Information Services & Use 40 (2020) 161–175
DOI 10.3233/ISU-200091
IOS Press

Introduction to SeamlessAccess

Tim Lloyd[*]
CEO, Liblynx, Alexandria, VA, USA

Abstract. This paper provides a general introduction to the SeamlessAccess (SA) initiative, covering the following four topics: Why are we here? A recap of why was SA created and the history of this initiative to date; How does it work? A walk through how Federated Authentication and SA work, so you are familiar with some of the technical terms and we can build on that knowledge when we address privacy and security; What is the current status? An update on the status of this initiative; How do I participate? Information on how you can participate to deliver more seamless access to your users.

Keywords: SeamlessAccess, federated authentication, user access experience, RA21

1. Why are we here?

1.1. Why SA is needed

Why do we need improvements to the user access experience? Isn't this a solution looking for a problem? Proxy solutions like EZproxy only cost libraries $500 a year and work just fine.

Library use of IP recognition was developed when off-site access to electronic resources was in its infancy and has changed little since then. EZproxy was a godsend at a time when it was difficult and unwieldy to get a proxy server setup and then support users trying to make configuration changes to their browser, and then needing to remember to undo these. But, after twenty years of IP authentication, there are better alternatives.

SA has a number of important goals.

Firstly, to improve remote access scenarios. There has been a considerable increase in remote access to online resources over the years - from a multitude of devices. With IP authentication we force researchers to start from, or at some point circle back through, the library's web site to find a proxy-prefixed URL, extra friction that simply deters users. This is not how researchers research. We should aim for delivery at the point of discovery.

Secondly, to improve the usability of access workflows. Current issues include the numerous clicks to reach content behind an authentication barrier, and the numerous user credentials scattered over a multitude of platforms; libraries have not kept pace with the consumer web. If an institution is using VPN, Shibboleth, and EZproxy, users face multiple, inconsistent access experiences and can quickly feel confused and overwhelmed.

Access is currently so complicated that even fully-entitled end-users are turning to alternative resources, such as SciHub, ResearchGate, etc.

[*] E-mail: tim@liblynx.com.

Finally, we want to personalize user services capabilities and enhance user privacy mechanisms, and we will address both issues later in this document.

1.2. Security concerns and workflow issues

There are also some important security concerns and workflow issues with traditional IP authentication.

The perceived wisdom is that IP authentication is more privacy-preserving than federated authentication. It is not. IP addresses can directly identify individuals with a dedicated IP address (GDPR considers IP addresses as Personal Data in some circumstances), and an inadvertent side effect is that users can be forced to pass over individual credentials in order to benefit from personalization (personal data that, all too often, can compromise their privacy and security).

Most libraries still use some type of URL rewriting methods (EZproxy, WAM) which have security vulnerabilities, such as:

- IP Spoofing/Man-in-middle attacks - especially over wireless networks.
- Clickjacking - user credentials are sometimes prompted within an iframe - easy to exploit.
- Session hijacking - session cookies are sometimes sent in the clear.

Blocking IP addresses is a blunt tool: if a vendor identifies one compromised user account, access is cut off for the entire institution - instead of for one compromised or malicious user account.

Finally, IP addresses themselves are subject to change (sometimes without the knowledge of the library), often stored in a variety of formats, and pass through many hands between IT, the library, and service providers. “On average, 58% of the IP ranges held by publishers to authenticate libraries who license their content are inaccurate” (Publishers Solutions International, 2017 [1]).

You will find similar concerns expressed by groups like FIM4L (Federated Identity Management for Libraries [2]), a library-led working group looking to improve access scenarios via federated authentication.

1.3. Compromised user credentials

Another major concern is the volume of compromised user credentials available on the web, typically including a link to the list of library resources that can be accessed using them. It took the author only five minutes of searching to find multiple examples of compromised credentials available on shared Google docs and web-based chat groups.

While compromised credentials impact both IP authentication and Federated authentication, the big difference is that you can far more easily identify and shut down access to compromised credentials with Single Sign-On. In contrast, it can take days of painstaking analysis through logs to trace the access back from an IP address through the proxy server to a physical computer, and then back to a specific login.

1.4. Who can benefit?

Access issues are often presented in the context of patrons accessing library resources. While that is an important use case impacting millions of users across the world, the benefits of SA are much wider. For example:

(a) Research collaborations

A common feature of research is the need to collaborate with colleagues in the field at other institutions - for example, to share research and datasets. Unlike most library access, research collaborations involve sharing much more detailed information, such as a user's name, email address, role, and department. The authentication challenges faced by researchers collaborating across institutions has long been recognized.

Examples of organizations working to address researcher needs include:

- AACR (Authentication and Authorization for Research and Collaboration [3]), an initiative launched in May 2015 to address the increased need for federated access and for authentication and authorization mechanisms by research and e-infrastructures); and
- FIM4R (Federated Identity Management for Research [4]), a collection of research communities and infrastructures with a shared interest in enabling Federated Identity Management for their research cyber infrastructures - ranging from the arts & humanities (DARIAH, see: https://www.dariah.eu/) through to photon & neutron science (Umbrella, see: https://umbrellaid.org/).

(b) Institutional workflows

Another area where SA can help is in institutional workflows that require users to confirm their institutional affiliation with third parties. A good example is the growing need to authorize the use of institutional funds for open access publishing fees, such as Article Processing Charges.

1.5. What was the RA21 project?

Let us briefly recap the work that has been done to date, starting with the Resource Access in the 21st Century project (or RA21) that was initiated 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 sixty organizations over three years.

It identified that Federated Authentication held the most promise for providing a robust, scalable solution for remote access to scholarly content. It also investigated barriers to take-up, developed best practices, and piloted technical approaches to simplifying access.

RA21's conclusions were published as a draft NISO Recommended Practice last April, receiving >200 comments that helped identify further areas for investigation and confirmed the value of testing a beta service. A final NISO Recommended Practice was published in June 2019 [5].

1.6. What is seamlessaccess?

SA was created in July 2019 as a community-driven effort to enable seamless access to information resources, scholarly collaboration tools, and shared research infrastructure.

To date, we have five founding organizations:

- The National Information Standards Organization (NISO [6]); organizers of the conference at which this paper was presented.
- GÉANT [7]: a European Research & Education (R&E) network that operates a service called eduGAIN [8] that connects over sixty R&E identity federations around the world. GÉANT also support related initiatives, such as FIM4L.
- Internet2 [9]: a US Research & Education network that operates the US identity federation, InCommon, among many other activities.

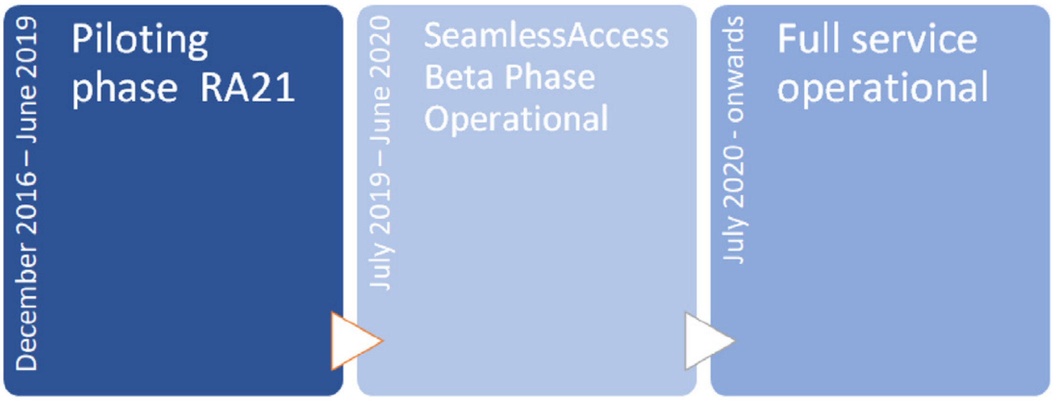


Fig. 1. SA timeline.

- ORCID [10]: a non-profit that provides researchers with a digital, persistent identifier; and
- The International Association of STM Publishers [11].

We are actively seeking a partner to represent the library community.

Think of us as the operational successor to the RA21 project, delivering an operational service plus best practices and standards.We have a full-time implementation team including an experienced librarian technologist, dedicated to library outreach. There are governance and advisory committees with representatives from across the stakeholder groups, as well as an outreach committee including six institutional participants. We also have two cross-industry working groups that I will describe in more detail lower down.

Figure 1 shows an overarching timeline for our activities to date and going forward.

To summarize, the RA21 project developed and piloted ideas from the end of 2016 until last June, when it wrapped up. SA is now in the process of testing these ideas in the light of community feedback and developing best practices around the use of federated authentication. This beta phase is scheduled to run until June 2020. From July 2020, our goal is to have the service fully operational.

2. How does it work?

2.1. A brief refresher on Federated Authentication

While IP authentication is an easy concept to grasp, Federated Authentication is more complex and much less well understood. It is laden with jargon and acronyms that obscure a series of transactions that really are not complex to grasp at a basic level.

If you are unfamiliar with the term Federated Authentication, you may recognize the name Shibboleth instead - Shibboleth [12] is an open source software commonly used to implement Federated Authentication. Let us start by walking through a simple analogy for how Federated Authentication works:

- Bob runs a conference booth that provides books to anyone who studies at a subscribing institution.
- Amy comes up to the booth and says “Hi, can I have a book?”
- Bob says, “Sure” and asks her if she is at a subscribing institution.
- Amy says that she is a student at ABC College.

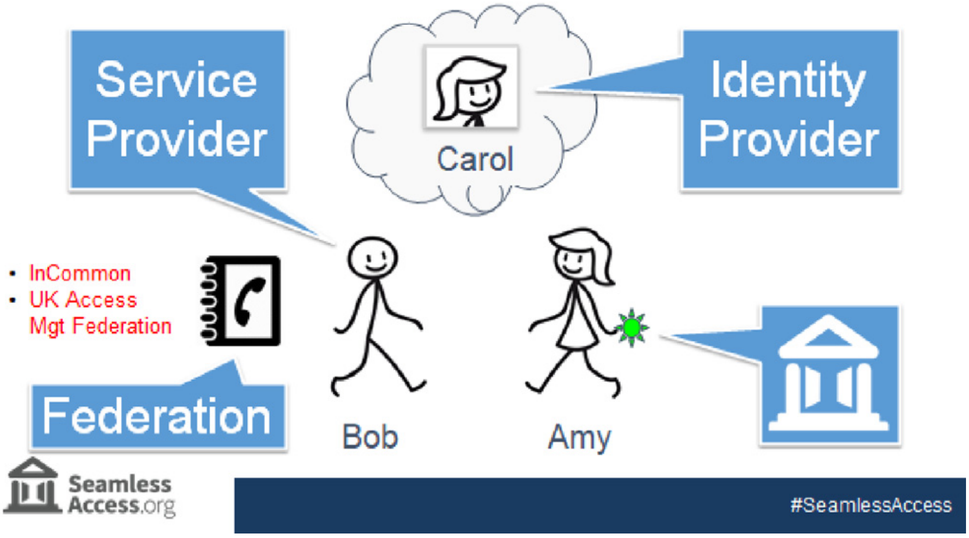


Fig. 2. How federated authentication works.

- However, Bob does not know Amy, so he needs to verify that she is registered with ABC College.
- Luckily, he has a phone book where he can look up someone who can help him. In the case of ABC College, the person to talk to is Carol.
- Bob calls Carol to ask if she can confirm that the person at his booth is a student at ABC College.
- Carol asks him to pass the phone to the student so she can talk to her directly.
- Carol talks to Amy and can confirm that she is a valid student.
- Amy passes the phone back to Bob so that Carol can confirm to him that she is a student at ABC College.
- Bob would ideally like to know her name so that he can learn more about her interests and recommend other books to her in future.
- However, ABC College's policy is not to release student names and so Carol cannot provide Bob with any additional information on the student.
- Bob has now verified that the student in front of him is at ABC College.

Bob gives Amy her book, and also gives her a bright green badge to wear that says "I'm with ABC College" - Bob tells her that if the other booths see that badge, it will save some time as she won't need to tell every booth at which institution she studies This simple scenario is actually very close to how federated authentication works (see Fig. 2)!

Bob is the Service Provider or SP that needs to check a visitor's institutional affiliation before providing access to services. His phone book is an identity federation - a trusted list that details how to talk to a set of vetted institutions and vendors. Examples of identity federations in Higher Education include InCommon [13] in the United States, and the UK Access Management Federation [14].

Carol is the Identity Provider or IdP - the institution's federated authentication service that confirms a visitor's identity. And while our characters in this scenario speak English, in reality Bob, Carol and the Federation communicate using a language called Security Assertion Markup Language, or SAML for short.

Finally, the badge that Bob gives to Amy is what SA is about - making it easier for Amy to deal with other service providers.

It is important to note that Carol, as the Identity Provider, was in control of Amy's identity and opted not to share any information about Amy with Bob, such as her name. All Bob got was confirmation that Amy was affiliated with ABC College and, as Bob trusts the phone book, he trusts Carol is the right person to confirm that.

In Federated Authentication, Identity Providers control user privacy by deciding whether to share extra user information, known as Attributes, with a Service Provider. An attribute might be affiliation information, such as a department or role, or more personal information such a name or email address. In this example, no attributes were shared.

2.2. *Privacy, Attributes and why they are important*

Having understood the basics of what a Service Provider, an Identity Provider, and a Federation are, let us talk about how user privacy works in Federated Authentication.

Attributes is the term used to describe data about an authenticated user, and 'Attribute release' is the process by which that data is shared by an Identity Provider (such as a research and education institution) with a Service Provider (such as a publisher) as part of the authentication process. The format an attribute takes depends on the underlying technology. For example, Security Assertion Markup Language [15], or SAML for short, is the technology that underpins Shibboleth and OpenAthens, but there are other technologies that support Federated Authentication, such as OpenID Connect [16], which is used by consumer-focused services like Facebook and Google.

Here are some examples of the types of attribute that can be passed as a result of a successful user authentication:

- An anonymous token is one that is uniquely generated for every login and for each Service Provider, regardless of whether the user is new or returning. This token cannot be used to support personalization because it changes every time the user signs in, and so 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 that they visit (but cannot be used to build a pattern of usage across Service Providers). This can 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.
- And there are also personal data fields, such as your name and email 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:* an institution 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 hassle) 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 email address, and without asking for a password.

Attribute release is optional - an Identity Provider can simply assert that a user is a member of their organization and do nothing more. And it only happens after a user is authenticated. A Service Provider cannot pull attributes - they only receive what the Identity Provider chooses to send. Attribute release is configured by the Identity Provider for each category of Service Provider. Library resource access is only one of several valuable use cases for federated authentication. For example:

- Research Collaborations involving researchers across different institutions would typically share some personal data, such as a name and email address.
- Institutional workflows that require users to confirm their institutional affiliation with third parties may involve scenarios where it is appropriate to share a much broader range of user data, such as authorizing the use of institutional funds for open access publishing fees.

With library resources, our recommendation is for a much more limited set of attributes. Because the Identity Provider is in control, any special needs for attributes need to be agreed in advance so that attribute release can be configured appropriately.

Having explained what Attributes and Attribute Release are, let us talk briefly about the problem with them, which revolves around configuring access.

To avoid Identity Providers having to manually configure exactly which attributes to send to each Service Provider, configuration is managed through Entity Categories. An entity category is a metadata tag used to group entities like Service Providers or Identity Providers so that profiles can be built and applied at the group level, rather than the individual entity level.

However, the most well-known entity category in use today is the REFEDS Research & Scholarship (or R&S) entity category [17]. REFEDS [18] is the Research and Education FEDerations group, which represents the global research and education identity federations. This entity category only applies to Service Providers that are "operated for the purpose of supporting research and scholarship interaction, collaboration or management, at least in part." It cannot be used for access to licensed online resources. This means there are no standards for how Identity Providers should release attributes for the many use cases that fall outside of the R&S entity category - such as library access to licensed resources.

2.3. *How SA works: the user experience*

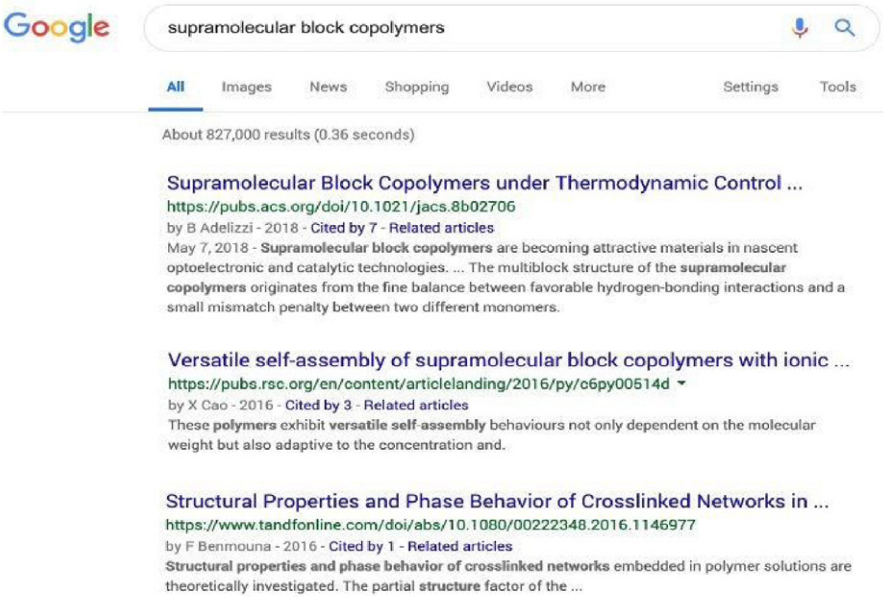
Now let us look at how SA works, starting with how a user experiences the service. Currently, it is a bit of a Wild West out there when it comes to the user experience for institutional authentication. Users may not understand terminology like "institutional login." The institutional login can be hard to find, and the authentication experience can feel quite different across different sites.

Let us use the example of a user searching Google for 'supramolecular block co-polymers' (see Fig. 3). In this example, the first result is an American Chemical Society (ACS) journal article.

Clicking on that article gets us to an article page where we are invited to 'Access through your institution' using a standard SeamlessAccess login button (see Fig. 4).

Clicking on that button for the first time prompts us to search for our institution (see Fig. 5). As we type, a list is dynamically generated. In this case, we are selecting Boston College as our institution.

We get forwarded to the Boston College Shibboleth software where we enter our Boston College credentials and are authenticated as a Boston College user. We are then returned to that same ACS article page and, this time, the 'Access through your institution' button is replaced by a link to view the full text as a pdf (see Fig. 6).



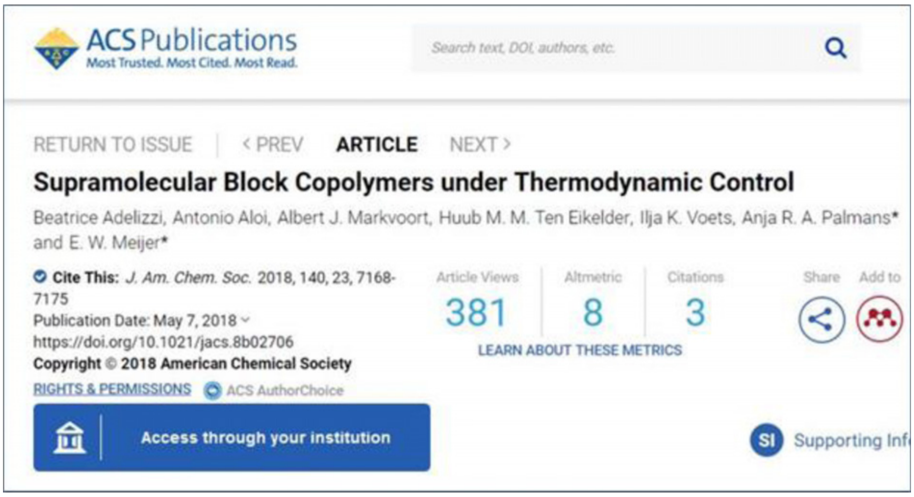
About 827,000 results (0.36 seconds)

Supramolecular Block Copolymers under Thermodynamic Control ...
https://pubs.acs.org/doi/10.1021/jacs.8b02706
by B Adelizzi - 2018 - Cited by 7 - Related articles
May 7, 2018 - **Supramolecular block copolymers** are becoming attractive materials in nascent optoelectronic and catalytic technologies. ... The multiblock structure of the **supramolecular copolymers** originates from the fine balance between favorable hydrogen-bonding interactions and a small mismatch penalty between two different monomers.

Versatile self-assembly of supramolecular block copolymers with ionic ...
https://pubs.rsc.org/en/content/articlelanding/2016/py/c6py00514d
by X Cao - 2016 - Cited by 3 - Related articles
These **polymers** exhibit **versatile self-assembly** behaviours not only dependent on the molecular weight but also adaptive to the concentration and.

Structural Properties and Phase Behavior of Crosslinked Networks in ...
https://www.tandfonline.com/doi/abs/10.1080/00222348.2016.1146977
by F Benmouna - 2016 - Cited by 1 - Related articles
Structural properties and phase behavior of crosslinked networks embedded in polymer solutions are theoretically investigated. The partial **structure** factor of the ...

Fig. 3. How SA works - an example google search.



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

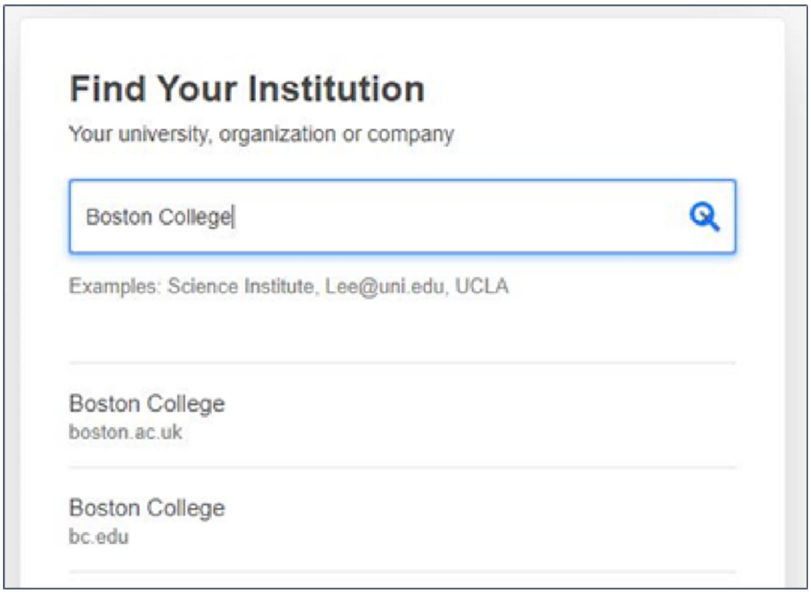
Article Views	Altmetric	Citations
381	8	3

Access through your institution

Fig. 4. How SA works - ACS website example, prior to authentication.

Great, so let us go back to our search results and try the 2nd result, which is a Royal Society of Chemistry (RSC) article. When we arrive at the RSC journal article page, we see the same standard SA button to ‘Access through your institution’ except that, now, we are offered the chance to access via Boston College (Fig. 7). No need to repeat the same process to find our institution.

Note that underneath this button is an option to ‘Access through another institution’ in case I wear several hats and need to authenticate with one of my other institutional identities.



Find Your Institution

Your university, organization or company

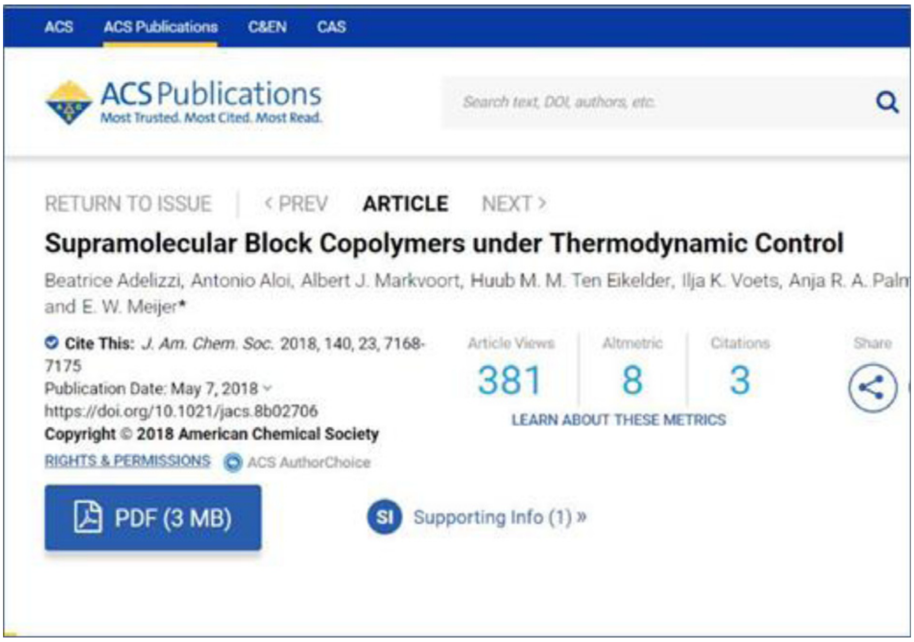
Boston College

Examples: Science Institute, Lee@uni.edu, UCLA

Boston College
boston.ac.uk

Boston College
bc.edu

Fig. 5. How SA works - searching for your institution.



Supramolecular Block Copolymers under Thermodynamic Control

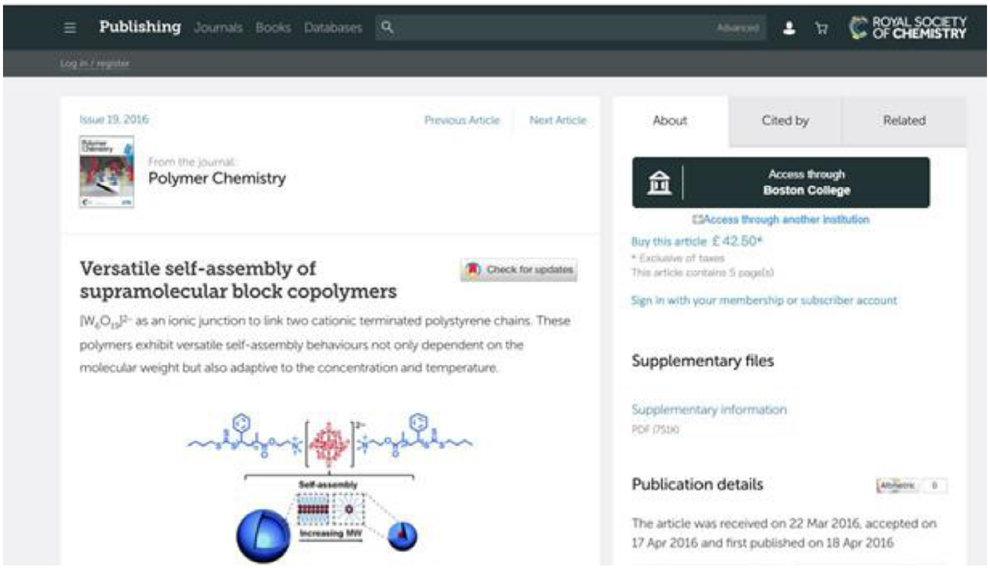
Beatrice Adelizzi, Antonio Aloi, Albert J. Markvoort, Huub M. M. Ten Eikelder, Ilja K. Voets, Anja R. A. Palm
and E. W. Meijer*

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 381 · Altmetric 8 · Citations 3

Fig. 6. How SA works - ACS website example, after authentication.

We will click to access and seamlessly get to the full text this time. This is because I am already logged in to the Boston College network and so, behind the scenes, the SAML-handshake can proceed without me needing to re-enter my credentials.



Versatile self-assembly of supramolecular block copolymers

$[W_6O_{19}]^{2-}$ as an ionic junction to link two cationic terminated polystyrene chains. These polymers exhibit versatile self-assembly behaviours not only dependent on the molecular weight but also adaptive to the concentration and temperature.

Buy this article £ 42.50*
* Exclusive of taxes
This article contains 5 page(s)

Sign in with your membership or subscriber account

The article was received on 22 Mar 2016, accepted on 17 Apr 2016 and first published on 18 Apr 2016

Fig. 7. How SA works - RSC website example, prior to authentication.

It is worth noting one wrinkle to this, which is outside of the control of SA. Identity Providers control how long they maintain your session once you have logged in with your institutional credentials - it may be an hour; it may be a day. I have heard of some institutions that force re-entry every time …yikes.

2.4. How SA works: under the hood

Let us look at what is happening under the hood to enable this. Firstly, SA has a standard visual cue for how a user accesses resources that require an institutional affiliation. This access button either displays a generic ‘Access through your institution’ message or a custom message listing your more recent institutional choice, if relevant. Users always have the option to re-select an alternative institution.

Secondly, the SA Identity Provider Discovery Service offers a standard method for finding your institution that incorporates best-practice design, such as dynamic search results as you type, alternative spellings and acronyms, and displaying institutional logos to simplify selection.

Finally, the SA Persistence Service stores your institutional choices on your computer in local browser storage. This information can only be accessed by applications coming from the SeamlessAccess.org domain and are not stored remotely. And users will be able to opt out (which means they will need to select their institution every time they login). That all looks well and good, but what about Security & Privacy?

SA has adopted the GÉANT Data Protection Code of Conduct [19], which provides specific guidance to Service Providers on how they should handle personal data in the context of federated authentication. It covers four fundamental principles: Purpose limitation, Data minimization, Deviating purposes, and Data retention

In a nutshell, Service Providers should:

- Only use attributes necessary for access.
- Use as little data as possible, wherever possible.

- Not do anything, but provide access with this data.
- Delete/anonymize this data when it is no longer needed.

The Code of Conduct document is remarkably readable, and I recommend it to anyone interested in this area. It also aligns very closely with the American Library Association's library privacy guidelines found in the ALA Code of Ethics [20].

We have also audited our approach. The RA21 project created a dedicated Security & Privacy Working Group to assess the technical security and privacy risks using industry-standard assessment models. This group comprised a variety of stakeholders from across the information spectrum and completed their analysis in July 2018. For security, when applying standard security and data protection practices, no significant risks were found. For data privacy, a data protection impact assessment was performed compliant with the EU's GDPR to determine if "high risks" were involved, and none were found. You're welcome to read the full report at the RA21.org website [21].

2.5. How SA works: integration options for Service Providers

Let us review how Service Providers integrate SA into their Federated Authentication experience. We are testing three options during the beta phase: Limited, Standard, and Advanced. Service Providers will only be able to integrate once they have agreed to our Terms of Service, which we are in the process of developing (and more on that later). Note that the final integration options available at launch will likely be simpler. Let us explore the differences between them.

(a) Limited Integration

In a Limited Integration, the Service Provider website simply sends users to the SA Identity Provider Discovery Service (service.seamlessaccess.org) to determine their preferred Identity Provider. SA does this either by identifying a previously stored choice in their local browser storage, or by asking them to select from a list. Then the user returns to the Service Provider site and their SAML solution manages the authentication process as normal. It is the easiest to implement - in Shibboleth, you simply configure SA as your Identity Provider discovery service.

However, this offers users the least-seamless experience because they are sent to the SA interface for Identity Provider discovery. As this gives the Service Provider no branding or control over this experience, we're expecting this limited option to only be used in scenarios where a Service Provider is unable to use the standard integration, e.g. they don't have the necessary control over the website or SAML solution.

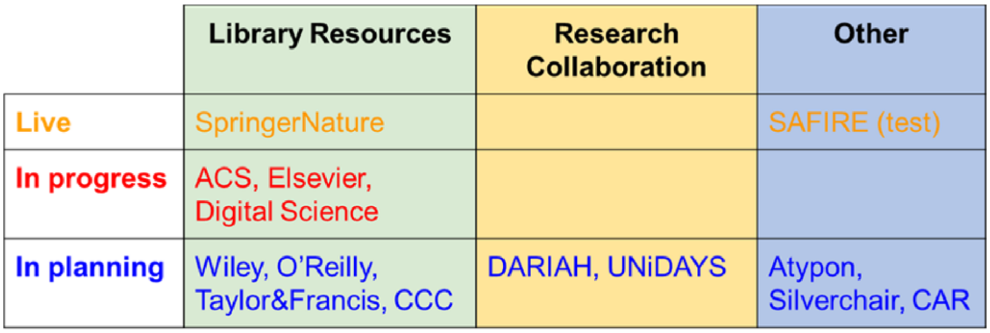
(b) Standard Integration

A Standard Integration is essentially the same process except that the Identity Provider discovery experience is embedded within the Service Provider's existing web experience. This requires more control over the Service Provider application to implement but is not particularly complex. It is 'standard' because we expect most Service Providers to opt for this method as it offers a nice balance of simplicity and control.

(c) Advanced Integration

The Advanced integration offers Service Providers more granular control over their login and discovery user experience. For example, to customize:

- The user workflow; e.g. to display extra info/steps after certain actions (vs the default workflow from SA).
- The list of Identity Providers available in the discovery phase (vs the default list from SA).
- The visual look of the login button within specific parameters (vs the default button from SA).



	Library Resources	Research Collaboration	Other
Live	SpringerNature		SAFIRE (test)
In progress	ACS, Elsevier, Digital Science		
In planning	Wiley, O'Reilly, Taylor&Francis, CCC	DARIAH, UNiDAYS	Atypon, Silverchair, CAR

Fig. 8. Current SA beta implementations.

The API access enabling this more advanced configuration is only available to a set of 'trusted sites,' i.e. whitelisted applications. However, these options can be significantly more complex to implement because of the more specialist technical knowledge required. *Note:* whitelisting is currently by domain, but more secure options will be available when the service is formally launched.

3. What is the current status?

3.1. The beta phase

We are currently in the beta testing phase that you saw in Fig. 3. This means we are:

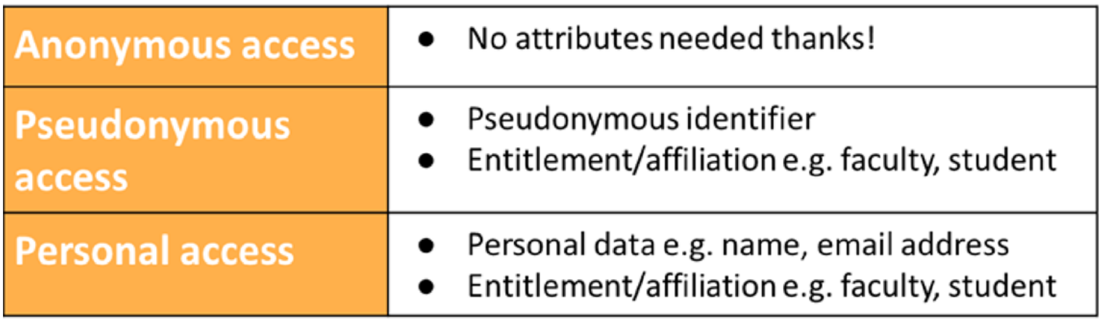
- Testing our ideas in practice.
- Soliciting feedback to help improve our solutions.
- Changing our approach as needed.

We have live implementations in progress because we need to test the service with real users to fully understand how it works in the real-world and to iron out the kinks. This is not a mature, fully developed service yet .... Figure 8 shows you organizations that are currently working on a beta implementation, with their organization type along the top and their implementation status down the left-hand side. As you can see, we are currently live with two implementations (SpringerNature and SAFIRE).

Note that while most of these organizations are library resource providers (green column), we also have a couple of examples of research collaborations, as well as some platform providers and an Internet2 initiative.

Notes:

- DARIAH [22]: Digital Research Infrastructure for the Arts and Humanities (DARIAH) - a network to enhance and support digitally enabled research and teaching across the Arts and Humanities.
- UNiDAYS [23]: Student affinity network with 13m+ registered students across 100+ countries.
- SAFIRE [24]: national academic identity federation for the South African research and higher education community.
- CAR [25]: Consent-informed Attribute Release - an Internet2 initiative.



Anonymous access	• No attributes needed thanks!
Pseudonymous access	• Pseudonymous identifier • Entitlement/affiliation e.g. faculty, student
Personal access	• Personal data e.g. name, email address • Entitlement/affiliation e.g. faculty, student

Fig. 9. Draft recommendation for new entity categories.

3.2. The attribute release working group

The Attribute Release Working Group is tasked with identifying distinct categories of scholarly information resources and users, determining their needs, and developing standard attribute release policies and profiles applicable to each category. Once agreed, we want these categorizations and profiles endorsed and adopted by global research and education federations, library communities, and other stakeholder groups. Example categories include library walk-ins, hospital/clinical settings, research collaborations, and corporate libraries.

This group includes over twenty members from across industry stakeholders including Service Providers, Identity Providers, libraries, federations, and consultants. The end goal is to greatly simplify the process of configuring Service Provider access for Identity Providers, lowering costs, and minimizing the risk of errors. Hot off the presses following recent meetings, we have some draft recommendations that I can share with you. Please note that this is work in progress, and nothing will be completed until it goes through a broad community consultation phase.

The draft recommendation is for the creation of three new entity categories (Fig. 9).

These categories would be asserted by the Service Provider. And, unlike the R&S entity category that we mentioned earlier, which is controlled at the federation level (Identity Providers either implement it for any Service Provider or for none), these entity categories would be controlled at the local institutional level. This means that each identity provider can choose whether to accept them for individual Service Providers. These proposed new entity categories nicely match the attribute types mentioned in Section 2.2.

The Anonymous Access category would be used by a Service Provider who does not need any user attributes. Just confirmation of their organizational affiliation. The Pseudonymous Access category would be asserted by a Service Provider who needs to personalize their service, and would also allow for additional entitlement and affiliation data that could provide more control over access, such as a user's role (faculty vs student). And the Personal Access category would be asserted when a Service Provider needs personal data, such as a name and email address, in addition to the entitlement and affiliation data. This category would assume an appropriate contract is in place between the Service Provider and Identity Provider. As will all these categories, the Identity Provider remains in control and can decide whether to support it.

3.3. *The Contract Language Working Group*

The Contract Language Working Group is tasked with developing contract language templates for library use based on Attribute Release/Entity Descriptions from above. This will give libraries a mechanism to ensure Attribute Release compliance.

This working group is still collecting members and will start work once the Attribute Release recommendations are ready, later in the Spring pf this year.

3.4. *Issues we are working through*

Here are some examples of the sorts of issues we are currently working through:

(a) Terms & Conditions for Service Providers

We are finalizing the terms and conditions that Service Providers will be required to sign to use the SA service. These will control how they can use the service, including the user's institutional choices, and include the ability to deny access to SPs that abuse the service. Every Service Provider needs to be registered with a federation.

(b) User Consent workflow

We are working on the user consent workflow to enable users to opt out of storing their institutional choice in their local browser storage. Opting out simply means your Service Provider will need you to identify your institution each time you login. We will also flag if access to your local browser storage is blocked (intentionally or not) and we are unable to store your institutional choice.

(c) Access to Identity Provider choices prior to authentication

This issue relates to an unintended feature arising from the flexibility available under the Advanced Integration. We are exploring ways to address this, including changes to the advanced API, expanding consent requirements, and clearer prohibitions on use within Service Provider T&Cs. More information can be found on this issue at the SA website [26].

(d) Personal Data

Our collective understanding of what constitutes Personal Data (PD) has changed over the life of the project. The EU's GDPR regulations [27] only came into force two years after the initial RA21 project started in 2016, and since then technologies such as browser fingerprinting have greatly expanded the information that enable organizations to identify an individual. As a result, we are investigating two possible areas where the data we store may fall within the definition of PD.

The first is that the institutional choices we are storing locally (in the form of SAML entity IDs) may be classified as PD when combined with other information held by Service Providers. The second is that user IP addresses may be PD when they can be tied back to an individual, rather than an institution. While we do not formally store IP addresses, they do persist within server logs used in our cloud-based infrastructure. As a result, we are carefully reviewing our approach to fully understand the implications before deciding what steps we need to take.

(e) Feature requests

An example of the feature requests that we are considering as a result of feedback from the beta testing is the ability to customize the list of institutions that users can search within. For example, to exclude organizations that do not have access, or add-in additional organizations that are not listed.

4. How do I participate?

If you are a Service Provider interested in finding out how to make your access experience more seamless, please visit our website [28] and contact Heather Flanagan at contact@seamlessaccess.org.

If you are an Identity Provider wanting your users to benefit from more seamless access, then you simply need to support federated authentication. Nearer the launch date, we will provide more information and materials to help inform users about the forthcoming improvements to their access experience.

If you do not already support Federated Authentication, then you will need to plan for a transition. As with any other software implementation, you will likely want to talk to your IT department, evaluate vended solutions, understand your organization's privacy policies etc.

We are aware this can get complex, and we are working to simplify that process by standardizing approaches to attribute release and contract language.

Finally, if you just want to learn more as an individual, please sign up for our monthly email updates and consider joining the Contract Language working group. Contact us at contact@seamlessaccess.org.

References

[1] Statistic cited at https://www.psiregistry.org/newsl, last accessed June 7, 2020.
[2] https://libereurope.eu/strategy/research-infrastructures/fim4l/, last accessed June 7, 2020.
[3] https://aarc-project.eu/, last accessed June 7, 2020.
[4] https://fim4r.org/, last accessed June 7, 2020.
[5] "Recommended Practices for Improved Access to Institutionally-Provided Information Resources: Results from the Resource Access in the 21st Century (RA21) Project", NISO Recommended Practice published June 21, 2019, available at: https://www.niso.org/publications/rp-27-2019-ra21, last accessed June 7, 2020.
[6] https://www.niso.org/, last accessed June 7, 2020.
[7] https://www.geant.org/m, last accessed June 7, 2020.
[8] https://edugain.org/, last accessed June 7, 2020.
[9] https://www.internet2.edu/, last accessed June 7, 2020.
[10] https://orcid.org/, last accessed June 7, 2020.
[11] https://www.stm-assoc.org/, last accessed June 7, 2020.
[12] https://www.shibboleth.net/, last accessed June 7, 2020.
[13] https://incommon.org/, last accessed June 7, 2020.
[14] https://www.ukfederation.org.uk/, last accessed June 7, 2020.
[15] https://en.wikipedia.org/wiki/Security_Assertion_Markup_Language, last accessed June 7, 2020.
[16] https://en.wikipedia.org/wiki/OpenID_Connect, last accessed June 7, 2020.
[17] https://refeds.org/category/research-and-scholarshipm, last accessed June 7, 2020.
[18] https://refeds.org/, last accessed June 7, 2020.
[19] https://geant3plus.archive.geant.net/uri/dataprotection-code-of-conduct/V1/Documents/GEANT_DP_CoC_ver1.0.pdf, last accessed June 7, 2020.
[20] http://www.ala.org/advocacy/sites/ala.org.advocacy/files/content/proethics/codeofethics/Code%20of%20Ethics%20of%20the%20American%20Library%20Association.pdf, last accessed June 7, 2020.
[21] https://ra21.org/wp-content/uploads/2018/07/RA21-Security-Privacy-Final-Report.pdf, last accessed June7, 2020.
[22] https://www.dariah.eu/last accessed June 7, 2020.
[23] https://www.myunidays.com/, last accessed June 7, 2020.
[24] https://safire.ac.za/, last accessed June 7, 2020.
[25] https://spaces.at.internet2.edu/display/CAR/CAR%3A+Consent-informed+Attribute+Release+system, last accessed June 7, 2020.
[26] https://seamlessaccess.org/posts/2020-01-13-clarifications/, last accessed June 7, 2020.
[27] https://gdpr-info.eu/, last accessed June 7, 2020.
[28] https://seamlessaccess.org/services/for-service-providers/, last accessed June 7, 2020.