Collaborative Librarianship

Volume 10 | Issue 3 Article 3

12-4-2018

A Problem None Can Solve Alone: RA21 as Collaborative Effort

Jill O'Neill
National Information Standards Organization, joneill@niso.org

O'Neill: A Problem None Can Solve Alone

From the Field

A Problem None Can Solve Alone: RA21 as Collaborative Effort

Jill O'Neill (joneill@niso.org)
Director of Content, National Information Standards Organization

Abstract

Stakeholders in the information community recognize the challenges surrounding user authentication in the context of licensed information resources. Resource Access in the 21st Century (RA21) is one cross-sector initiative that is intended to reduce those challenges for both academic libraries as well as content and technology providers. Further collaboration by stakeholders may assuage some of the hesitations regarding RA21.

Keywords: collaboration, rights access management, scholarly communication

This article is opening with a truism. We operate in an increasingly complex digital working environment; systems interact invisibly behind the "wall" of an interface without much attention from the user. It matters little whether the worker is a corporate vice president or a college freshman. Users assume the system will recognize (with minimal interrogation) their right to access information at the point of need. Should it matter what network or computing device (desktop, laptop, mobile, etc.) is in use? The exception being, the inconvenienced user, there is not much attention paid to the labyrinthine processes required of stakeholders to enable authenticated access.

Behind the scenes, of course, there is not a simple one size fits all mechanism for identity and authentication management (IAM). In particular, academic institutions are of different sizes with diverse populations. Each operates under different sets of budgetary constraints. For practical reasons, the approach to authentication management in place at a small teaching college is not the same approach adopted by a statewide system or by a single Carnegie, I research university. An administrative choice of relying on IP ranges (proxy servers) or single sign-on is a decision point according to such variables as staff and financial and technological resources. The perception for many information professionals is that IAM is really a concern to be addressed at the administrative level (who will likely assign it to some local IT unit) or remotely by the platform provider.

Even at the library level, there are many practical questions. What are the trade-offs? Is the user experience so very negative that a change is necessary? What solutions are on offer? Which of the options are affordable by a particular institution? Can the change in delivery management occur at the library level or does that change require buy in from other administrators on campus? Will it scale?

System security and maintenance frequently fall outside of the library's purview and other campus IT units may not have a sufficiently close working relationship to understand why identity/access management represents a concern for libraries. Administrators may not view barriers faced by users in accessing information as a particular priority. Finally, as institutions of higher education continue to wrestle with budgetary constraints, there is little enthusiasm for incurring new overhead costs associated with membership in one of the emerging trust federations or for extended licensing and support of advanced identity management software. Particularly in the United States, given the wide range of institutions in terms of resources, enrollments, and funding, it is unlikely that there can be an easy, out-of-the-box solution suited to all.

Unquestionably, this problem meets the definition of a wicked problem – one where there may be conflicting requirements or constant flux in determining the appropriate criteria for a solution. In such a situation, a solution can only work if it is the result of cross-sector cooperation and collaboration. Librarians, platform providers, and IT professionals gathered under the aegis of an organization such as NISO, working on such initiatives as the Resource Access in the 21st Century (RA21), are more likely to identify a sensible pathway towards meeting the user expectation of "how things work".

Current technology and user behaviors for working with technology have reached a point where it is advisable that the information community revisit existing mechanisms for managing authentication and access. Many within the community have begun thinking about how best to address the problem and preferred solutions

Rationale for RA21

The initial impetus behind the RA21 initiative arose from librarians themselves, those deeply involved with researchers and the workflows of scientific research and development. The Pharma Documentation Ring (http://www.p-d-r.com/) is an association whose members represent the corporate information centers of the large, international companies working in drug development. The information centers of these firms were seeing that researchers required the means to connect to appropriately licensed materials via external (off-campus) networks. Further, those connections required the use of a variety of devices and occurred in changing geographic locations. The demonstrable inadequacies of IP-based authentication in that context deepened further by on-going cycles of mergers and acquisitions of companies. Managing IP addresses required extensive oversight by content providers as well as their pharmaceutical customers. An authenticated IP address belies the assumptions that the user is from a particular geographic location. Even when working in a realm of proxy servers and virtual private networks, such an assumption was faulty. Logically, approved access should be a matter of user identity rather than physical positioning.

Ultimately, two cross-sector membership organizations – the International Association of Scientific, Technical, and Medical Publishers (STM), headquartered in Europe, and the National Information Standards Organization (NISO), headquartered in the United States – jointly undertook the responsibility of shepherding discussions. The intent was not for content and technology providers to dictate to customers, but rather to expand awareness of proven approaches to authenticate and spotlight successful practices already in place at some research universities. What kind of implementations would cause the least disruption to the information community? What best practices or guidelines academic institutions working to accommodate non-traditional workflows – those in use by researchers and faculty as well as rising student populations, might use? How might trade-offs as well as costs be minimized?

O'Neill: A Problem None Can Solve Alone

The International Association of STM (https://www.stm-assoc.org/) has over 120 members from 21 countries. That membership encompasses a diverse group of commercial, society, and professional association publishers, enabling access for a variety of disciplinary communities to scholarship and research data hosted on multiple platforms.

NISO (https://www.niso.org) numbers in its membership more than 130 national, academic, and research libraries in addition to a cross-section of content and technology providers.

In both of these communities, the stakeholders – libraries as well as those who supply and host content – have a desire to improve the user experience equally and to improve system efficiencies and functionality.

What is the Answer? A Collaborative Approach

User frustrations arise when the individual faces interruption in accessing sought-after materials by a series of confusing and potentially obscure prompts. Users may not have all of the needed log-ins or passwords to pass through such security measures. Indeed, all the user may mentally register is that the process requires more than five clicks (and frequently as many minutes) to reach the desired content. Given users' perception that access should be nearly instantaneous, this process is both frustrating as well as time-consuming. Such frustration makes it difficult for users to resist the convenient access offered by alternative (if potentially problematic) services, such as SciHub.

Ideally, achieving a better pathway for users occurs through the implementation of federated identity or access management. Trust federations are groups of stakeholders coming together in collaborative agreements to enable a common channel of exchange of data in situations of high volume. The most immediately recognizable example of a trust federation would be cooperative arrangements between banks and merchants in extending credit. A trust federation in the context of higher education in the United States would be InCommon (www.incommon.org). A student, researcher, or faculty member receives a single credential or token indicating authorized recognition of his or her entitlement to access specific systems and resources. The credential or token exchange occurs swiftly between the Identity Provider (the institution or corporate enterprise) and Service Provider, resulting in rapid access for the authorized individual. Speed of access would not be the sole benefit. Implementation could allow a more robust integration between information resources and automated workflows; consequently, there would be fewer interruptions to the researcher's discovery and thought processes. The enhancement is a system security that eases the concerns of libraries and their parent institutions as well as service providers. In a period when libraries produce indicators of return on investment, single sign-on systems offer further usage and behavioral data as well as increased visibility to the patron community.

The underlying technology favored by those interested in fostering this type of single sign-on approach and the one currently in use in existing RA21 pilots is SAML (Security Assertion Mark-Up Language). SAML – an open standard – emerged from a cross-sector collaboration between corporate enterprises, non-profits, and government agencies. Systems that have implemented that standard can exchange information with minimal information about the user being required, essentially needing just a numerical ID token and an indicator of affiliation with an authorized group. The technology satisfies the need for efficiency in that the user no longer has to supply the information multiple times. Content suppliers gain the information associated with the user's attributes (whether the individual is a walk-in user, a faculty member, and a

graduate student) rather than obtaining the data indicative of the user's identity.

That is important, given that there are specific use cases requiring special attention. For example, those working in medical libraries are well aware that the use of physician credentials allow an individual to sign-on to hospital systems in highly specific contexts. Only through mutual discussions and collaboration can information community stakeholders develop the best means for enabling access in specific use cases. It is for that reason that NISO, with its diverse membership of libraries, technology providers, and content suppliers, holds the position to build trust and facilitate development of best practices and guidelines.

Aside from RA21, emerging from corporate entities such as Google, are initiatives such as Google CASA (Campus Activated Subscriber Access), which works in conjunction with Google Scholar's Subscriber Links program to solve the same issue for researchers seeking access to materials while working off campus. A number of content providers serving a variety of research communities – JSTOR, Project Muse, Gale − as well as platform providers such as Highwire have enabled support for CASA.

Resisting the Shift

The resistance of academic libraries to embrace the enhancements that single sign-on approaches appear to offer has surprised some.

At a recent NISO Forum on the topic, Cody Hanson, Director of Web Development for the University of Minnesota Libraries, provided his own and others' rationale for resisting a call to move to single sign-on approaches. (A recording of his presentation is freely accessible at https://showcase.dropbox.com/s/Digital-Libraries-Authentication-Access-Security-for-Information-Resources-UqqQcjJMiwIEI0wfPBaKI.) Hanson stressed the following value derived from adhering to the older technological approach:

(1) Privacy Protections. Citing the American Library Association's code of ethics as well as legal requirements imposed by the laws of Minnesota as well as by federal authorities, Hanson noted that protection of patron privacy is a central tenet for the profession. User interactions with resources, either a digital resource or a physical publication, remains confidential. Even when inadvertent, usage becomes associated with identity, and that connection becomes potentially sensitive information. For the University of Minnesota Libraries, use of a proxy server acts as a firewall for individual identity as it is submerged in the aggregated dataset.

(2) Security. Academic institutions are *keenly* aware of the need to guard against security breaches resulting from compromised user accounts and, for some university systems, use of Shibboleth protections may already be in place. The protection of campus information systems containing relevant data pertaining to student financial aid or medical information, and banking information associated with payment of taxes and payroll, is carefully undertaken. However, in the context of wrongful access to information resources, the verification of claims of abuse need to delineate that the issue is because of bad intent rather than user error or overly sensitive algorithms. Sticking with IP ranges and proxy servers allows the library the diligent opportunity to verify claims of abuse on behalf of publishers and content providers without unnecessarily revealing the identity of the user.

(3) Business Intelligence. The perception is usage data independently gathered through server logs are an invaluable mechanism for analyzing which user populations are accessing a particular resource or as an independent check against vendor-supplied usage statistics. Given the millions spent at academic institutions across North

O'Neill: A Problem None Can Solve Alone

America, such reliance on internal reporting simply represents good stewardship.

Speaking at the same event, Tim Lloyd, CEO of Liblynx, provided his own experience-driven rationales for why consideration of the shift to single sign-on might still be in the best interests of an institution.

(1) Cost of Maintenance. One significant consideration put forward was that the continued use and regular maintenance of web proxies represent a substantial drain on the institution. Changes to web-based practices and protocols (such as the use of cookies in the wake of the European privacy legislation, general date protection regulation, or the move to https://) require the time and attention of information technology staff for appropriate handling. Given that institutions handle the deployment of staff so differently depending upon their missions and enrollments, the costs associated with maintaining proxy servers is significant.

(2) Reporting User Behavior. Current proxy servers allow usage tracking by the library alone. That represents a defense for libraries concerned with patron privacy. However, there are misunderstandings about the benefit to the publisher's access to user data generated via the single sign-on approach. Properly implemented by the institution, the approach does not yield the volume of data needed to generate robust personalized services.

As an additional protection, Lloyd referenced the "scaleable consent" initiative whereby an organization provides tools and policies for the user that allows the individual to make their own choices about what information or attributes of the user releases to the service provider. Such an approach exists at Duke University. (See Lloyd's presentation as well as those of other speakers at this event at https://www.niso.org/events/2018/05/digital-libraries-authentication-access-security-information-resources)

(3) Risks to the User. Continued use of IP filtering represents a risk to users as well as to their institutions. IP filtering is insecure and easily exploited by bad actors. Lloyd characterized such systems as "messy" and "soft targets for fraudulent access", noting at least one instance where the access licensed by one UK institution had actually been attributed to an entirely different one. The virtue of moving to a federated trust approach would be an improved capability for libraries to formulate security protocols that would be more precise and secure. Users, prone to using duplicative or weak passwords, would have just the single numerical token, readily tracked by the institution.

The Call to Become Involved

For librarians who are committed to ensuring positive user experiences in discovering and navigating to relevant content, recognizing the investigating and implementing authentication management systems is a priority. User behaviors are unlikely to change; mobility and device preferences will continue to drive demands placed on libraries and information resources.

That being the case, it is critically important that professionals from the academic community engage with content and system providers to ensure that proposed approaches satisfy library needs. Engaging in such constructive discussions allows library professionals to offer key insights and ensures that appropriate solutions for the academic environment emerge in timely fashion. At the same time, the engagement with providers allows universities to prepare and budget for this opportunity to improve the user experience. As with any collaborative effort, the first step is to express interest in participating in the process. Reach out to professionals working with RA21 or similar initiatives and engage them of your enthusiasm for building a better

set of solutions for identity and authentication management.

There are other ways to signal your interest in engaging with this issue.

- Ensure familiarity of library colleagues, campus IT professionals, and administrators with the advantages associated with membership in one of the emerging Trust Federations for institutions of higher education. Those advantages might include more reliably secure web-based access to information services, reduction in IT costs associated with support and maintenance of proxy servers, and minimization of abuse claims that require staff intervention.

- Document data collection from subscription services and engage in discussions with those providers about their rationale for collection and/or transferal of specific data elements through system APIs. Invite those working in campus IT to join in those discussions. Precisely identify the data for collection that each party feels necessary or (alternatively) wishes to avoid retaining. Doing so may help to eliminate risks for all stakeholders.

- Work through cross-sector associations such as NISO to develop useful materials (something like an Authentication Management 101 course or primer) with vendors and other stakeholders. Doing so can ensure common understanding of terminology and processes in investigating and implementing authentication management systems. Such aids might include an overview of the existing landscape, a glossary of terminology, the identification of the roles of participants, the identification of challenges and options, and the outline of a decision path.

Most importantly, bear in mind that these discussions about upgrading the means of authenticating users are not because of a desire to add superfluous bells and whistles to existing infrastructure. Rather, the intent is to enhance the "under the hood" processes that support the active use of the infrastructure. Digital environments do not remain static; even in the absence of true disruption, human engagement with those environments drives change.

The position held by academic librarians allow them to observe how students, faculty, and research professionals react to technical "speed bumps" that hinder their workflow. By working closely with campus IT groups, librarians can influence the thinking about solutions to better satisfy the needs of their users, and through early collaboration with content and platform providers, librarians can ensure the applicability of those solutions to the use cases encountered in their institution, which will prove in the best interests of their communities.