Chapter 1

The Access and Delivery Problem for Libraries

Introduction

The problem of discovery has been a long-standing area of study for academic libraries. In the online era, we transitioned from abstracts and indexing subject databases to full-text databases. This was followed by the attempt to aggregate access to all the disparate e-resources in one centralized platform. To that end, we went from federated search engines to web-scale discovery engines in the 2010s, and the promise of semantic linked data now beckons as we struggle to define our online services in comparison to Google.

However, as interesting as discovery is to academic libraries, it tells only part of the story. While discovery—the process of helping users find what is potentially useful to them—is important, access and delivery are equally important. The access and delivery process allows users to easily check if they have access to the content that they are interested in and to quickly access the content. If a user does not have access, they should be supplied with useful options such as getting access via document delivery.

In the online environment, these are some typical scenarios that users face:

- *Scenario A:* A faculty member googles an article and lands on a JSTOR journal article landing page. Because they are off campus, JSTOR is unable to determine that they have institutional access, and they get a paywall. How does the faculty member authenticate and access the full text?
- *Scenario B:* Same scenario as above, except this time the link leads to an article on a journal provider's platform that the subscriptions of the faculty member's institution do not cover, though the same article can be accessed on another platform. How does the faculty member know this?
- *Scenario C:* A postgraduate finds a paper of interest by searching on the Mendeley online platform. How do they know whether they have access to the full text through their institution?

While I have used access to a journal article as examples of content the user wants to access, similar scenarios apply for users trying to access e-books, databases, and other resources licensed online by the library. The fact that today's online context involves the user moving between multiple devices (e.g., desktop, mobile phone, tablet) further complicates matters.

In all three scenarios, if the users are off campus, it may not be obvious to them how they can access the content or even determine if they have access to it. Alternatively, they may be given options that result in broken links, which may mean inaccurate metadata was supplied. While this report focuses on newer authentication workflows, many of these methods ultimately rely on the accuracy of metadata that is shared throughout the supply chain. This report will briefly mention some of the issues related to inaccuracy of metadata and how some systems, like GetFTR, try to improve on traditional methods. I recommend referring to other texts on the topic such as Pacific University Press's Managing Licensed E-Resources web page for more guidance on the issue.

> *Managing Licensed E-Resources*
> http://www.lib.pacificu.edu/create/pup/pacific-university-press-all-books/pup-managing-licensed-eresources/

In the online environment, solving such issues typically falls under access management, and understanding the concepts of authentication and authorization

will be helpful. We will cover that topic in the next chapter. However, for now, let us consider why it is worthwhile to spend effort studying and understanding the access and delivery issue.

Why Care about the Access and Delivery Issue?

- While discovery is getting less important, delivery and access are becoming steadily more important for academic libraries.
- Delivery and access library solutions need to be more seamless to increase patron awareness of library resources.
- Libraries need to be aware of what new alternatives are available to help solve the problems of access and delivery.
- Publishers are affected by competitors such as Sci-Hub and the fear of leakage.

While Discovery Is Getting Less Important, Delivery and Access Are Becoming Steadily More Important for Academic Libraries

In the mid-2000s, Lorcan Dempsey coined and popularized the phrase "Discovery happens elsewhere" and a few years later introduced the idea of library services, particularly discovery services, moving to the network level, which he dubbed "web-scale."[1]

Both trends collectively foreshadowed the decline in the prominence of library discovery services and the rising popularity and importance of large web-scale gateway services such as Google, Google Scholar, and ResearchGate. Since then, various surveys of researchers have confirmed the rising importance of academic search engines and academic social networking sites and the declining importance of library discovery services.[2]

One of the first academic libraries to take this seriously was Utrecht University, which argued provocatively in a series of talks that it was time to start "thinking the unthinkable" and even consider "doing away with the library catalogue." Simone Kortekaas and Bianca Kramer of Utrecht University argued back in 2014

> At Utrecht University we strongly believe that academic libraries have lost their role in the discovery of scientific information and should focus on delivery instead. . . . We have to admit that others can do a better job on discovery, so don't spend too much time on this. Make a priority of your delivery task and rethink the way you can provide value for your users.[3]

As a result, Utrecht University closed down its custom-made discovery tool, Omega, and did not replace it with then-trendy web-scale discovery services (EBSCO Discovery Service, Summon, Primo, etc.) that most academic libraries were implementing or already had in place. Instead, it focused on supporting delivery of items for users who were using Google Scholar and OCLC's WorldCat. It also developed a simple JavaScript bookmarklet for users and a Chrome browser extension called UU Easy Access, both of which helped users gain access when they were on a page without needing to go back to the library home page. The browser extension, developed by a staff member at Utrecht University, was later spun off to become Lean Library, one of the leading access broker browser extensions, which we will feature in chapter 3.[4] While Utrecht's approach was not very popular, one does not need to take an extreme position on discovery to recognize that focusing on improving delivery may also be a good idea. (This is not to say libraries should totally ignore the discovery issue. Academic libraries increasingly have to engage in activities to promote content from their community, such as to enhance the discovery of content deposited into institutional repositories.)

WorldCat
https://www.worldcat.org

Lean Library
https://www.leanlibrary.com

Delivery and Access Library Solutions Are Not Seamless Enough

Whether or not you accept the idea that the role of library discovery is diminishing, is there any reason our delivery and access options need improving? In a 2015 analysis that was influential, at least in the publisher world, titled *Meeting Researchers Where They Start: Streamlining Access to Scholarly Resources*, Roger C. Schonfeld of Ithaka S+R systematically detailed the problems researchers face when trying to access library resources.[5] Some of the issues he identified were

- access solutions being overreliant on users starting at the library websites or doing research only on campus
- overly complicated, unintuitive workflows as well as inconsistent, confusing labeling for proxy solutions and federated SAML solutions such as Shibboleth
- access and delivery solutions that are often not optimized for mobile devices or multiple devices
- unstable and unreliable linking mechanisms

Table 1.1. Some current fundamental issues with library delivery and proposed solutions

Need	Traditional solution	Weaknesses	Emerging solution
User lands on a content owner site and tries to access the content. Site needs to authenticate user.	IP and proxy solution	• Unintuitive to use when off campus and when not starting at library home page	• Access broker browser extensions (chapter 3) • Campus Activated Subscriber Access (CASA; chapter 5) • Content syndication with ResearchGate (chapter 5)
	Federated access	• Nonstandard and unintuitive log-in screen	• SeamlessAccess (chapter 4)
User lands on a discovery platform site and tries to check if they have access anywhere. Site needs to determine where to send the user where they may have access (including free-to-read copies).	Library link resolver (typically OpenURL)	• Poor link reliability (various reasons) • Generated links not automatically leading to full text	• GetFTR (chapter 4) • LibKey infrastructure (chapter 3) • Access broker browser extensions (chapter 3)
User needs to access resources on multiple devices, including mobile devices.	None		• Campus Activated Subscriber Access (CASA) validation shared across all devices with the same Google account (chapter 5)

Most of these issues are probably not news to librarians, who often need to help users with e-resource troubleshooting. In particular, some of these problems—such as the inadequacy of IP authentication with proxy-based solutions and the unreliability of library linking solutions, particularly OpenURL—had been well known for over a decade,[6] although the issue with mobile devices was relatively new.

Still, Schonfeld's analysis was one of the first by a nonlibrarian to bring it all together. Promoted further by Schonfeld, a relatively prominent member of the publisher community, in talks as well as on the popular blog *Scholarly Kitchen*,[7] it got a lot of attention among librarians and also caught the interest of publishers and researchers active in scholarly communication circles. We will consider a deeper analysis of some of these problems in the next chapter, but for a brief summary of issues and some possible solutions, refer to table 1.1.

Availability of New Alternatives to Solve the Problems of Access and Delivery

Even if we admit that the library access and delivery process has room for improvement, are solutions available? Indeed, since the mid-2010s, a whole slew of tools and technologies have started to emerge that either try to improve on existing technologies—IP authentication and proxy solutions, federated access solutions—or introduce new methods for access and delivery. Some of these solutions, which will be covered in future chapters, are

- library access broker browser extensions—chapter 3
- federated access solutions—SeamlessAccess and GetFTR—chapter 4
- Google's Campus Activated Subscriber Access (CASA), content syndication partnership between Springer Nature and ResearchGate—chapter 5

While these solutions may not handle every issue identified by Schonfeld's analysis, they do mostly address the meat of the issue, which is to help provide more seamless access to content.

Competitors Such as Sci-Hub and the Fear of Leakage

While some of the solutions mentioned above are journal-publisher-independent approaches, many publishers are now supporting federated identity approaches to access via the SeamlessAccess coalition, which was the successor organization to 2016's RA21, as well as the GetFTR initiative in 2019. The SeamlessAccess coalition also seeks to eventually eliminate IP-based authentication methods of access.[8] Why this sudden interest by publishers in improving the delivery process?

SeamlessAccess Coalition
https://www.seamlessaccess.org

GetFTR
https://www.getfulltextresearch.com

While we can speculate on why publishers decided on such a move, one reason is probably competition from Sci-Hub. Sci-Hub, a website that illegally provides free access to academic content, first rose to prominence in 2015 and 2016, when big publishers such as Elsevier brought lawsuits against it. While a certain amount of piracy was expected, an analysis

of Sci-Hub logs from September 2015 through February 2016 led to the surprising finding that many of its users were academic users who probably already had legal access via their institutions.[9]

Similarly, an analysis by Bianca Kramer of Utrecht University, using the same logs, attempted to answer a similar question: "Do people use Sci-Hub to get papers they do not otherwise have access to, or do they (also) go to Sci-Hub for convenience: a one-stop shop to get access, without having to navigate library and publisher websites?"[10] By restricting analysis to entries in the Sci-Hub log that correspond to the IP addresses from the university, Kramer found that over 60 percent of Sci-Hub accesses were to content that could be downloaded via library subscriptions and inferred that many users were using Sci-Hub for convenience rather than access.[11] If this surprising finding is typical, it suggests that library access and delivery options are perhaps often too unintuitive for many researchers so that the researchers prefer to use Sci-Hub rather than jumping through library access and authentication hoops to gain access to what they want.

So why do these findings bother publishers? Since the value of big deals that libraries sign with publishers reflects the usage reports (typically downloads from COUNTER reports) generated by publisher platforms, any usage that happens off the platform on sites like Sci-Hub would lower the value of the deal and would be considered what Roger Schonfeld called publisher "leakage."[12] Recent surveys of readers revealed that

> If you compare the data which shows that 80% of people are viewing the version of record with the data that shows only around 45–50% of article downloads come from the publisher web site, we can conclude that a significant proportion of Version of Record articles are being obtained from resources outside of publisher control.[13]

To be fair, if leakage is defined as any usage that happens off the publisher platform, it would include downloads by users from both legitimate sources such as SSRN and institutional repositories and illegal sites like Sci-Hub. As we shall see in chapter 4, publisher-supported initiatives like SeamlessAccess and GetFTR help plug this leakage by providing more seamless access to downloads available via publisher platforms.[14]

Does Open Access Make the Access and Discovery Question Moot?

The open-access (OA) penetration rate has been steadily rising, and for the publication year 2020, over 50 percent of publications tracked by Digital Science's Dimensions product were found for the first time to be OA.[15] (As of December 2021, Dimensions tracks over 100 million articles.) A conservative projection is that by 2025, 70 percent of all article views will be of OA articles.[16] With developments like Plan S in the wings, it seems likely that the momentum toward OA will continue. Does this mean that focusing on access and delivery is likely to be wasted effort since in a soon-to-be OA world, access and delivery problems will automatically solve themselves? Not necessarily.

First, such projections tend to use a broad definition of OA and usually include both version of record (VoR) and accepted manuscript, if not even earlier versions. Authors are known to have clear preference for VoR, so this is not a trivial distinction. If we achieve high levels of OA in the near future, but a big part of this OA is achieved via non-VoR, there is still a place for access and delivery mechanisms to guide the user to VoR copies.

Second, as it currently stands, when we talk about OA, we are talking mostly about journal articles and to a much lesser extent about books. However, the library provides access to paywalled e-resources that go beyond journal articles and books. For example, the library provides access to A&I databases, financial and business databases, and image databases, all of which will continue to require good access and delivery mechanics for our users.

Last, the rising availability of OA copies of various versions of resources also means that the library can play a role in guiding users to help them discover and access OA copies when they are stumped by a paywall that cannot be bypassed by library subscriptions.

Conclusion

In this chapter, we started by distinguishing the problem for academic libraries of supporting discovery from that of supporting access and delivery. We asserted that delivery and access are an important area of library service to focus on and provided four reasons for that assertion. Chief among the reasons was the less-than-intuitive nature of our current access and delivery systems, which may have led some of our readers to prefer to use Sci-Hub, which does not have any authentication system. We also briefly mentioned the major different tools and solutions available to libraries and ended with a section explaining why even with the rise of OA, libraries should still focus on delivery and access.

In the next chapter, we will explain the basic concepts of authentication and authorization for library e-resources.

Notes

1. Lorcan Dempsey, "Discovery Happens Elsewhere," *LorcanDempsey.Net* (blog), September 16, 2007, https://www.lorcandempsey.net/orweblog/discovery-happens-elsewhere/; Lorcan Dempsey, "Thirteen Ways of Looking at Libraries, Discovery, and the Catalog: Scale, Workflow, Attention," *Educause Review*, December 10, 2012, https://er.educause.edu/articles/2012/12/thirteen-ways-of-looking-at-libraries-discovery-and-the-catalog-scale-workflow-attention.
2. Tracy Gardner and Simon Inger, *How Readers Discover Content in Scholarly Publications: Trends in Reader Behaviour from 2005 to 2021* (Renew Consultants, July 2021), https://renewconsultants.com/wp-content/uploads/2021/07/How-Readers-Discover-Content-2021.pdf.
3. Simone Kortekaas and Bianca Kramer, "Thinking the Unthinkable: Doing Away with the Library Catalogue," *Insights* 27, no. 3 (November 3, 2014): 244, https://doi.org/10.1629/2048-7754.174.
4. Johan Tiistra, "Lean Library's Browser Extension: Seamless Delivery for Users," guest post, *Aaron Tay's Musings about Librarianship* (blog), May 30, 2017, https://musingsaboutlibrarianship.blogspot.com/2017/05/guest-post-lean-librarys-browser.html.
5. Roger C. Schonfeld, *Meeting Researchers Where They Start: Streamlining Access to Scholarly Resources*, issue brief (New York: Ithaka S+R, March 26, 2015), https://sr.ithaka.org/publications/meeting-researchers-where-they-start-streamlining-access-to-scholarly-resources/.
6. Cindi Trainor and Jason Price, *Rethinking Library Linking: Breathing New Life into OpenURL* (Chicago: American Library Association, 2010).
7. Roger C. Schonfeld, "Dismantling the Stumbling Blocks That Impede Researcher Access to E-Resources," *Scholarly Kitchen* (blog), November 13, 2015, https://scholarlykitchen.sspnet.org/2015/11/13/dismantling-the-stumbling-blocks-that-impede-researcher-access-to-e-resources/.
8. Todd A. Carpenter, Heather Flanagan, and Chris Shillum, "Myth Busting: Five Commonly Held Misconceptions about RA21 (and One Rumor Confirmed)," *Scholarly Kitchen* (blog), February 7, 2018, https://scholarlykitchen.sspnet.org/2018/02/07/myth-busting-five-commonly-held-misconceptions-ra21/.
9. John Bohannon, "Who's Downloading Pirated Papers? Everyone," *Science* 352, no. 6285 (April 25, 2016): 508–12, https://doi.org/10.1126/science.352.6285.508.
10. Bianca Kramer, "Sci-Hub: Access or Convenience? A Utrecht Case Study (Part 1)," *Library Blog*, Utrecht University, June 20, 2016, https://ubublog.sites.uu.nl/2016/06/20/sci-hub-utrecht-case-study-part-1/.
11. Bianca Kramer, "Sci-Hub: Access or Convenience? A Utrecht Case Study (Part 2)," *I&M/I&O 2.0* (blog), June 20, 2016, https://im2punt0.wordpress.com/2016/06/20/sci-hub-access-or-convenience-a-utrecht-case-study-part-2/.
12. Roger C. Schonfeld, "Is the Value of the Big Deal in Decline?," *Scholarly Kitchen* (blog), March 7, 2019, https://scholarlykitchen.sspnet.org/2019/03/07/value-big-deal-leakage/.
13. Gardner and Inger, *How Readers Discover Content*, 32.
14. Roger C. Schonfeld, "Publishers Announce a Major New Service to Plug Leakage," *Scholarly Kitchen* (blog), December 3, 2019, https://scholarlykitchen.sspnet.org/2019/12/03/publishers-announce-plug-leakage/.
15. Daniel Hook, "Open Access Surpasses Subscription Publication Globally for the First Time," *Dimensions* (blog), February 24, 2021, https://www.dimensions.ai/blog/open-access-surpasses-subscription-publication-globally-for-the-first-time/.
16. Heather Piwowar, Jason Priem, and Richard Orr, "The Future of OA: A Large-Scale Analysis Projecting Open Access Publication and Readership," preprint, bioRxiv, October 9, 2019, 795310, https://doi.org/10.1101/795310.