Purdue University

Purdue e-Pubs

Charleston Library Conference

Introducing SeamlessAccess.org: Delivering a Simpler, Privacy-Preserving Access Experience

John W. Felts
Coastal Carolina University, jfelts@coastal.edu

Tim Lloyd
LibLynx, tim@liblynx.com

Emily Singley
Boston College, emily.singley@bc.edu

John W. Felts, Tim Lloyd, and Emily Singley, "Introducing SeamlessAccess.org: Delivering a Simpler, Privacy-Preserving Access Experience" (2019). *Proceedings of the Charleston Library Conference.*
http://dx.doi.org/10.5703/1288284317205

Introducing SeamlessAccess.org: Delivering a Simpler, Privacy-Preserving Access Experience

John Felts, Head of Library Technology and Systems, Coastal Carolina University, jfelts@coastal.edu

Tim Lloyd, CEO, LibLynx, tim@liblynx.com

Emily Singley, Head of Systems & Applications, Boston College Libraries, emily.singley@bc.edu

Abstract

Managing access to subscribed services in an era of abundance is a major challenge for libraries. Users have come to expect a seamless, personalized experience on their mobile devices, but traditional approaches to access management force librarians to choose between the anonymous ease of onsite IP authentication or the access friction experienced by users authenticating across multiple resources with single sign-on. Building on the work of the RA21 initiative, a recent NISO Recommended Practice on Improved Access to Institutionally Provided Information Resources charts a way forward. It will enable libraries to provide seamless, privacy-preserving, and one-click access to their subscribed content from any device, any location, and from any starting point in the research process. The implementation of these recommendations will be led by SeamlessAccess.org, starting with a beta phase implementation in the fall of 2019. But how are user and data privacy protected? How is access simplified? How will numerous library use case scenarios be accommodated, and will current accessibility standards be implemented and supported? This paper discusses how these concerns are being addressed by a consortium of industry partners including librarians, access providers, publishers, and standards organizations. It also discusses how the coalition will manage this service for publishers and libraries while continuing to improve this user experience, provide governance on data policy and privacy issues, and maintain core Web services specific to this initiative.

History

The Resource Access for the 21st Century project (RA21) was created in 2016, initially to explore the challenge of remote access. It involved stakeholders from the publishing, library, software, and identity communities, and took input from 60 organizations over a three-year period. This initiative identified that federated identity management (FIM) held the most promise for providing a robust, scalable solution for remote access to scholarly content and investigated barriers to adoption, developed best practices, and piloted technical approaches to simplifying access. Its conclusions were published as a draft NISO Recommended Practice in April, receiving over 200 comments that helped identify further areas for investigation and confirm the value of testing a beta service. A final NISO Recommended Practice was published in June and is available on the NISO website: https://groups.niso.org/apps/group_public/download.php/21892/NISO_RP-27-2019_RA21_Identity_Discovery_and_Persistence.pdf

Why Do We Need SeamlessAccess?

IP recognition has been around since the 1970s, and library use of IP recognition was developed when offsite access to electronic resources was in its infancy and has changed very little since then. Although it wasn't originally designed for remote access, it proved very beneficial since prior to its implementation, libraries found it difficult and unwieldy to implement and support a proxy server. But after 20 years of IP authentication, technologies have matured to where we can now do a better job.

SeamlessAccess seeks to improve remote access scenarios and create a better user experience. There has been a considerable increase in off-campus access to library resources over the years and from a multitude of devices. Using IP authentication forces researchers to begin their research from or at some point navigate through the library's website to click on the proxy-prefixed URL necessary for remote access. This simply is not how researchers conduct their research. Current obstacles to access include forcing the user to click through numerous pages to access content behind a paywall, plus typically users have credentials scattered over a multitude of platforms that become increasingly difficult to manage. Also, if an institution provides numerous solutions for content access (e.g., VPN, EZproxy, Shibboleth, etc.) then users can become overwhelmed with complicated instructions regarding which protocol to select and how to implement these on their local devices. By providing such complicated

https://doi.org/10.5703/1288284317205

procedures for navigating beyond a paywall, libraries could inadvertently be pushing fully entitled end users to turn to alternative resources such as Sci-Hub or ResearchGate to obtain easily accessible content.

SeamlessAccess also seeks to enhance user privacy mechanisms. Typically, IP authentication is thought to be more privacy-preserving than federated identity management, yet this is not the case. Since every EZproxy .log file conflates a username with a user's IP address and therefore by physical location, this information alone is considered personally identifiable information and is not in compliance with the EU's General Data Protection Regulation. Also, many libraries still use some type of URL rewriting method, which has significant security vulnerabilities such as IP spoofing and man-in-the-middle attacks, clickjacking, and session hijacking. Another security challenge attributed to IP address authentication often stems from the theft of the personal credentials to proxies that serve as gateways to IP addresses. There are numerous blogs, sites, and boards on the open Web that list institutions with compromised user credentials that allow unfettered access to large amounts of content.

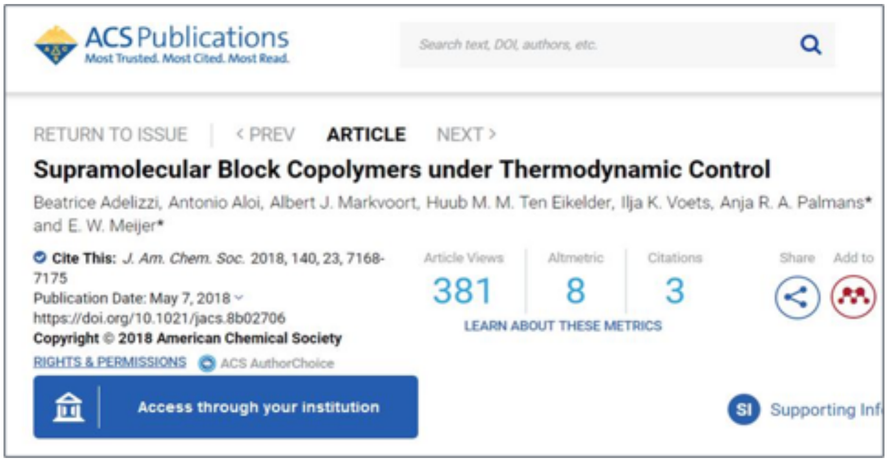
Also, when using IP authentication a critical library service issue occurs when a vendor identifies one compromised user account accessing content on their platform. If this happens, the service provider is forced to block access for the IP number on which the proxy server resides, which in turn blocks access for the entire institution. If federated identity management were implemented, then access could be blocked for that one compromised user account rather than for the entire organization.

What Is SeamlessAccess?

SeamlessAccess is a community-driven effort to enable seamless access to information resources, scholarly collaboration tools, and shared research infrastructures. To date, this initiative has five founding organizations: the National Information Standards Organization (NISO), GÉANT, Internet2, ORCID, and the International Association of STM Publishers, and features a full-time implementation team that includes an experienced library technologist dedicated to library outreach. There's also a governance committee with representatives from across the stakeholder groups, an outreach committee that includes six institutional participants, and a new cross-industry working group that will explore ways to improve the release of user attributes. SeamlessAccess is now in the process of testing ideas that were developed and piloted by RA21 in the light of community feedback and developing best practices around the use of federated identity management. This beta phase is expected to run until June 2020, and from July 2020 SeamlessAccess expects to have a fully operational service.

User Experience

Because the overall user authentication experience is currently inconsistent, confusing, and replete with jargon, SeamlessAccess seeks to implement a standard for digital authentication based on a single sign-on through the user's home institution. Regardless of where the end user begins their research, they will encounter consistent imagery, language, and log-in placement, along with a standardized Identity Provider Discovery flow (Figure 1).



ACS Publications
Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

RETURN TO ISSUE | < PREV ARTICLE NEXT >

Supramolecular Block Copolymers under Thermodynamic Control

Beatrice Adelizzi, Antonio Aloi, Albert J. Markvoort, Huub M. M. Ten Eikelder, Ilja K. Voets, Anja R. A. Palmans* and E. W. Meijer*

Cite This: *J. Am. Chem. Soc.* 2018, 140, 23, 7168-7175
Publication Date: May 7, 2018
https://doi.org/10.1021/jacs.8b02706
Copyright © 2018 American Chemical Society
RIGHTS & PERMISSIONS ACS AuthorChoice

Article Views	Altmetric	Citations
381	8	3

LEARN ABOUT THESE METRICS

Figure 1. **Consistent imagery, language, and placement.**

Once authenticated using their preferred sign-in credentials, the end user will not be required to sign in again across all SeamlessAccess-enabled sites. Additionally, an important milestone has been reached: Springer Nature is the first publisher to implement the SeamlessAccess service, which will provide vital user experience feedback to the SeamlessAccess implementation team as it continues to enhance and improve these core services and overall usability.

Data Privacy and Attributes

In federated identity management, attributes represent extra data about an authenticated user. Attribute release is the process by which that data is shared as part of the sign-in process by an identity provider (IdP), such as an institution, with a service provider (SP), such as a publisher. The format an attribute takes depends on the underlying technology. For example, SAML is the technology that underpins Shibboleth and OpenAthens, but there are other technologies that support federated identity management, such as OpenID Connect, which is used by consumer-focused services such as Facebook and Google.

Table 1 gives some examples of the types of attributes that can be passed as a result of a successful user authentication.

- An anonymous token is one that is uniquely generated for every log-in and for each service provider, regardless of whether it is a new or returning user. This token can't be used to support personalization because it changes every time the user signs in and therefore retains user anonymity.
- A pseudonymous identifier is unique to each person and for each service provider, so it masks their true identity, but it does enable that user to be identified by the same service provider the next time they visit. However, it's important to note that this identifier can't be used to build a pattern of usage across service providers. It can also be used to personalize a user's experience.
- There are a variety of organizational data fields that can be provided where necessary, such as a user's home organization, their entitlements (or rights), role, department, or location.
- There are also personal data fields, such as name and e-mail address.

Table 1.Attribute types.

Anonymous token	• Unique for every visit • Real identity unknown (anonymous) • No personalization
Pseudonymous ID	• Unique ID for every person • Real identity unknown (pseudonymous) • Personalization possible
Organizational	• Home organization, entitlements, role, department, location, etc.
Personal	• Name, e-mail address

Attributes are important because they give both sides of the authentication transaction greater control. This control can be valuable in a variety of different ways. For example:

- **Access control**: A library can choose to make a resource available only to users who are full-time staff and students, preventing, say, alumni or contractors from access.
- **Cost control**: A library can limit resource access to users with a certain role or from a certain department.
- **Risk control**: Pseudonymous IDs allow users to benefit from personalization without exposing them to the risks and inconvenience of separately registering yet another username and password. The service provider can recognize a returning pseudonymous ID and personalize that user's experience accordingly without receiving any personally identifiable data, without needing to store their e-mail address, and without asking for a password.

Attribute release is optional. An IdP can simply assert that a user is a member of their organization and do nothing more, and this only occurs after a user is authenticated. A service provider can't download additional attributes; they only receive what the IdP chooses to send. Attribute release is configured by the IdP for each "category" of SP.

Access to library resources is only one of a number of valuable use cases for federated identity management. For example, research collaborations need more identity data so that research and opinions can be associated with individuals. In these cases, a configuration would typically send a name and e-mail address. Federated identity management can also be used by an organization to authenticate access to services provided by vendors, such as payroll or communications platforms. In these cases, it may be appropriate to share a much broader range of user data. With library resources, SeamlessAccess's recommendation is for a much more limited set of attributes to be shared with a vendor or publisher. Because the IdP is in control, the institution needs to agree to any special needs for attributes in advance so that the SP profile can be configured appropriately.

Publishing Use Cases

Table 2 gives examples of how the use of attributes translate into the real world, along with some example publishing uses cases.

- Scenario 1. Users are accessing full-text articles on a platform where there is no option for personalization. They simply need to confirm that they are members of your organization. In this case, the vendor simply needs an anonymous token.
- Scenario 2. Users can get content recommendations in the user interface based on their prior search history so the vendor needs to recognize them when they return to the platform. In this case, a pseudonymous identifier will enable this.
- Scenario 3. Select users are allowed by the library to have access to prepaid funds in order to buy e-books for their department. In this case, an attribute for a user's role could be used, in addition to the pseudonymous ID.
- Scenario 4. Clinicians are participating in online training to earn continuing education credits and they need to receive a certificate by e-mail. In this case, user consent should be sought to obtain an e-mail address in addition to the pseudonymous ID.

User Privacy and Security

SeamlessAccess shares with librarians the belief that preserving patron privacy and intellectual freedom is a fundamental principle of library service. In support of this, RA21 (as the operational predecessor to SeamlessAccess) convened the RA21 Security & Privacy Working Group to conduct a technical analysis of possible security and privacy risks associated with this service. No significant risks were found and any residual risks from both a security and privacy perspective were considered low. The nature of the data involved was found to not be directly attributable to an individual and appropriate safeguards were deemed to be in place to mitigate any confidentiality concerns. The complete technical report is located on the RA21 website: https://ra21.org/wp-content/uploads/2018/07/RA21-Security-Privacy-Final-Report.pdf

Also, on January 31, 2019, RA21 adopted the GÉANT Data Protection Code of Conduct, which provides specific guidance to service providers regarding how they should handle personal data in the context of federated identity management. To summarize, this Code of Conduct states that service providers are required to only use the attributes and data necessary for providing access, use as little data as possible, and to delete or anonymize data when it is no longer needed. This document is located on the

Table 2. Publishing use cases.

Scenario	Attributes
1. Users access full-text articles with no options for personalization	• Anonymous token
2. Users get content recommendations in UI based on prior history	• Pseudonymous ID
3. Faculty have the ability to purchase e-books using library funds	• Pseudonymous ID • User role
4. Clinicians receive e-mail confirmation of continuing education credits received	• Pseudonymous ID • User e-mail address (with user consent)

GÉANT Project website: https://geant3plus.archive.geant.net/uri/dataprotection-code-of-conduct/V1/Documents/GEANT_DP_CoC_ver1.0.pdf

Considerations for Librarians

A stringent, minimalist approach to handling attributes is excellent for user privacy but may overlook the robust benefits that are possible when utilizing federated identity management. Ultimately, it will be incumbent on each organization to identify the right balance between providing only access vs. providing access in addition to other value-added services. To assist libraries in understanding how to manage attribute release with their vendors and publishers, SeamlessAccess is convening a working group that will be dedicated to generating reusable contract language for libraries. It will consider attribute release policies, provide assurances that these policies are being adhered to by our vendors and service providers, and will provide libraries with templates that can be used in contract negotiations regarding the sharing of personally identifiable information during authentication and access.

In addition, SeamlessAccess is currently working on standardized attribute policies and configurations for various library use cases. These will include templates for generating more granular and robust resource usage data, for implementing chargeback agreements with other campus departments, and for potentially building more robust personalization services. These could potentially include allowing researchers to save data, references, and advanced search queries, or for customizing the user's library experience based on coursework, interests, and demographics.

It's important for librarians to understand that SeamlessAccess is a mutually beneficial initiative that benefits all stakeholders including researchers, library patrons, librarians, service providers, and various support organizations. As librarians, it's not necessary to believe that this initiative is being foisted upon us by a third party whose primary concern isn't the library or what's in the best interests of the library patron. There will continue to be numerous opportunities for librarian participation in upcoming SeamlessAccess initiatives.