An Introduction to RA21

Todd Carpenter (NISO)
Robert Kelshian (American University Library)
Don Hamparian (OCLC)
Ann Gabriel (Elsevier)
RA21 Steering Committee

November 10, 2017
Moving Content From Print to Digital
IP -Address Authentication

Your IPv6 Address Is:
2606:f180:0:45b:45b:be23:53f3:a2db

Your IP Details:
ISP: EGIHosting
City: San Jose
Region: California
Country: United States

Leaflet | © OpenStreetMap Terms

See your IP address? See your location?
That’s why people use a VPN.

Learn More
It worked well in this environment
Until, people began connecting from everywhere
IP Address Authentication

Your IPv6 Address Is:
2606:f180:0:45b:45b:be23:53f3:a2db

FAIL!!!
Librarians are smart!

Proxy Server
These solutions worked well
These solutions worked well

Until they didn’t
Who had their proxy server shut down when a content provider notices unusual activity from that server?
Behind the scenes: Does the user have access rights? Yes or No?

Do you have a login? Yes or No?

Where are you from?

???????
And patrons are just getting annoyed
All this and no data...

Who is using the content libraries acquire? Are those user’s getting what they need? Are they achieving their goals? Successful in their outcomes? Are they really secure? Are they really private?
RA21 Principles: Improve User Experience

• From any location on any device
• Beginning from any entrance point
• Ending with the desired content
• With a consistent user interface
• With greater privacy, security and personalization
RA21 Principles: It must be open

• The solution can not 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
Pilot program

• Pilot program through Q1 2018-ish
  — 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
RA21 Pilots

• Corporate Pilot (Universal Resource Access “URA”)
• Two Academic Pilots
  — Privacy Preserving Persistent WAYF Pilot
  — WAYF Cloud Pilot

• All seek to address the User Experience for off-campus access
Progress of our corporate pilot...

- New **user experience flow** developed by publishers
- Agreed requirements for **granular usage statistics**
- Exploring federation for **easy flow between publishers**
Privacy Preserving Persistent (P3) WAYF 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 via a shared discovery service
  - Populate shared discovery service hints from the Service Providers regarding what Identity Providers are likely to work in an authorization scenario
  - Enable cross-provider persistence of WAYF choice using browser local storage

**Pilot participants (confirmed so far)**

### Project Management
- GÉANT

### Educational Access Management Federations
- Sunet & SWAMiD (Swedish Federation)
- The samlbits.org project
- eduGAIN
- EduServ

### Publishers
- Elsevier
- American Chemical Society

### Subscribing institutions
- MIT
- University of California, Davis
- University of Arizona (tbc)
- University of Florida (tbc)
- University of Denver (tbc)

### Service Providers
- ProQuest
- Ping
- LibLynx
- Ebsco
Preserving Privacy

<table>
<thead>
<tr>
<th>Built upon ”SAML-BITS” technology in production</th>
<th>Email address*</th>
<th>Only domain part of email address needs to be transmitted from browser to publisher platform to select IDP</th>
<th>Need to define and test a standardized UI that makes this clear to users</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><a href="mailto:a.professor@mit.edu">a.professor@mit.edu</a></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IdP preference is stored locally in the browser, retrieved using centrally served javascript, not on a central server

Need to adapt Account Choose mechanism to support SAML IdPs vs OpenID Connect Authorization Servers

+ Add Account
WAYF Cloud Pilot

• Goal
  — Seamless Access as close to IP Authentication as possible
  — Eliminate steps which users have to repeat at every publisher
  — Support for remote access

• Methodology
  — Leverage existing organizational systems/protocols for user authentication
  — Look to form a potential industry standard for WAYF data exchange
    - Data Format
    - Modern Interface Specification
  — Create an infrastructure for sharing WAYF data amongst publishers. The WAYF Cloud software
    - embrace OpenSource Software development
    - easy integration points with service provider platforms
  — User Interface - Reference design
What is the WAYF Cloud?

**What is it?**
- Data Format Definition
- Data Access Interface Specification
- Software Component (Free/Opensource)

**What does it do?**
- Allows publishers to exchange information with each other
Want to get involved?

• Visit: https://www.RA21.org

• Mailing lists:
  – P3W community list: https://lists.refeds.org/sympa/subscribe/p3w-community
  – WAYF Cloud community list: TBD

• Everyone: Register your interest in participation by emailing:
  Julie Wallace: Julia@RA21.org and
  Heather Flanigan: Heather@RA21.org
Contact:

Todd Carpenter
Executive Director
National Information Standards Organization (NISO)
tcarpenter@niso.org

@TAC_NISO
An Academic Library Perspective

• What’s the problem that needs solving?
  – I know through turnaway stats what users want that I don’t have.
  – I don’t know how often users can’t get to what I do have.
  – I don’t know every user’s background, experience, and how they navigate.
  – If I had access to alternatives that provided more robust access and didn’t use them, I would be in trouble.
So many paths to the same end...

• Changes in how and who is accessing what:
  – Large percentage of access is happening from off-campus.
  – Increase in number of online programs with populations who may never receive library instruction.
  – Fewer opportunities to find out what I don’t know.
    • Increasing number of students with whom we never have contact.
• Tendency toward judging how user’s navigate
  – What do you mean they don’t go through the library website?
  – Are we getting the most for our money if we’re not trying to ensure access at the end of the road?
Security and Privacy

- Using Shibboleth as a front layer to EZproxy has worked well for us.
  - Still counting on user’s to navigate through the website.
- We care a lot about privacy. Maybe more than our users do.
- Both academic pilots for RA21 have emphasized user privacy as a priority for the project. Additional focus on ease of implementation will be important as the project progresses as well.
- Has to be easy to implement to make it a viable option.
Contact:

Rob Kelshian
Director of Access Services
American University Library
calvin@american.edu

Thanks!
Charleston: EZproxy and RA21

Don Hamparian
Senior Product Manager, EZproxy and Identity Management,
OCLC
A Quick Show of Hands

Who has experience with:

Using EZproxy to access licensed content?

Administrating EZproxy?

Before this panel, who knew what:

The RA21 Initiative is?

At your institution:

Does your institution have a SAML-based Authentication system (Shibboleth or CAS)?

Are your library systems integrated?

EZproxy and RA21 – OCLC’s Commitment

- EZproxy support for standard specifications from RA21
- EZproxy enhancements for security features for IP authentication
- Compatibility with RA21 SSO schemes and traditional IP authentication
- Improving our database configuration process
Securing EZproxy Today

- Most common problem (by far) is compromised credentials at the institution
- EZproxy is not “hacked”
- Four-part strategy to secure access
  - Protect & Prepare (Proper configuration)
  - Detect and close compromised credentials
  - Educate (Provider, Institution Admin, Users)
  - Collaborate (OCLC, Institution, Provider)
Institution Best Practices – Protect & Prepare

• Password Policies
  – Biggest opportunity – get this done and EZproxy provides a very secure content access path
  – Require hard-to-guess passwords
  – Consider Multifactor Authentication (MFA)

• Commit resources for EZproxy server management

• Exercise detection process before publisher calls you
Institution Best Practices – Protect & Prepare

• Keep your evidence (configure log / audit files with data you need) and back them up
• Use SSL for authentication and content access where possible
• Keep server OS upgraded
• Install the current version of EZproxy
• Keep your system time correct
Educate & Collaborate
Involving Publishers, OCLC, Libraries

• Improve education (staff & patron)
• Improve technology
• Improve support of our mutual customers
• Use OCLC Support Site & Community Center
• Your ideas?

https://www.flickr.com/photos/quinnanya/111201180
References

- EZproxy Support Site: https://www.oclc.org/support/services/ezproxy.en.html
- Managing your EZproxy: https://www.oclc.org/support/services/ezproxy/documentation/manage.en.html
- Securing your server: https://www.oclc.org/support/services/ezproxy/documentation/example/securing.en.html
- EZproxy - Publisher Support Page: https://www.oclc.org/support/services/ezproxy/contentproviders.en.html
- RA21 Initiative: https://ra21.org/

http://www.clker.com/clipart-13969.html
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
    • don’t disappoint the user
Process – Publisher Perspective

• With User Consent Publishers will have the opportunity to provide a better, more customized experience.

• Publishers will have the ability to provide granular and differentiated access for better reporting to governing bodies and customers.

• Publishers will be able to work with purchasing departments to more carefully manage licensed access and together could quickly target any instances of fraudulent or illegal activity.

• RA21 aims to increase the ability of publishers to ensure the integrity of content on both institutional and commercial platforms.
What about Privacy?

User privacy is one of the guiding principles of RA21

- The General Data Protection Regulation and the e-privacy regulation proposal arising from the EU provide act as powerful input to the expectations of the project.
- SAML federated authentication technology has in-built mechanisms for preserving privacy.
- The institution is fully in control of what personally identifiable information is disclosed to a resource provider
- Academic SAML Identity Providers provide unique, persistent, but OPAQUE identifiers
- Publishers may then ask for additional personal information from users via their normal registration processes, disclosing their privacy policies.
Outreach Activities

CNI - December 2016, April 2017
STM - December 2016, July 2017, December 2017
SSP - May 2017
JISC - July 2017
AGLIN Forum - August 2017
SURF - September 2017 Utrecht
Internet2 - October 2017 San Francisco
Charleston Conference - November 10, 2017
UKSG - November 16, 2017
CCC - hosted webinar November 16, 2017
ALA Midwinter - January 20-24, 2018 Denver
PSP - February 7-9, 2018 DC
MLA - May 18-23, 2018, Atlanta
SSP - May 30 - June 1, 2018, Chicago

RA21 in the News

UKSG Insight – Opinion Pieces: “Easy access to the version of record (VoR) could help combat piracy: views from a publishing technologist” Author: Tasha Mellins-Cohen. 10 July 2017.


Index Data: “RA21 Project aims to ease remote access to licensed content” Author: Peter Murray. 19 December 2016.
RA21 Position Paper

Still in DRAFT

• Audience: Identity Providers and Federation Operators

• Topic: Metadata Extensions and Session Management recommendations

• Goal: areas that identity providers and federation operators can work on today to immediately start improving the user’s identity discovery experience
Pilot Participation

Agreed goals:

- Working demoable system, with publisher platforms integrated with the WAYF Cloud

- A public Github repository with the source of the WAYF Cloud with an Open Source License

- A report with the results from user testing

- A set of recommendations for the governance and other operational aspects of the WAYF Cloud
Who’s Involved

• Constituted as a joint NISO/STM initiative
• Initial funding provided by participating publishers

• Two dedicated staff brought on board to drive pilots:
  - Julia Wallace
    • Program Director
    • Julia@RA21.org
  - Heather Flanagan
    • Project Coordinator
    • Heather@RA21.org

Steering community with stakeholders from library, vendor and identity management communities:

  • Chris Shillum, Elsevier (Co-chair)
  • Meltem Dincer, Wiley (Co-chair)
  • Gerry Grenier, IEEE
  • Laird Barrett, Springer Nature
  • Ralph Youngen, American Chemical Society
  • 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
  • Library Representative (name TBA soon)
  • Todd Carpenter, NISO representative
  • Eefke Smit, STM representative
  • Ann Gabriel, Elsevier (Chair, RA21 Outreach and Communications Committee)
QUESTIONS???

Thank you!