Resource Access for the 21st Century (RA21)
Corporate Pilot Report
September 2018

Final

Contents

Executive Summary:

RA21 is a joint initiative of the International Association of STM Publishers (STM) and the National Information Standards Organization (NISO) with the goal of improving access to scholarly resources, from anywhere and on any device. Its purpose is to recommend an alternative to IP authentication based on a federated identity approach.

Three RA21 pilot projects were established; one corporate pilot, working specifically with representatives from pharma companies who are members of the Pharma Documentation Ring (P-D-R), and two academic pilots. Separate groups were established to further develop the User Experience (UX) work started under the corporate pilot; and also to evaluate the privacy and security issues of the technical architectures of the two academic pilots.

The corporate pilot was started in early 2017, which included a survey of P-D-R companies and confirmed the readiness of P-D-R companies for a federated identity management system.

The three key goals of the corporate pilot were:

- Improved user login experience at the publisher sites
- Provision for granular usage statistics reporting
- Ability to easily set up and maintain Single Sign On with multiple publishers.

During the first phase of the pilot in 2017, the initial User Experience (UX) development was tested with end users at the P-D-R companies. Key findings from the phase one (2017) testing included:

- Equal support for the use of institutional name and personal email address for identification at the publisher site.
- Privacy concerns raised around use of email address.
- Confusion identified around variety of names for an institution.
- Individual user registration seen as being more valuable for frequent users but could be a privacy issue for some.

The SAML protocol was tested with two publishers (Elsevier; Springer Nature) and additional attributes identified that would be required for department billing, differentiating between employee types and for granular usage reporting.

The second phase of the corporate pilot (January to June 2018) focused on further UX testing by the P-D-R pilot participant companies, the specification for granular usage reporting and the exploration of options for the set up and maintenance of a P-D-R-specific federation.

UX testing with a live prototype was done during May 2018 and the results of this have fed into ongoing UX development that will further streamline the Identity Provider (IdP) discovery process. While this UX testing was useful this was not sufficiently seamless to

meet the P-D-R companies' requirements. Further development to streamline the UX is planned for the latter part of 2018.

Specific SAML attributes as previously identified to meet the needs of the P-D-R community have been specified. In the case of the usage reporting attribute, a new SAML attribute has been specified and has been shared with the REFEDS community for consideration.

Successful testing was done by GSK with the OpenAthens access management federation, Elsevier, SpringerNature and Wiley; and discussions were held with a number of federation operators who in principle would be willing to work with P-D-R. An access management federation would enable easy setup and maintenance of Single Sign On (SSO) facilities with multiple publishers.

The corporate pilot has now concluded but work will continue under the overall umbrella of RA21 and this will involve P-D-R pilot participants in further UX testing. A recording of the end of project webinar on 23rd July 2018 can be found here.

Final best practice recommendations from RA21 are expected towards the end of 2018. Following that, the standard NISO process for recommended practices will be followed.

Section 1: Introduction

Publishers, researchers and libraries have relied on IP addresses to authorize content access for many years; but in today's distributed environment, more effective solutions are needed to facilitate a seamless, intuitive and consistent user experience. Single sign-on can significantly reduce the administrative burden on institutions and remove barriers to resource access.

For the consumers of scholarly resources, easy access is critical regardless of workflow, device or location. The underlying assumptions that led to the implementation of IP access are no longer valid; devices are not tied to one location and the user does not typically start their research at the company or institutional portal.

Since 2016 publishers, libraries and other interested parties have been working together, Resource Access for the 21st Century (RA21), towards improved user access to subscribed content across a range of content platforms. RA21 is a joint initiative of STM and NISO.

Under the umbrella of RA21, three pilot programs were undertaken—one focused on the corporate environment and two on the academic environment, P3W and WAYF Cloud. These led to the establishment of cross-pilot groups working specifically in the areas of UX; and user security and privacy concerns.

This is a report from the RA21 corporate pilot program, providing some background information, detailing goals of the pilot, activities, and initial results. This report also indicates some next steps in the RA21 project; and highlights some next steps specifically related to the P-D-R community.

Section 2: Stakeholders

2.1 RA21 (Corporate) Steering Committee Members

(This committee was formerly known as the URA Task Force and then the URA steering committee for the corporate (P-D-R) stream of RA21)

The following individuals, representing their organizations, served on the RA21 (Corporate) Steering Committee:

Jenny Walker
Independent Consultant

Tracey Armstrong
Copyright Clearance Center (CCC)

Elias Balafoutis
Atypon (Until December 2017)

Laird Barrett
Springer Nature (May 2017 – December 2017)

Andrew Clark
UCB (Until October 2016)

Meltem Dincer
John Wiley and Sons (Wiley)

Andy Halliday
Springer Nature (until April 2017)

Matt Kleiderman
Copyright Clearance Center (CCC)

Richard Northover
Elsevier (2017)

Helen Malone
Glaxo Smith Kline (GSK) and P-D-R Pharma Documentation Ring

Babis Marmanis
Copyright Clearance Center (CCC)

Audrey McCulloch
Association of Learned and Professional Society Publishers (ALPSP) (Until December 2016)

Chris Shillum
Elsevier (Until December 2017)

Eefke Smit
International Association of STM Publishers (STM)
(Until December 2016)

Lauren Tulloch
Copyright Clearance Center (CCC)

Rich Wenger
Massachusetts Institute of Technology (MIT) (until December 2016)

Ralph Youngen
American Chemical Society (ACS)

2.2 P-D-R Pilot Participants

The following P-D-R companies participated in the RA21 Corporate Pilot during 2017:

- AbbVie
- BASF
- Glaxo Smith Kline (GSK)
- Novartis
- Roche

The committee would like to thank the many individuals in these companies that contributed to the work of the pilot project.

2.3 Publisher Pilot Participants

- ACS
- Elsevier
- Springer Nature
- Wiley

The committee would like to thank the many individuals in these companies that contributed to the work of the pilot project.

2.4 Acknowledgements

The RA21 (Corporate) Steering Committee would like to offer a special thank you to the following organizations and individuals for their assistance:

- CCC, GSK, ACS and Elsevier for their generous financial assistance in support of the RA21 corporate pilot and their commitment to this project.
- CCC for their support in providing tools for the work of this committee; and in particular Marianne Bright at CCC for her unwavering administrative support and Matt Kleiderman for his technical advice, his editorial assistance, and webex expertise.
- Ralph Youngen, ACS, and Chris Shillum, Elsevier, for their work in adapting the EDUCAUSE survey on Identity and Access Management.
- Richard Northover and Inge Schoutsen of Elsevier for their input into—and preparation of--the User Login wireframes used in the initial UX testing.
- Bill Hess and Ralph Youngen of ACS for their creation of the script for undertaking the initial UX survey.
- Serena Rosenhan and Anna Rouben, both of ProQuest, for their roles in leading the RA21 UX cross-pilot group, and in testing the live prototype with P-D-R company users.

- Heather Flanagan and Julia Wallace, key principals in RA21, for their unflagging support for the corporate pilot; and in particular to Heather for her work on defining, and lobbying the Research and Education Federations Group (REFEDS https://refeds.org/) community for, a SAML attribute to facilitate more granular usage reporting.
- JISC and OpenAthens for their help and advice for GSK in relation to federations.

Section 3: Background

In June 2015, the Pharma Documentation Ring (P-D-R)[1], held a special meeting "Authentication Technologies – Challenges and Opportunities for the Scientific Corporate Information Centre". This meeting was attended by ~50 delegates from P-D-R Member companies, authentication vendors and publishers. In a direct response to the P-D-R who called for follow-up discussions with academic community leaders, publishers and technology vendors, the Copyright Clearance Center (CCC) co-ordinated and sponsored an event in Amsterdam on 8 June 2016. "*The Universal Resource Access Forum: Connecting Researchers to Scholarly Content,* included more than 40 delegates from three key stakeholder groups, and resulted in a commitment from many attendees to take specific actions toward identifying industry-wide solutions."[2]

One of the stated objectives of the Amsterdam forum in June 2016 was to define a way forward that would result in actionable next steps. This was satisfied by a nearly unanimous call to embark on one or more pilot projects that would explore an alternative solution to IP authentication. Since their June 2015 meeting the P-D-R companies had developed their requirements for a solution (See Appendix A).

In brief, one of the key P-D-R requirements is for the user experience to be seamless, intuitive, and consistent; corporate users at P-D-R firms expect easy access to resources regardless of workflow, device or location. Further, the systems must be secure, compliant and enable granular access. From a practical perspective, there was a desire for a single authentication solution supported by all STM publishers.

In July 2016, a cross-sector Task Force, URA, was formed comprising representatives from P-D-R companies, publishers, CCC and representatives of industry groups such as STM and ALPSP. Rich Wenger of MIT was invited to join URA to represent the academic community and Jenny Walker, an independent consultant, agreed to lead this group. The participating parties acknowledged the importance of finding new and practical solutions to address the current issues inhibiting effective universal access to resources.

The Task Force's initial objective was to define and agree on milestone(s) to be reached by the end of 2016; ideally to create pilot projects to address at least one of the many issues raised during the URA event in Amsterdam.

[1] The P-D-R is an association whose members represent the scientific information departments of the leading international R&D-based pharmaceutical corporations

[2] Luther, Judy. Universal Resource Access: Finding a Solution. A report prepared for the Copyright Clearance Center. 17th August 2016. https://www.informedstrategies.com/wp-content/uploads/2015/10/CCC_Universal_Resource_Access_Finding_a_Solution.pdf

By mid- 2016, a parallel initiative, RA21, was emerging from the STM association. Several STM members had similarly been considering the growing number of issues with IP authorization and the limitations of identity federation as the most logical next step for authorizing digital access to material. A number of STM members were already participating in the URA initiative working with P-D-R companies; and before the end of 2016, URA combined with the STM-led effort becoming the corporate pilot for the RA21 initiative; initially known as the URA steering committee for the corporate (P-D-R) stream of RA21

This combined group had a common goal of improving identity discovery – a key initial step to making federated identity viable for all parties.

After the appointment in early 2017 of Julia Wallace as RA21 Project Director and Heather Flanagan as RA21 Academic Pilot Coordinator, RA21 moved to validate the use cases, mission and goals of the project and sought to engage as many interested parties as possible across all stakeholder groups. Two academic pilots (the Privacy Preserving Persistent WAYF or P3W and WAYF Cloud pilots) were established to focus mostly on best practice for identity discovery. Meanwhile the corporate pilot project proceeded towards its goals, working with five of the 26 P-D-R companies and four major STM publishers.

In December 2017, the corporate pilot was extended to the end of June 2018 with the following objectives:

- Participation in further UX review and testing.
- Creating a specification for granular usage statistics.
- Further exploration of a P-D-R-specific federation.

The corporate pilot concluded 30th June 2018. A recording of the end of project webinar on 23rd July 2018 can be found here.

Section 4: RA21 Pilot Projects

4.1 General

Under the umbrella of RA21, three pilot programs were established—one focused on the business environment and two on the academic environment, P3W and WAYF Cloud. It was important to make sure that proposed best practices are implementable in a real world setting and to ensure that the different concepts meet the guiding principles of RA21[3] and are implementable in a secure and privacy preserving manner.

All three RA21 pilots proposed the use of Federated Identity Management (FID) based on Security Assertion Markup Language (SAML) technology. SAML is an established standard for authenticating users for the purpose of authorization. A key aspect of this approach is that a subscriber organization, such as a P-D-R company, will vouch for the relationship between that organization and a user without identifying the user to the publisher or content service provider. Regardless of whether the user starts their search at a publisher site or in Google Scholar, they are authenticated by their sponsoring organization only when they need to access subscribed content. When users attempt to access a resource, it is their organization that authenticates them; and it does so by vouching for the user's relationship with the sponsoring organization; the user's personal information and credentials remains by default with their organization, thus preserving privacy. Single sign-on, through this use of only one username and password to access resources across different platforms, applications and locations, can significantly reduce the administrative burden on institutions and remove barriers to access.

In June 2018 an in-depth technical evaluation was done on the differing architectures of the two academic pilot projects. Both academic pilots were successful in technical approaches to identity provider persistence, and a lot was learned from both projects. Both projects require the establishment and operation of some central infrastructure to support the proposed solutions. The P3W architecture, however, minimizes the value of data that is held centrally, thus minimizing the potential for security or privacy breaches. Further, the P3W architecture has a lighter technical footprint for the central infrastructure and this is more attractive to potential operations partners. RA21 has, therefore, chosen to proceed with the technical architecture prototyped by the P3W pilot. The WAYF Cloud project has now been closed.

The full evaluation and security review are available via the links below, and serve as the first formal outputs of RA21.

- WAYF Cloud and P3W Security & Privacy Recommendations - July 2018
- RA21 Academic Pilot Technical Evaluation - July 2018

[3] https://ra21.org/index.php/what-is-ra21/

The ongoing RA21 work will focus on how to further support more complex levels of integration of P3W technology into a service provider's site and to determine governance for a central P3W service model.

4.2 The Corporate Pilot

Four major STM publishers and five P-D-R companies participated in the corporate pilot. The publishers are ACS, Elsevier, Springer Nature and Wiley. The P-D-R companies are AbbVie, BASF, GSK, Novartis and Roche.

The focus of the corporate pilot project has been on access to publisher resources by authorised users using desktop or mobile devices *outside* the corporate network. The specific goals of the corporate pilot were as follows:

- To improve the current user experience at the publisher's site for redirecting the user back to their company's identity server for authentication. The user experience should be consistent across multiple STM publishers' sites; regardless of the user's location and device used.
- To explore the use of SAML attributes to meet usage statistics needs.
- Ability to easily set up and maintain Single Sign On with multiple publishers. (This 3rd goal was considered of secondary importance).

The pilot project undertook to consider both the privacy (eg GDPR 2018 requirements due for enforcement by 25th May 2018) and compliance aspects of a new authentication solution.

Three work streams were defined to proceed in parallel:

- Baseline SAML integration between the P-D-R companies and the publishers.
- Interface design for user logins at publisher sites.
- Defining attributes for usage statistics.

4.2.1 Corporate Pilot Activities and Results

Activities undertaken and outputs to date include the following:

4.2.1.1 Initial survey of preparedness

An initial survey was prepared and distributed in October 2016 to all P-D-R companies in order to understand the readiness of P-D-R companies for a federated identity management system as was being proposed by URA. The questionnaire on

Identity and Access Management was adapted from a survey created by EDUCAUSE Center for Applied Research (https://library.educause.edu/~/media/files/library/2010/3/esi10b-pdf.pdf) and was reused with the kind permission of Eden Dahlstrom, Chief Research Officer. Results from the survey were shared with the participating companies and anonymised, aggregated results were shared with EDUCAUSE.

Five companies responded to the survey; and the results were determined to be very helpful in validating the way forward. The results were encouraging with respect to the maturity of Identity and Access Management (IAM) implementations at the P-D-R participating companies and their suitability for the solutions being proposed by RA21 as alternatives to IP Authentication. All those who responded to the survey seemed to understand the value of IAM for regulatory and other purposes.

Some specific observations noted include:

- Many, if not all, respondents acknowledged challenges with vendor software.
- Most respondents used a common Single Sign-on (SSO) platform.
- Most respondents allow organizational users to access external resources through SAML and OAuth.

4.2.1.2 Testing SAML Integration

While the initial survey of preparedness indicated the adoption of federated management and the use of SAML by many of the P-D-R companies, this had largely not been used by these companies for access to information resources such as scholarly articles. A baseline SAML integration test was proposed in which each of the P-D-R pilot participants successfully tested SAML integration with Elsevier and Springer Nature. For this test, agreement was reached by the pilot participants on the use of the following attributes to be exchanged:

- First name
- Last name
- Unique Id
- Email address

At least one of the P-D-R pilot participants indicated that the sharing of these attributes with publishers may be of concern with regard to the GDPR requirements.

Two potential additional SAML attributes were identified that may be required by some P-D-R companies:

- To identify internal company users from contracted third-party users in order to manage differentiated access rights.
- To identify a user's department affiliation for purposes of departmental billing.

There are many attribute schema in the world today. One of the most ubiquitous (completely sector neutral) is called inetOrgPerson. This schema is defined in RFC 2798, which has been around since April 2000. This schema is supported "out of the box" by many directory servers eg Active Directory and OpenLDAP. inetOrgPerson supports two attributes that would address the needs identified above:

- departmentNumber
  This identifies a department within an organization and specifies the code for the department to which a person belongs. This can be strictly numeric (e.g., 1234) or alphanumeric (e.g., ABC/123).
- employeeType
  This is used to identify the employer to employee relationship. Typical values used will be "Contractor", "Employee", "Intern", "Temp", "External", and "Unknown" but any value may be used.

The specific values for these attributes could be agreed at a federation level. Pending the formation of a federation (see section 4.2.1.5), P-D-R customers would need to negotiate with each publisher for a specific value to be configured.

Note that although it was not identified at the time, a further attribute will be needed in order for publishers to generate granular usage reports. See also section 4.1.2.4

4.2.1.3 Improving the User Login Experience (UX)

A key goal of the corporate pilot was to improve the current user experience at the publisher's site for redirecting the user back to their company's identity server for authentication.

The corporate pilot participating publishers worked together—also with some of the P-D-R members—to define a new user login experience at the publisher sites. Agreement among publishers was important to the P-D-R participants to ensure a consistent user experience across multiple publishers' sites.

Based on discussions within the Steering Committee and with the P-D-R pilot participants, wireframes for the proposed new publisher login experience were developed. The results provided important input for the RA21 cross-pilot UX group that continued the development of the UX.

To date there have been two iterations of UX testing:

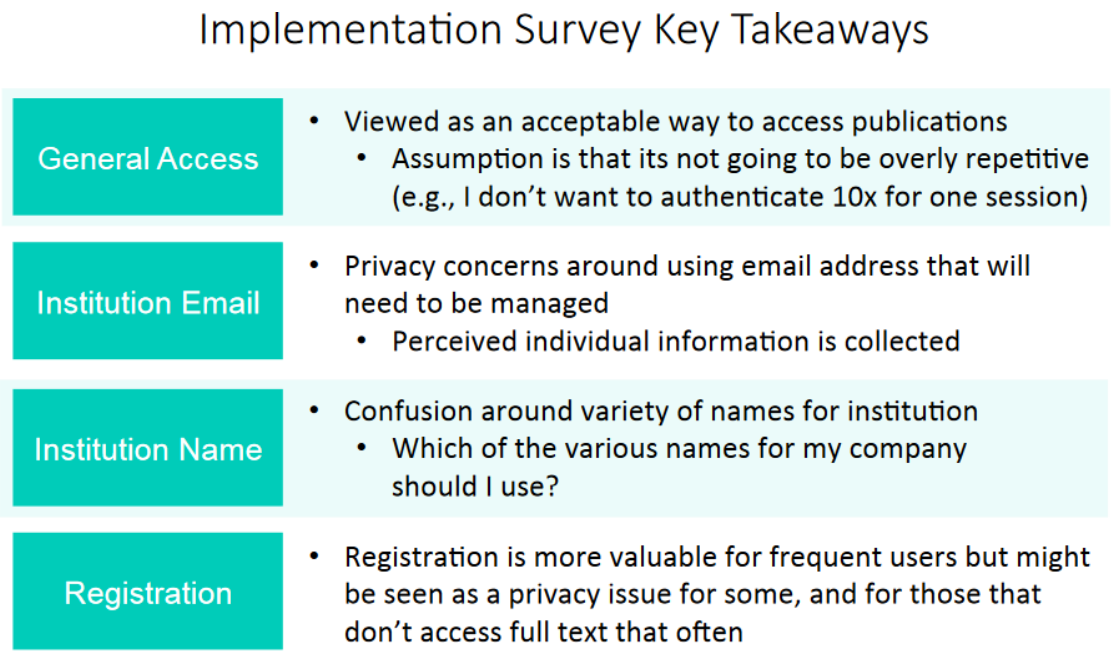
4.2.1.3.1 Wireframes testing with corporate pilot participants only

Wireframes showing a proposed new UX were incorporated into a survey script which was distributed to all five P-D-R pilot participant companies to be completed by end users. Staff could also respond and provide their feedback though these results were not included in the overall analysis.

The type of questions the survey was intended to answer are as follows:

- Do users prefer entering email or company name to identify their parent organization?
- Does the terminology used (e.g., “Check Access”) make sense on the publisher’s site?
- Would users like a registration option that would shortcut further downloads?
- Do users have questions or concerns about the flow presented?
- From a demographics standpoint: how often do users download full text articles?

The survey findings, together with the wireframe screen designs and a user login example from ACS, were presented at a workshop in Amsterdam on 1st September 2017. The findings were as follows:



There was in general equal support for institutional name and email address for identification at the publisher site. Users highlighted issues with both these options as indicated above. However, it was noted that a user's email , if used for initial identification, would not be stored by the publisher but rather used to identify the user's home institution and thereby their Identity Provider service where they would need to identify themselves with their institutional userid and password.

Concerns were raised by the survey respondents regarding the potential need to repeatedly enter their email or institution name at the publisher sites. The P-D-R pilot participants were informed that the focus of the RA21 academic pilots is to explore the persistence of user logins, thereby eliminating the need for users to repeatedly login.

4.2.1.3.2 Second iteration of UX testing (first iteration of UX cross-sector group) – all pilots included.

While the initial UX effort was led by the corporate pilot, by mid-2017 it was recognised that this work needed to span all of the pilots. Responsibility for the ongoing UX work was handed over to a UX cross-pilot work stream to lead through the next phase. This round of UX was tested live against the P3W platform; and using each user's corporate logins for access.

The study objectives were to evaluate the overall experience, organizational access cue, discovery/search page, remembered organization.

The research questions included:

- Can the user successfully access full text from different publisher pages through the call to action?
- Does the subsequent authentication experience match expectations set by the call to action?
- Are users likely to recognize the pattern over time?
- How do users search for their organization?
- Is remembering a previously selected organization effective?

Methodology:

- 30 minute moderated and unmoderated usability sessions where users were asked to complete tasks with the prototype and answer questions about their experience.
- Of the five P-D-R pilot participant companies, Novartis and AbbVie users were involved in this round of testing.

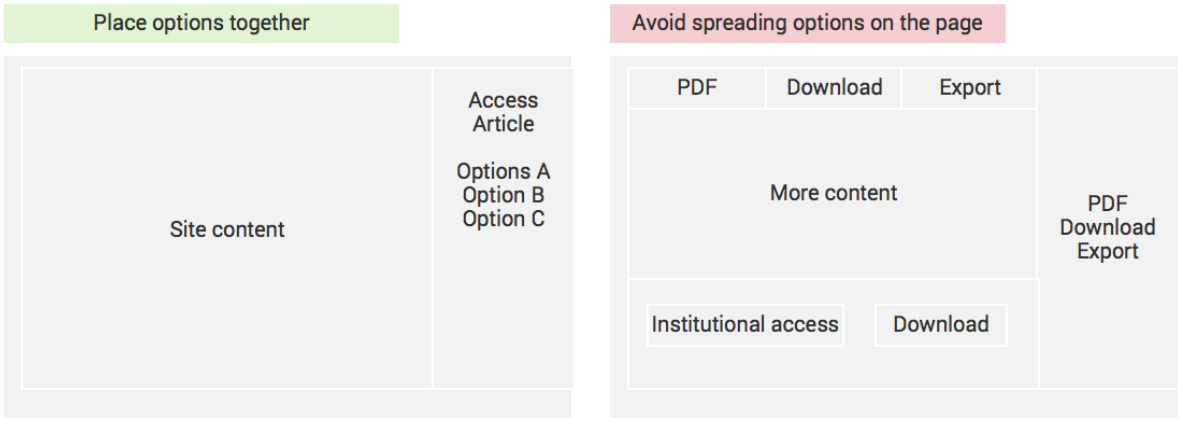
Findings and Recommendations:

Corporate users are very used to their current workflows. They are preconditioned as to what to expect and how to access full text. In some cases, when researchers see an article preview page instead if the full text, they assume they don't have access. The success of the proposed solutions will likely depend on additional internal communication steps to introduce the new solution to the researchers.

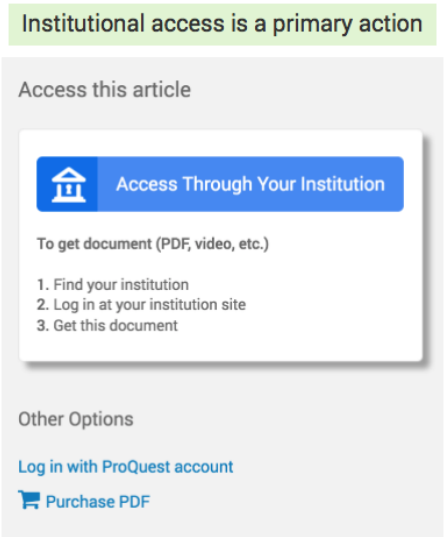
In addition, corporate users would benefit from a stronger indication that they have access through their organization. For example, instead of seeing a button with the label "Access through your institution", the button text would read "Access through Company ABC". Both of these steps will help establish the trust needed for the researchers to follow the new workflow.

Initial recommendations for article preview pages on the publisher sites:

- <u>Present document access options together</u>
  If the document preview page contains several options to access articles, present them together so that the user can quickly see all of the available options without scanning/scrolling.



- <u>Present article access options in hierarchical order</u>
  Some options (probably one) gives user a better experience of getting an article. For the majority, the ability to get an article for free is a better experience than going through the option to purchase an article. Present the option to get an article for free with more prominence. For example, if most of your users have institutional access, present institutional access as the primary call to action. Place other options nearby.



- <u>Use language that clearly differentiates the options</u>
  If the site has both free options and paid options, clearly distinguish them. If,

for example, PDF is not free, accompany it with text that makes it clear (e.g. Purchase PDF, Purchase and Export).

If users see PDF and Institutional Access, they are more likely to follow the PDF option even if it leads to a purchase. In the latter case it would mean backing out and looking for another option and starting over; and thereby leading to a longer path and a worse experience.

- Consistent look
  Users expect that things that look the same will work the same. Seeing the same call to action across publisher sites will reduce cognitive load and minimize friction, allowing users to click through authentication steps quickly and confidently, without losing focus on their research task.

The RA21 UX team is actively working on updating the live prototype to demonstrate a Level 2 experience whereby "Access through <companyName>" call to action is automatically populated using local browser storage. The UX design team is working on visual refinements (colors, fonts, etc.) to the prototype and collecting feedback from a wider set of publishers than those who participated in the initial round of testing the live prototype.

Another round of user testing on the next major iteration of the live prototype is targeted for September/October 2018. All P-D-R pilot participants will be invited to participate in this next round of testing.

However, these UX developments resulting from the overall research findings, while useful, were considered to fall short of the expectations of the P-D-R pilot participants. A more "seamless" experience is desired, similar to IP authentication for users currently connecting from within the corporate network. See 4.2.1.3.3 below.

4.2.1.3.3 Exploring options for further streamlining the login process for company-owned devices.

Discussions were held at a UX meeting in London on June 8th 2018 to explore the possibility of further improving the UX for users with company-owned devices.

The idea was that for company employees/researchers reading articles in publications for which the company has a subscription, to ensure that these users have information in their browser local storage that will point the publishers to the company IdP in order to authenticate the users. A few different approaches of distributing that pointer data for different populations of users and devices were suggested. The pointer being distributed would not contain any individually-identifiable data, would be identical for all users in the organization and does not actually grant access to articles, just informs the publisher how to authenticate the users.

- For fully corporate managed devices, using endpoint management software to push that pointer to user's browsers so that it's available the first time they encounter the publisher's paywall, and can be immediately re-directed to the company's authentication page.
- For devices not centrally-managed but allowed to access corporate resources, placing a code snippet on a well-known internal web page (e.g., a library portal home page) that downloads the IdP pointer into a user's browser local storage when they access that page.
- For devices largely detached from the corporate network, allowing users to email/SMS themselves the code snippet to download the pointer.

These options were discussed between Ralph Youngen and GSK information and IT staff. As a result of this discussion, it was proposed that a clickable URL containing the necessary IdP information could be shared with users (see second and third options above). On clicking the URL the local browser storage would be updated with the company's IdP information, thus eliminating the need, even on the first login at a publisher site, for the user to enter their company name or email address. This proposed solution would not be limited to company-owned devices. Clicking on the URL would need to be done only once per user's device.

Further testing of this approach will be required.

4.2.1.4 Granular Usage Statistics

GSK currently generate usage reports at the level of the individual user and would like to maintain this capability with any new authentication system that is introduced. The level of granularity sought is not provided by the industry COUNTER reports and alternative mechanisms are required.

The Steering Committee proposed the designation of a particular SAML attribute that would set the aggregation of reporting logic. This aggregation attribute could be the same as the unique user identifier, in which case the statistics will be generated per user (as required by GSK), or this attribute could match a department or other grouping, in which case the reporting would be at a group level. This could also be left blank. This proposal would make it easier for the publishers to develop a single reporting mechanism that could provide different reporting levels to different customers.

No known existing attribute was considered suitable and a proposal has been made to the REFEDS community by the editorial board of the SCHAC schema to add a new attribute, schacLocalReportingCode, which would be used for opaque codes, which do not affect the users' privacy. The recipient, in this case the publisher, would use the codes to generate the usage reports per such code.

The schacLocalReportingCode attribute would allow for multiple values because billing (and hence reporting) may be spread across different departments or cost centers in an organization.

Once approval has been given by the REFEDS community for the confirmation of this attribute, the SCHAC schema will be updated and made available to directory servers for use.

Heather Flanagan, RA21 Academic pilot coordinator, is responsible in her capacity as the SCHAC schema editor for the updating of the attribute. Heather will also produce a draft specification for the granular usage reports for review by the P-D-R and publishing members (past and present) of the RA21 corporate pilot.

4.2.1.5 Initial exploration of a P-D-R-specific federation.

The exploration of a P-D-R federation was a secondary goal of the corporate pilot but support for this is growing. While it is feasible for the P-D-R companies to work with each publisher directly, rather than via a federation, this becomes extremely onerous in cases where there are multiple, in some cases 100+, publisher subscriptions. It is also a large overhead for the publishers to deal directly with each company.

An access management federation would therefore make the proposed RA21 ecosystem work well without requiring individual communication and setup with each publisher. With a federation, metadata exchange would all be done with one entity (i.e., the federation). This would apply to company metadata and publisher metadata.

Discussions were held with a number of federation operators around the world, including those targeted predominantly at the academic space, those operating solely in the corporate world as well as those that straddle both environments.

Testing was done by GSK with two federations. Of these, one was very quick and easy to setup and access was gained to Wiley, Elsevier and SpringerNature sites. The other test proved more problematic due to an apparent incompatibility between the specific SSO service already in use and the particular federation under test

Some P-D-R companies are already members of a federation, though the information departments in these companies have largely not used these facilities. The extent of existing membership of a federation by P-D-R member companies may be an important factor in assessing a suitable federation operator of a P-D-R federation. Other factors to consider include:

- Global operations with 24X7 support available.
- The number of relevant publishers that are already members of the federation.
- The cost. Note that the P-D-R requirements stipulated a “cost-effective” service for both customers, such as their member companies, and also the

information providers. Typically the more organizations that join a federation, the cheaper federation membership tends to become.

- The level of security and compliance procedures. Many of the national academic consortia rely on self-assertion by members of compliance to the federation rules. At least one federation we spoke with that largely serves the corporate community stipulated a rigorous compliance process required by a third-party for members joining the federation. This would impact both the publishers and the P-D-R companies.

Further discussion will be needed by the P-D-R community on the issue of a P-D-R federation.

4.2.1.6 Outreach

Helen Malone and Jenny Walker participated in the RA21 Outreach group. Also, throughout the pilot period, a number of Steering Committee members spoke on RA21 and the work of the corporate pilot at industry events. A full list of events in which RA21 has participated is available on the RA21 website.

In November 2017 CCC hosted a webinar aimed at corporate information professionals to inform them of the RA21 objectives and activities to date. Well over 100 people registered for this event. A recording of the event can be found here.

In September 2018, Tracey Armstrong, CCC, Ralph Youngen, ACS, and Helen Malone, GSK, will present at the P-D-R 60th anniversary meeting in Vienna. They will present on the RA21 project in general and specifically on the outcomes of the corporate pilot.

4.2.1.7 Privacy and Security

Note that privacy and security issues were addressed by a dedicated RA21 working group, which included a number of CIOs from participating organizations. The evaluation included an in-depth security analysis, a privacy review, and a detailed technical architecture comparison. A detailed report describing the full results of this evaluation is available here.

In summary, it was determined that both the WAYF Cloud and P3W pilots are very low risk. The P3W architecture stores information about the user's choice of IdP in the user's browser while the WAYF Cloud architecture relied on the collection of the user's IdP choice in a central database. With no central collection of this information, the P3W architecture adheres to the privacy principle of data minimisation and this became the preferred option. See also further information in section 4.1.

The findings from the RA21 Corporate Pilot will be rolled into a final, RA21 NISO standardized set of recommendations.

4.2.2 Next Steps

Next steps for RA21 in general:

- The ongoing RA21 work will focus on how to support more complex levels of integration of P3W technology into a service provider's site; thus further improving the UX.
- Determine governance for a central P3W service model.

Next steps specific to the corporate pilot:

- Participate in and provide feedback on the next (level 2) iteration of the live UX prototype. This is targeted for September/October 2018.
- Test options for populating, in advance, the local browser storage with the company's IdP information and thus eliminating the need, even on the first login at a publisher site, for the user to enter their company name or email address.
- Discuss options for a P-D-R-specific federation.
- Review granular usage reporting specification once released and provide feedback on this. Once agreed with the publishers, set up a test.

Appendix A: P-D-R Requirements

User Experience Requirements:

1. A user access experience that is seamless, intuitive, simple, efficient, and effective.
2. Consistent inside and outside the network and across multiple publisher/vendor platforms.
3. Easily accessible from any device.
4. Supports multiple workflows.

Technical Requirements:

1. Supports single sign on access inside and outside the corporate network.
2. Supports open standards (eg SAML) supported by all parties.
3. Copyright and licence compliant access – multiple permission sets.
4. Comprehensive, complete and granular usage reporting – leverage existing corporate HR directories.
5. Flexible and user-friendly administration interface, allowing use of this authentication solution for all types of information resources, license models, and company structures.

Vendor Strategy:

1. Single Authentication Solution across all providers.
2. All providers (including smaller publishers and information vendors) need to be able to support single sign on and not be IP dependent.
3. Simple content licensing models that enable consistent permissions for internal staff and external partners/collaborators.
4. Vendors need to embrace advancement in authentication and make it a priority in their development roadmap.
5. Cost-effective solution (for both customers and information vendors).

Appendix B:

Identity and Access Management Questionnaire for P-D-R (with results shown)

Key:

1=P-D-R company A

2= P-D-R company B

3= P-D-R company C

4= P-D-R company D

5= P-D-R company E

Section 1: About you and your organization

1.4 What is your position?
( ) CIO (or equivalent)
( ) Vice president or equivalent (non-CIO)
(1,4) Director of IT
( ) Chief information security officer
( ) Manager of IT networking
(2,3,5) Other IT management
( ) Other administrative management

1.5 At my organization, IT is:
( ) Highly centralized
(1,2,4) Centralized
(3,5) Balanced
( ) Decentralized
( ) Highly decentralized

1.6 What BEST characterizes your organization in terms of adopting new technologies?
(2) Early adopter
(1,4) Mainstream adopter
(3,5) Late adopter

1.7 I am personally very involved in Identity and Access Management decisions at my organization.
( ) Strongly disagree
( ) Disagree
( ) Neutral
(3,5) Agree

(1,4) Strongly agree

Section 2: Organizational Perspectives on Identity and Access Management (IAM)

2.1 What is your opinion about the following statements?
As used here, IAM means "the business processes and technological capabilities required to support the use of centralised digital identities both within your organisations and when interacting with third parties in order to identity and authenticate authorised users "

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
a. My organization's senior management understands the benefits of investing in Identity and Access Management				3	1,2,4,5	
b. My organization's senior management understands the costs of Identity and Access Management			3		1,2,4,5	
c. My organization is providing the resources needed for Identity and Access Management			3	2	1,4,5	
d. My organization has the identity and access management infrastructure needed to effectively manage access to internal organizational resources.				2,3	1,4,5	
e. My organization has the identity and access management infrastructure needed to effectively manage access to extra-organizational resources.		2,3		5	1,4	
f. My organization has the identity policies needed to effectively manage access to internal organizational resources.				1,2,4	3,5	
g. My organization has the identity policies needed to effectively manage access to extra-organizational resources.				1,2,4,5	3	

2.2 What is the status of the following activities?

	Completed	In progress	Planning to do	Not planning to do	Don't know
a. Documented business case for any area of Identity and Access Management	1,2,3,4,5				
b. Documented plan for Identity and Access Management	1,3,4,5	2			
c. Released an RFI or RFP for Identity and Access Management	1,3,4,5	2			
d. Risk assessment of data access security and privacy practices	1,4,5	2,3			

2.3 Have you implemented, or are you currently implementing, any of these online self-service functions? *Check all that apply.*

[1,2,3,4,5] a. Updating personal information
[ ] b. Setting privacy preferences for release of identity information

2.4 Do you have documented policies for establishing identity (e.g., how user IDs are issued)? *Required.*
( ) No documented policies
( ) Policies are in progress or partially completed
(1,2,3,4,5) Policies are completed
( ) Don't know

2.5 Do you have documented policies for user authentication (e.g., guidelines, responsibilities for passwords)? *Required.*
( ) No documented policies
( ) Policies are in progress or partially completed
(1,2,3,4,5) Policies are completed
( ) Don't know

2.6 Do you have documented policies for user authorization (e.g., what groups are allowed what access)? *Required.*
( ) No documented policies
( ) Policies are in progress or partially completed
(1,2,3,4,5) Policies are completed
( ) Don't know

2.7 What are the primary motivators at your organization for pursuing Identity and Access Management? *Select up to three.*
[1,2,3,4,5] a. Regulatory compliance (e.g., HIPAA, GLB Act, FERPA)
[1,2,3,4,5] b. Security/privacy best practices
[1,4] c. Enhanced user services and satisfaction
[3] d. Cost reduction/increased efficiencies
[ ] e. Strategic value/opportunities
[ ] f. Improvements in our technical environment
[ ] g. Strategy of early adoption/experimentation
[ ] h. Keep current with generally accepted IT directions
[2] i. Position the organization for implementation of federated identity
[5] j. Reduce vendor dependencies
[ ] k. Other

2.8 What are the primary challenges to your organization in pursuing Identity and Access Management? *Select up to three*.
[ ] a. Lack of acceptable ROI
[ ] b. Adequate funding is not available
[2] c. Higher IT priorities
[ ] d. Lack of IT staff expertise
[ ] e. Lack of organizational senior management's support
[2] f. Technical solutions are too immature
[1,2,3,4] g. Problems with vendor software and support
[ ] h. Problems with our organization's technologies/infrastructure
[3] i. Data integrity problems (consistency, accuracy, etc.)
[2] j. Difficulty developing campus policies and procedures
[1,3,4] k. Difficulty implementing campus policies and procedures
[ ] l. Lack of ownership of Identity and Access Management by a central group

[ ] m. Other

2.9 Which BEST describes your organization's current thinking about Identity and Access Management solutions?
( ) We probably will not use vendor solutions, but will build solutions using in-house developed or open-source software.
(3) We will address our short term needs with best-of-breed vendor point solutions and integrate these various products in-house.
(1,2,4) We will first identify our long-term business and architecture strategy and then decide on a solution or set of solutions for the organization.
( ) We will probably buy the vendor suite solution that best aligns with our network, infrastructure, and hardware vendors.
(5) We will probably buy the vendor suite solution that best aligns with our administrative applications and ERP vendors.
( ) Other
( ) Don't know

Section 3: Identity and Access Management Importance and Capability

This section presents 13 benefits related to Identity and Access Management for your evaluation. Please rate each item's importance to your organization and your organization's current capability to deliver it.

3.1 Capability to immediately enable all authorized services for a new user

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			2	5	1,3,4	
b. Rate your organization's current capability			3	1,2,4,5		

3.2 Capability to immediately change authorized services for a user who changes roles

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			5	2,3	1,4	
b. Rate your organization's current capability	3	1,2,4,5				

3.3 Capability to immediately disable all services and user IDs when a user is no longer affiliated with the organization

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization					1,2,3,4,5	
b. Rate your organization's current capability			3		1,2,4,5	

3.4 Capability to give visitors/guests only the specific access they require and disable that access at the correct time

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization		2,5			1,3,4	
b. Rate your organization's current capability		2,5	3	1,4		

3.5 Prior to issuing credentials (e.g., user account, ID card, etc.), have the appropriate level of confidence (based on type of constituent) that a user is who he or she claims to be

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			2	3,5	1,4	
b. Rate your organization's current capability			2,3	1,4,5		

3.6 Capability to directly track illegal or unauthorized network activity back to the person responsible

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			5	2	1,3,4	
b. Rate your organization's current capability			1,3,4,5	2		

3.7 Single sign-on

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization				5	1,2,3,4	
b. Rate your organization's current capability			3	1,4,5	2	

3.8 Capability to provide self-service functions (e.g., password reset, profile management)

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			5	3	1,2,4	
b. Rate your organization's current capability			1,3,4,5		2	

3.9 Capability of strong authentication (e.g., strong passwords, two-factor authentication)

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization				5	1,2,3,4	
b. Rate your organization's current capability				1,3,4,5	2	

3.10 Have a single authoritative source of information for all persons affiliated with the organization (as an organizational asset)

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			3	2,5	1,4	
b. Rate your organization's current capability		2,3		1,4,5		

3.11 User authentication and authorization processes that are scalable

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			3	5	1,2,4	
b. Rate your organization's current capability			1,4	2,3,5		

3.12 Capability to allow organizational users to access external resources that require their own authentication and authorization (e.g., licensed information services, cloud services, etc.)

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			2,5	3		
b. Rate your organization's current capability		2	3	5		

3.13 Capability to allow non-organizational users access to our organizational resources for which we require authentication and authorization

	Very low	Low	Medium	High	Very high	Don't know
a. Rate the importance to your organization			2	5	1,3,4	
b. Rate your organization's current capability			3,5	1,2,4		

Section 4: Establishing Identity and User Authentication

4.1 Which of the following user authentication methods does your organization use when providing access to network services?

	Using	Planning to use	Not planning to use	Don't know
a. Conventional password/PIN	1,2,3,4,5			
b. Strong password	1,2,3,4,5			
c. Kerberos	1,2,3,4,5			
d. PKI certificate (software)	1,2,4,5	3		
e. PKI hardware token	2,5	1,4	3	
f. Onetime password/token	1,3,4,5			
g. Other multi-factor authentication methods	1,2,3,4,5			
h. Biometric identification		1,2,4,5	3	

4.2 Do any organizational web resources at your organization support sign-in using web-based user-centric identifiers (e.g., OpenID, Facebook Connect, Google Account, Windows Live ID)?
(1,4,5) No
(2,3) Yes
( ) Don't know

4.3 **To what extent is your organization considering or implementing an enterprise directory? By *enterprise directory*, we mean an organizational directory service that has the capability to include all persons affiliated with the organization and to be used by multiple applications.**
( ) Not considering
( ) Currently evaluating
( ) Planned, but won't start within the next 12 months
( ) Plan to start within the next 12 months
( ) Implementation is in progress
(3) Partially operational
(1,2,4,5) Fully operational

Section 5: Single Sign-On

5.1 To what extent is your organization considering or implementing single sign-on? *Required.*

( ) Not considering	*Go to 5.3*
( ) Currently evaluating	*Go to 5.2, then to 6.1*
( ) Planned, but won't start within the next 12 months	*Go to 5.2, then to 6.1* ( )
Will start within the next 12 months	*Go to 5.2, then to 6.1*
(3) Implementation is in progress	*Go to 5.2, then to 6.1*
( ) Partially operational	*Go to 5.2, then to 6.1*
(1,2,4,5) Fully operational	*Go to 5.2, then to 6.1*

5.2 What is, or will be, your approach to implementing single sign-on? *Check all that apply.*
[ ] a. Not yet determined
[ ] b. Use open-source software (e.g., Kerberos, CAS, PubCookie)
[ ] c. Use homegrown software developed at your organization or another organization
[1,2,3,4,5] d. Use commercial vendor software (e.g., Ping, Okta, Microsoft)
[ ] e. Other

5.3 What are the primary reasons your organization is not considering single sign-on? *Check up to three.*
[ ] a. Capabilities not required at this time
[ ] b. We are not that far along in our Identity and Access Management work
[ ] c. Adequate funding is not available
[ ] d. Technical solutions are too immature
[ ] e. Problems with our organization's technologies/infrastructure
[ ] f. Data integrity problems (consistency, accuracy, etc.)
[ ] g. Difficulty developing campus policies and procedures
[ ] h. Difficulty implementing campus policies and procedures
[ ] i. Other

Section 6: Federated Identity

Definition: *Federated identity permits automated management of identity information between your organization and other organizations, or internally, to facilitate collaborative or business initiatives.*

6.1 Over the next 12 months, demand for Software-as-a-Service applications will increase my organization's need for federated identity services.
( ) Strongly disagree
( ) Disagree
(2) Neutral
(5) Agree
(1,3,4) Strongly agree

6.2 To what extent is your organization considering or implementing a federated identity solution? *Required.*

( ) Not considering — *Go to 6.10*
( ) Currently evaluating — *Go to 6.9*
( ) Planned, but won't start within the next 12 months — *Go to 6.9*
( ) Plan to start within the next 12 months — *Go to 6.9*
( ) Implementation is in progress — *Go to 6.3–6.9*, then 7.1
(3) Partially operational — *Go to 6.3–6.9, then 7.1*
(1,2,4,5) Fully operational — *Go to 6.3–6.9, then 7.1*

6.3 How many external service providers does your organization exchange identity attributes with via a federated identity solution?
( ) 0
(2) 1 to 50
(1,3,4,5) More than 50
( ) Don't know

6.4 Twelve months from now, how many external service providers do you expect your organization to be exchanging identity attributes with via a federated identity solution?
( ) 0
(2) 1 to 50
(1,3,4,5) More than 50
( ) Don't know

6.5 How many internal resources/applications are currently enabled for sign-on via a federated identity solution at your organization?
( ) 0
(3,5) 1 to 50
(1,2,4) More than 50
( ) Don't know

6.6 Twelve months from now, how many internal resources/applications do you expect will be enabled for sign-on via a federated identity solution at your organization?
( ) 0
(5) 1 to 50
(1,2,3,4) More than 50
( ) Don't know

6.7 Overall, which best describes your organization's experience with implementing federated identity?
(1,3,4) Policy/process issues are more challenging than technical issues.
(5) Policy/process issues are about equally as challenging as technical issues.
(2) Technical issues are more challenging than policy/process issues.

6.8 Has your organization implemented, or is it planning to implement, any of the following?*

	Already implemented	Currently implementing	Planning to implement	Not planning to implement	Don't know
Service A				1,2,3,4,5	
Service B				1,2,3,4,5	
Service C	3,5			1,2,4	
Service D	1,4			2,3,5	
Service E		1,4		2,3,5	
Service F	1,2,3,4	5			
Service G				1,2,3,4,5	
Service H				2,3,5	1,4

*The list of federated identity, Single Sign On and SAML based services has been anonymised for the purposes of this report.

6.9 What are the primary motivators for your organization to evaluate or implement federated identity solutions? *Check up to three.*
[1,2,4] a. Single sign-on within the organization
[ ] b. Provide access to internal resources
[1,2,3,4,5] c. Provide access to external administrative applications (e.g., HR, benefits)
[2,3,5] d. Provide access to external service applications (e.g., travel, expenses)
[ ] e. Provide access to external information resources/research tools/data resources
[1,2,3,4,5] f. Provide for external collaboration
[ ] g. Provide access to external training resources
[2] h. Enable access by external users to organizational resources

6.10 What are the primary reasons your organization is not considering a federated identity solution at this time? *Check up to three.*
[ ] a. Capabilities not required at this time
[ ] b. We are not that far along in our Identity and Access Management work
[ ] c. Adequate funding is not available
[ ] d. Technical solutions are too immature
[ ] e. Problems with our organization's technologies/infrastructure
[ ] f. Data integrity problems (consistency, accuracy, etc.)
[ ] g. Difficulty developing policies and procedures
[ ] h. Difficulty implementing policies and procedures
[ ] i. Other

6.11 Do you think your organization will consider implementing a federated identity solution at some point in the future?
( ) No
( ) In the next 12 months
( ) More than 12 months from now
( ) Don't know

Section 7: Conclusion

7.1 May we contact you to obtain further insights or clarifications on your responses? *Required.*
( ) No *Go to 7.3*
(1) Yes *Go to 7.2, then to 7.3*

7.2 What is your e-mail address?___

7.3 If you have any other comments or insights about Identity and Access Management, please share them with us._____