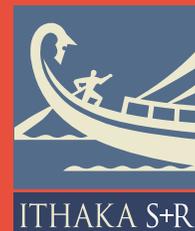


ISSUE BRIEF

Meeting Researchers Where They Start: Streamlining Access to Scholarly Resources

March 26, 2015
Roger C. Schonfeld



LOGIN
PASSWORD

START

PROXY

ENTER
PASSWORD

ERROR
ACCESSING
FILE

404 ERROR

UNABLE TO
CONNECT

LOGIN
PASSWORD

DO NOT ENTER

FONT
TOO SMALL



Ithaka S+R is a strategic consulting and research service provided by ITHAKA, a not-for-profit organization dedicated to helping the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways. Ithaka S+R focuses on the transformation of scholarship and teaching in an online environment, with the goal of identifying the critical issues facing our community and acting as a catalyst for change. JSTOR, a research and learning platform, and Portico, a digital preservation service, are also part of ITHAKA.

Copyright 2015 ITHAKA. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of the license, please see <http://creativecommons.org/licenses/by-nc/4.0/>.

ITHAKA is interested in disseminating this brief as widely as possible. Please contact us with any questions about using the brief: research@ithaka.org.

Instead of the rich and seamless digital library for scholarship that they need, researchers today encounter archipelagos of content bridged by infrastructure that is insufficient and often outdated. These interconnections could afford opportunities to improve discovery and access. But in point of fact, the researcher's discovery-to-access workflow is much more difficult than it should be.¹

...researchers' expectations are being set not by improvements relative to the past but rather by reference to consumer internet services

A different paper might reasonably celebrate the great strides that have been made over the past two decades by libraries and content providers to expand and facilitate access to scholarly resources online. But researchers' expectations are being set not by improvements relative to the past but rather by reference to consumer internet services that enable our use of multiple devices anywhere and effective switching between them. Given my own work conducting research, I am called to the unhappy task in this paper of emphasizing six points of failure.²

The library is not the starting point

A common understanding about discovery practices is a prerequisite for libraries, content providers, and intermediaries to facilitate and improve content access.³

First, the index-based search services, typically powered by EBSCO, Ex Libris, OCLC, or ProQuest and increasingly finding a prominent place on the library's home page, account for a relatively minor share of search-driven discovery. These services may be important for certain user types or practices, but they do not initiate even a substantial minority of content accesses to major content platforms.

¹ I use the term "researcher" inclusively to refer to all academic users, including students, scholars, and others. Some aspects of the workflows I discuss here will more relevant to one type of user or another.

² Although I have been working on the issues in this paper for some time, I offer special thanks to Jill O'Neill of NFAIS for inviting me to speak at the March 2015 virtual seminar Making Content Portable, Making It Usable, which was the catalyst to set down these thoughts in writing. I first experienced a user's presentation of his access difficulties in a presentation by Ashley Crowson at a Sage-organized panel at the 2014 London Book Fair. I thank Alex Humphreys, Kimberly Lutz, Deanna Marcum, Barbara Rockenbach, Nadaleen Tempelman-Kluit, and an anonymous reader for helpful comments on a draft of this paper.

³ I presented a variety of data on this topic in "Does Discovery Still Happen in the Library? Roles and Strategies for a Shifting Reality," *Ithaka S+R Issue Briefs*, September 24, 2014, available at <http://sr.ithaka.org/blog-individual/does-discovery-still-happen-library-roles-and-strategies-shifting-reality>.

Second, Google and Google Scholar are comparatively important discovery starting points. Each provides not only search but also various types of anticipatory discovery, including reading recommendations and keyword, author, and citation alerts. Google Scholar is especially strong for articles and article-like materials, and it has real limitations for certain other content types.

Third, on-platform discovery—i.e. the search engine of a major content platform, its alerting services, and its on-platform referrals—can be a significant starting point.

Fourth, various types of anticipatory discovery services, provided by third parties such as Academia.edu and ResearchGate, appear to be growing in importance.

Finally, scholarly conferences and professional networks are significant sources for the discovery of scholarship, both in process as well as the formally published literature.

Mechanisms for content access succeed only when they conform to Lorcan Dempsey's observation that "discovery happens elsewhere."⁴ Authentication and authorization to licensed e-resources must work effectively without regard to the researcher's starting point.

The campus is not the work location

A second vital assumption is about where research takes place. Most universities authenticate users to licensed content differently on-campus than they do off-campus. While that authentication is often invisible on-campus, it typically requires some kind of login from off-campus. It is thus absolutely essential to recognize that a substantial share of research takes place off-campus, in the sense that the preferred location is outside the campus's IP range. Therefore, researchers can easily find themselves at this type of screen on a licensed e-resource (see Figure 1). This particular screen is from Project Muse, but it is common across most licensed e-resources.

⁴ Lorcan Dempsey, "Discovery Happens Elsewhere," *Lorcan Dempsey's Weblog*, September 16, 2007, <http://orweblog.oclc.org/archives/001430.html>.

Figure 1

This login page is a mystery to most researchers. They can be excused for wondering “what is Shibboleth?” even if their institution is part of a Shibboleth federation that is working with the vendor, which can be determined on a case by case basis by pulling down the “Choose your institution” menu. Many research universities have at least begun to deploy Shibboleth for these purposes, and, in these cases, clicking on one’s institution produces an institutional login page that should unlock access to the item being sought. Unfortunately, in many more cases, even when the technology infrastructure is in place among both parties, the bilateral agreements required between each institution and each content provider have simply not been concluded, leading many higher education institutions in the United States absent from many platforms’ Shibboleth lists.

The proxy is not the answer

Instead, in many cases, researchers must use a proxy to authenticate.⁵ Proxies have been popular because they are comparatively easy for the library to configure and maintain. They rewrite the content provider URL, rather than actually providing the researcher with a portable credential, so the researcher must click through the proxy server before arriving at the licensed content resource. When a researcher arrives at a content platform in another way, as in the example above, it is therefore a dead-end.

It would in theory be possible to add a dropdown menu that takes the researcher through an institutional proxy server in just exactly the same way as Muse’s screen does for Shibboleth. This is less common, perhaps because it would require tracking thousands of different proxy servers across hundreds of different sites.

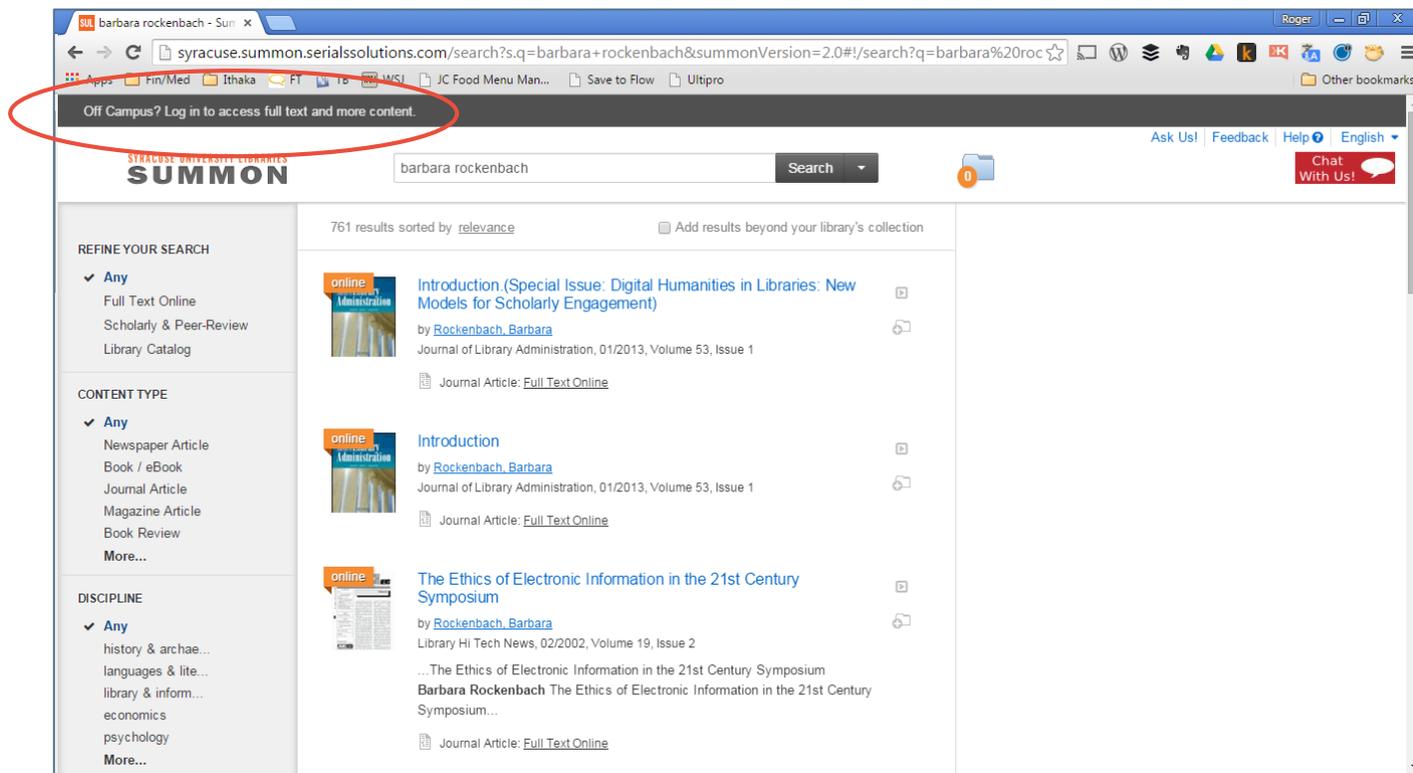
⁵ At some larger universities, it is not uncommon to have multiple mechanisms for authentication, causing further confusion.

The dead-end just described naturally enough leads to frustration. Having encountered this barrier, some off-site researchers might now search Google to see if there is an open access version of the item. Others, seeking authentication, might turn to the library home page, where they will typically find a search box powered by one of the major index-based search services.

There, a researcher might type the article title or author into the search box, and the discovery service, noting that the researcher is not on a recognized campus location, might suggest logging in through the proxy (as indicated in Figure 2, which shows the Summon configuration available through Syracuse University, where the login banner helpfully appears at the top). Although the number of web page clicks from discovery to access is far too high, once logged in through the proxy this approach generally succeeds.

Of course, in other scenarios, it is only after all this effort that the researcher realizes the item is not available through the university library. In this case, the workflow might lead to ILL or an article purchase or rental service.

Figure 2



The index is not current

And this approach also regularly fails for another reason, when the item being sought is new, such as when it is discovered through an alert. For example, a Google Scholar keyword or citation alert indicates newly available content typically days if not weeks before it is indexed by one of the index-based search services. In those cases, the researcher experiences yet another dead end.

The realization that these search indices are not updated immediately for all indexed content has served to limit their potential as more fully-fledged discovery services. They cannot offer alerting or other forms of anticipatory discovery until they are updated more frequently.

The PC is not the device

A declining share of scholarly research is conducted using desktop or laptop PCs and Macs. While there are other factors, the growth of the iOS and Android operating systems, both for phones and tablets, is a leading cause. This shifting dynamic poses some challenges but also suggests new opportunities for serving researchers.

Perhaps the most obvious issue is the PDF. The PDF file format is understandably ubiquitous for scanned serials and books, since it preserves the page image. It is also quite common as the sole format even in digital-first and digital-only publishing. PDFs offer several benefits, not least of which is that they typically contain the scholarly content and little else.⁶

But when a user wishes to read the content of a scholarly publication, PDFs have their limits, especially when working on smaller screens such as those on a mobile phone. The problem with the PDF is that the text does not naturally reflow. Figure 3 shows the impossibly small text size when viewing a complete PDF in portrait mode. Zooming in can work effectively by turning the device into landscape mode, for single-column text, as shown in Figure 4. But PDFs fail entirely on small devices when text is split into multiple columns. As a workaround today, many researchers may avoid reading scholarship on a mobile device and instead mark an item of interest to be read for later but, but there is every reason to believe that these are stop-gap solutions waiting for an

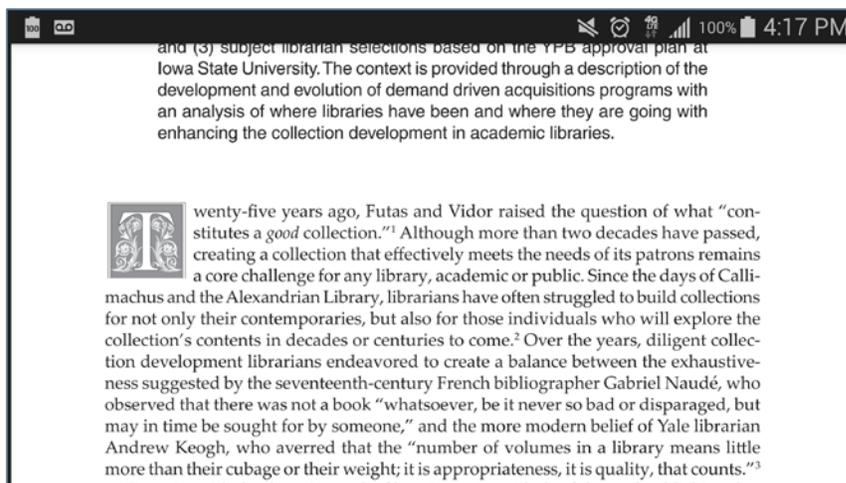
⁶ See for example Justin Kiggins's response on *Quora*, January 10, 2012, <http://www.quora.com/Why-do-scientists-tend-to-prefer-PDF-documents-over-HTML-when-reading-scientific-journals/answer/Justin-Kiggins>.

improved small-screen reading experience.⁷

Figure 3



Figure 4



⁷ The New Media Consortium has reported the shift among libraries towards mobile solutions as "fast trend." See L. Johnson, S. Adams Becker, V. Estrada, and A. Freeman, "Prioritization of Mobile Content and Discovery, *NMC Horizon Report: 2014 Library Edition* (Austin, Texas: The New Media Consortium), 2014, pp. 8-9, <http://cdn.nmc.org/media/2014-nmc-horizon-report-library-EN.pdf>.

Just as PDFs are optimized for large screens, so are many websites. Here are two examples of websites that do not work well on a mobile phone. Figure 5 shows a proxy login screen, for the EZProxy system as configured by Syracuse University, which is workable only because the web browser saves my password so I only have to click the "Authenticate" button and am saved from typing my credentials into the unreadable boxes. Figure 6 shows a reflowable-text article but without responsive design, so that the page is not automatically resized for a smaller screen, in this case from an article in *portal: libraries and the academy*, which is available through Project Muse.

Figure 5

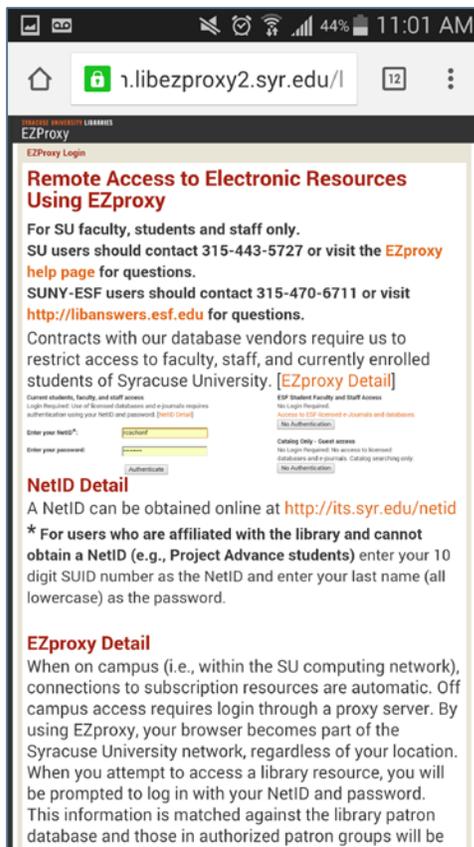


Figure 6



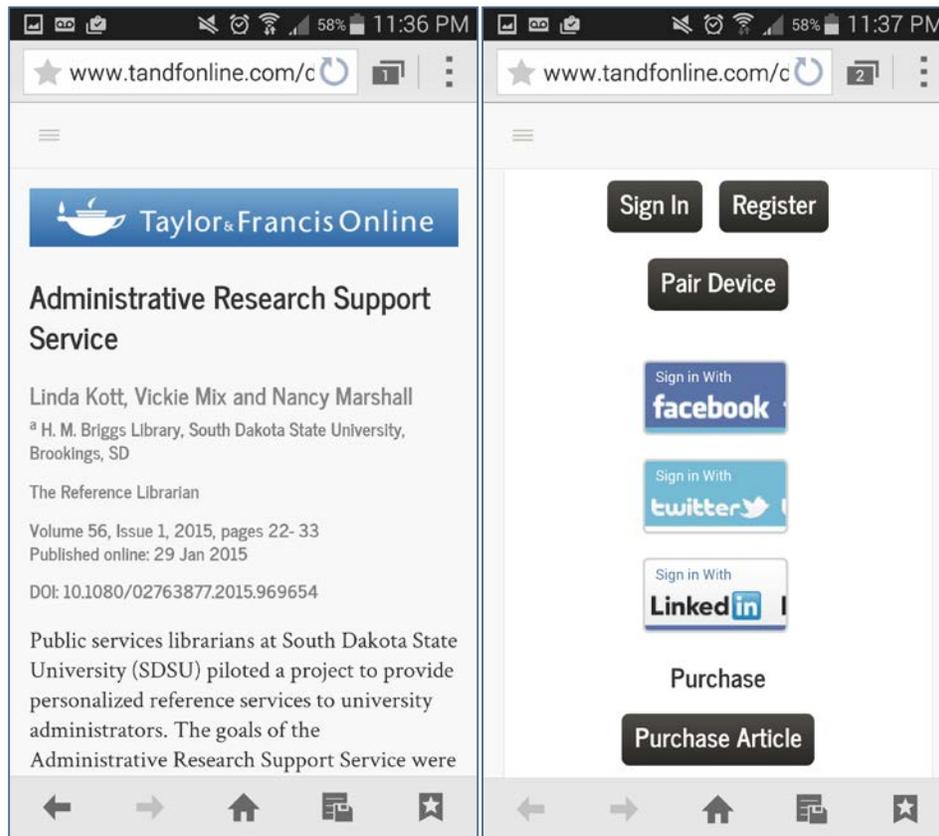
Other content providers and intermediaries have made real efforts to develop interfaces that work well on a mobile device. For example, Figure 7 shows a library homepage that uses a form of responsive design, in this case Yale University, which can also be found among several content platforms and index-based search services.

Figure 7



Other providers try to load an HTML “app” on smaller screen devices. While there may be some functionality advantages to these approaches, there are tradeoffs. In the case of Taylor & Francis, whose mobile app interface is shown in Figure 8, a researcher can pair the device with one’s account while “inside” a T&F authorized site, rather than simply using one’s account credentials, proxy, or other mechanism to log-in. The pairing process requires that you authenticate through your PC, navigate to the pairing website, and then type a code into your phone, a process that may need to be repeated if it expires. On several recent days, the T&F app would not successfully load on my preferred mobile browser (Google’s Chrome), requiring instead the default Samsung browser to load successfully, even then requiring more than a moment to load. While an HTML “app” represents a real investment to try to serve mobile users, it may not be the best approach.

Figure 8



The failure to accommodate mobile devices and use cases is especially pronounced in the overall inability for services to make use of the array of sensors that they offer. Location services, the camera, and the microphone offer fairly obvious opportunities for discovery well beyond the text search box, and a full realization of mobile device needs would begin to take these into account as well.⁸

⁸ For some sense of the broader changes that mobile could bring to scholarly publishing, see Joseph Esposito, "The Elephant in the Room Is a Phone," *The Scholarly Kitchen*, February 12, 2015, <http://scholarlykitchen.sspnet.org/2015/02/12/the-elephant-in-the-room-is-a-phone/>.

User accounts are not well implemented

Many content platforms and discovery services allow researchers to establish a user account to receive access to customized features, including the ability to establish content alerts (when content is added into a certain journal, that matches a search term, cites one's works, or meets other characteristics). These accounts have so much potential beyond these alerts to improve discovery, reduce barriers to access, and increase research efficiency.

In the first place, for off-campus researchers who have had to authenticate via Shibboleth or a proxy to gain access, this subsequent login page for the user account is likely confusing to many. In some cases, these user accounts are linked to a given institutional account in ways that make them all but impossible to use while visiting another institution or upon changing institutional affiliations. In practice, this can mean needing to set up an entirely different account, including alerts and otherwise, when a researcher changes jobs or when a student moves on to graduate school. Connecting a user account with an institutional affiliation is desirable only if it empowers seamless remote access and simple reconfiguration for another university. Shibboleth has been seen as principally an authentication mechanism, but its potential to allow for a single sign-on to a user's content platform account would be valuable if realized.

Today, each platform has its own account, rather than each researcher having a single account for that works across platforms. Looking ahead, there should be an opportunity for the researcher to have control of one's own user account data and carry it seamlessly across platforms to access advances services.⁹

⁹ Roger C. Schonfeld, "Data for Discovery," *The Scholarly Kitchen*, February 5, 2015, <http://scholarlykitchen.sspnet.org/2015/02/05/data-for-discovery/>.

Failure is not inevitable

In recent years, scholarly resource providers such as publishers have devoted extensive effort to understanding researcher behaviors on their platforms. But platform providers are not recognizing one important fact: a researcher does not use a single platform on a standalone basis. Rather, researchers discover scholarly resources through a variety of third-party services; and researchers access scholarly resources within, or working at cross-purposes with, an academic library's systems environment.¹⁰

... a researcher does not use a single platform on a standalone basis.

The switch to remote and mobile work practices has added complexity. Working from a PC in the library, a researcher has great advantages in accessing information resources once they are discovered. But take that researcher out of the library and off-campus, exchange the desktop for a phone, and the system breaks down.

On finding an article one would like to read that is available online and licensed by one's library, a researcher should never have to click seven, ten, or a dozen times, as is completely common today when working off-campus, to gain access to an article that, even so, cannot be read comfortably on a small screen. Let alone to click so many times only to find the article is not available through one's university library! Some of these problems are driving researchers away from using licensed e-resources and towards materials that are available on the open web, although others are encountered equally with open access providers as well. As an information ecosystem, libraries, content providers, and intermediaries, are collectively failing to meet the needs of their users.

These problems need to be addressed from several directions in parallel.

- » Interface design needs to be improved, for example optimizing all screens for mobile devices and eliminating mention of terms that researchers are unlikely to recognize such as "Shibboleth."

¹⁰ At the same time, there has been tremendous consolidation in publishers and platforms, at least for scholarly journals, while some of the multi-publisher content providers such as EBSCO and ProQuest appear to be continuing to grow. Consolidation onto a smaller number of larger platforms may tend to mitigate some of the challenges discussed in this paper, while also raising other challenges.

... interface and experience are baseline requirements for a content platform just as much as a binding is for a book.

- » Libraries need to develop a completely different approach to acquiring and licensing digital content, platforms, and services. They simply must move beyond the false choice that sees only the solutions currently available and instead push for a vision that is right for their researchers. They cannot celebrate content over interface and experience, when interface and experience are baseline requirements for a content platform just as much as a binding is for a book. Libraries need to build entirely new acquisitions processes for content and infrastructure alike that foreground these principles.
- » To understand researcher practices, user experience specialists both in a library and a content provider setting should examine the researchers' actual practices. Rather than trying to focus on specific tasks related to the system that their current project covers, as is all too often the approach taken, a more holistic, ethnographic perspective is vital. One place to start understanding the issues addressed in this paper is to try oneself to conduct some research from off-campus - from one's living room, perhaps - since this is the simplest way to understand how practices and experiences differ when away from one's desk and outside one's campus network.
- » Libraries and content providers must cooperate more effectively to overcome the impediments on linking from the no-access dead-ends on content provider webpages to Shibboleth and proxy authentication systems. A reliable knowledge base of these authentication systems that could automatically update content provider response pages is needed. Shibboleth is currently implemented painstakingly through bilateral agreements; multi-lateral agreements, perhaps facilitated by intermediaries or alliances, would reduce this barrier.
- » If the index-based search services are to serve as the library's primary point of entry to its collections, they must be updated more frequently. It makes little sense that the free Google Scholar knows about a new article before the library's vendor-provided index. Even if lags have declined, these indices cannot realize the vision of serving as a primary search starting-point for researchers if they are not up to date, and the lack of real-time updating hobbles their potential to provide reliable alerting and other more advance anticipatory discovery services.
- » Finally, it is time for a major commitment from the scholarly information ecosystem of libraries, publishers, university IT, and intermediaries, perhaps under the auspices of NFAIS, NISO, the Shibboleth Consortium, or another not-for-profit organization, to develop a single user account for all scholarly e-resources. This account would not only provide authentication via a researcher's institutional credentials but also would be the vehicle through which a variety of additional data-driven services could be provided on an opt-in basis. The account itself as well as the data it contains would be under the control of the researcher, and it would therefore travel with the researcher when changing institutional affiliations.

Research practices continue to push the boundaries of the platforms and infrastructure developed under a previous set of assumptions. We must strive collectively to get ahead of them if we are to serve researchers' needs effectively.¹¹

¹¹ For a recent survey of librarians on many of the issues discussed in this issue brief, see OpenAthens, "Librarians' experiences and perceptions of Identity and Access Management," <http://www.eduserv.org.uk/~media/Insight/Reports/OpenAthens%20Librarian%20ReportIAM1180.pdf>.