

Helen Malone
Director, Information Hub

6th December 2017
Objectives for the RA21 Corporate Pilot

- Test **Single Sign On** access with pilot publishers
- Improve the **user experience** at pilot publisher sites
- Explore ways to capture **granular usage statistics**
Corporate Pilot:
Pharma Companies & Publishers

AbbVie
BASF
Roche
GSK
Elsevier
Novartis
ACS
Springer Nature
Wiley
Example of a Potential New Access Model: Inside the Corporate Network

Step 1:

Enter your work email address:
Helen.J.Malone@gsk.com

Or find your company name:
GSK

No personal information stored

Example: 90-day cookie for future 1-click access

SSO (eg Ping Federate)
Example of a Potential New Access Model: Outside the Corporate Network

Step 1:

Enter your work email address:
Helen.J.Malone@gsk.com

Or find your company name:
GSK

No personal information stored

Step 2:

GSK Login Page

Username

Password

Login »
Easy Set Up between Companies and Publishers?

One Corporate Federation

- Abbvie
- BASF
- Roche
- GSK
- Novartis
- ACS
- Springer Nature
- Wiley
- Elsevier
Usage Statistics:
Knowing what and when our users download

1. User **Login ID** or Email Address

2. Publisher **Bibliographic** Info
3. Publisher **Date / Time** Stamp

Monthly **Automated Secure File Transfer**

Meaningful **Usage Reports**
Usage Statistics:
An additional attribute

User Level Statistics
Working Together in Partnership

Customers

Technology Experts

Publishers

Phased Approach to Implementation
User Experience
“You have to start with the customer experience and work your way back to technology.”
— Steve Jobs
RA21’s User Experience Challenge

• Today:
  – Awesome user experience on campus
  – Awful user experience off campus

• Tomorrow:
  – Consistent user experience anywhere on any device

• Challenge:
  – On campus user experience will become slightly less seamless
RA21’s User Experience Goals

• Improve the Where Are You From (WAYF) User Experience
  – Encourage consistency across all publisher websites

• Reduce the number of times a user encounters the WAYF
  – Share users’ preferred identity providers across publishers subject to both user privacy and publisher confidentiality concerns
RA21 UX Development Across the RA21 Pilots

- Corporate Pilot
  - Developed prototype WAYF UX based upon institution name or institution email address lookup
  - Tested prototype with ~50 corporate scientists
RA21 UX Development Across the RA21 Pilots

• UX work that began under Corporate Pilot is continuing as a single track across both the P3W and WAYF Cloud pilots
  – Heavy emphasis on how to accomplish cross-publisher sharing of prior identity provider choices

RA21 UX Demo

Warning: Work in Progress!
## Pilot Approaches to Cross-Publisher Sharing

### P3W vs. WAYF Cloud

<table>
<thead>
<tr>
<th>Sharing Approach</th>
<th>User Experience Impact</th>
<th>Security/Privacy Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P3W</strong></td>
<td>Tradeoff between UX options at publisher site and browser compatibility. May require iFrames or other approaches that stretch browser compatibility.</td>
<td>Less impact. Only IdP choices stored. All data stored in local browser.</td>
</tr>
<tr>
<td>Prior identity provider choices are stored in local browser storage.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WAYF Cloud</strong></td>
<td>Less impact. Prior identity provider choices are retrieved via backend call to centralize service.</td>
<td>Potential concern. Requires trusted third party to protect data. May not be compatible with privacy regulations.</td>
</tr>
<tr>
<td>Prior identity provider choices are stored in a centralized service.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Evaluation Criteria:

- UI/UX Flexibility
- User Privacy
- Publisher Privacy
- Browser Compatibility
- Implementation Complexity
- Transparency
- Resilience
- Etc.
Questions?
The P3W Pilot
Privacy Preserving Persistent WAYF
P3W Pilot Goals

To improve current SAML (Shibboleth) Identity Provider (IdP) discovery process

• Incorporate additional “WAYF hints” such as email domain and IP address into federation metadata
• Use both browser information and shared metadata hints to narrow down IdP options for the user without tracking the user
• Improve sign-in flow by using smart search and asking for minimal information up front
• Implement consistent, familiar UX across participants
• Enable cross-Service Provider persistence of WAYF choice using browser local storage

Pilot participants

SUNet (lead) 
Geant (project management) 
American Chemical Society 
CANARIE 
EBSCO 
Elsevier 

Johns Hopkins 
LibLynx 
myunidays 
OpenAthens 
ProQuest 
University of Nottingham
### P3W Components

**IdP Search**
- “Smart” search service making use of IdP metadata and browser hints and knowledge of which service providers work with which IdPs

**IdP Choice Persistence**
- Remembers previously used IdPs in browser local storage
- Gives user control over which service providers they share this information with

Services may be used separately in deep integration model
P3W Integration models

**Central discovery service**
– Service provider redirects user to central site to handle IdP selection
– Very simple integration model for SPs

**Deep integration**
– Service provider integrates search and/or IdP choice persistence into their own UI using shared Javascript
– Allows for more seamless UX
UI Flow – 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 of IdPs
## Preserving Privacy

<table>
<thead>
<tr>
<th>Technique</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full email address (if entered) does not leave the browser: Only domain</td>
<td>Need to define and test a standardized UI that makes this clear to users</td>
</tr>
<tr>
<td>needs to be shared with discovery service and only IdP choice needs to</td>
<td></td>
</tr>
<tr>
<td>be shared with service provider</td>
<td></td>
</tr>
<tr>
<td>IdP preference is stored locally in the browser, retrieved using centrally</td>
<td>Need to ensure that this model is compatible with latest browser initiatives to tighten up cross-domain</td>
</tr>
<tr>
<td>served javascript, not on a central server</td>
<td>communication</td>
</tr>
</tbody>
</table>
Challenges

• **Architecture for deep integration option**
  – There are several different models for integration, e.g
    • iFrames to render part of UI
    • iFrames for inter-domain messaging
  – Need to find right balance between UI consistency and flexibility and browser security model

• **Local accounts**
  – Most SPs need to support a variety of integration models
    • Local usernames/passwords
    • Non-federated IdPs
    • Need to ensure that these options can be smoothly accommodated in the UI flow

• **IdP Metadata**
  – Need IdPs to ensure necessary information (email domains, logos, etc) is accurately and consistently included in federation metadata
  – Need feedback process when metadata is incorrect, incomplete or inconsistent
Progress and Next steps

• SUNet’s pyff.io pilot platform has been extended to support:
  – Cross-domain shared settings based on browser local store and hidden iFrame messaging
  – Low-level discovery client API
  – jQuery widget to provide customizable discovery API

• Several other pilot participants are now working to integrate with this service in a sandbox environment
The WAYF Cloud Pilot

December 6, 2017

Meltem Dincer
VP, Platform Capabilities, Wiley
RA21 Co-chair
WAYF Cloud Pilot Goals

To provide a seamless user experience as close as possible to IP Authentication

- Eliminate steps users have to repeat at every publisher
- Leverage existing organizational systems/protocols for user authentication
- Create an infrastructure for sharing WAYF data amongst publishers
  - Embrace OpenSource Software development
  - Establish easy integration points with service provider platforms
- Look to form a potential industry standard for WAYF data exchange
  - Data Format
  - Modern Interface Specification

Pilot participants

- Atypon
- OpenAthens
- RINGGOLD
- SAGE
- Silver Chair
- UC Davis
- Wolters Kluwer
- WILEY
Desired User Access Experience

Private Experience

Target Institutional Experience

Log In

Email

Password

Remember Me

Forgot Password?

Log In

Log In

Email

Password

Continues with University of Wassamatta

Remember Me

Forgot Password?

Log In
The WAYF Cloud at a glance

What is it?
• Data Format Definition
• Interface Specification
• a server component

What does do?
• allows platforms to communicate with each other by
  • storing data shared by the platforms
  • serving the data back to the platforms

Architecture:
• Shared Infrastructure
• Decentralized Trust Model
How does it work?

First visit

Second visit

Manual Selection

WAYF Cloud

publisher platform

publisher platform

publisher platform

publisher platform

publisher platform

Past selection injected in the login page

discover

share
WAYF Cloud Components

• **WAYF Cloud Widget:**
  – Transfers the unique identifier of the device in the domain of the service provider
  – Service provider simply incorporates the WAYF Widget URL into relevant HTML pages

• **WAYF Cloud API**
  – Interface used by the service providers to Create, Discover, Share and keep up to date a user's WAYF history

• **WAYF Cloud**
  – Centralized service that assigns a global ID to the device and maintains the relationships with the local IDs
  – The global ID is stored at the device in the form of a cookie and its carried in all requests made by this device (i.e web browser) to the WAYF Cloud server.
  – Uses the information provided by the WAYF Widget to build relationships between a user's global ID and the different local IDs used by the different service providers for this device
  – **The relationship enables the seamless user experience**
WAYF Cloud Challenges

• Security/Privacy
• Maintaining an open sourced common code base
• Operating a shared service
RA21 – Security & Privacy – For all pilots

Privacy Track (non-technical)
• Analyze data collected for intended use and storage to ensure compliance with data privacy regulations (GDPR, GLBA, etc.)
• Perform privacy impact assessments
• Validate privacy controls are commensurate with data values per best practices

Final Results: Recommendations for privacy controls

Security Track (technical)
• Assess pilot against information security & web development best practices:
  • Adherence to W3C web development standards
  • Secure coding practices
  • Vulnerability management
  • Penetration testing
  • Authentication standards

Final Results: Recommendations for following W3C standards with proper security controls
## Operating an open sourced shared service

<table>
<thead>
<tr>
<th>Development</th>
<th>Run Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Contributor License Agreement and documented contributing process</td>
<td>• How do we know what’s running is what’s on GitHub?</td>
</tr>
<tr>
<td>• Copyright ownership</td>
<td>• Who runs the service?</td>
</tr>
<tr>
<td>• Organization to receives the contributions</td>
<td>• Who takes the responsibility for failure?</td>
</tr>
<tr>
<td>• Governance</td>
<td>• Who owns the data?</td>
</tr>
<tr>
<td>• Development process</td>
<td>• How and who manages SLAs (performance, security, privacy, etc.)?</td>
</tr>
<tr>
<td>• Release process</td>
<td></td>
</tr>
<tr>
<td>• Testing process</td>
<td></td>
</tr>
</tbody>
</table>
Progress – Development

Open Sourced
Licensed under Apache v2.0

API documentation web-site for vendors interested in integrating with the WAYF Cloud
Progress - Sandbox

- Sandbox system & working demo
  https://wayf-cloud-sandbox.literatumonline.com
Working Groups and Next Steps

• Privacy & Security – Face to face workshop on Dec 8th
• Interface Specification / Realization
• Testing & Usability Evaluation
• Operating the shared service
Questions?