Chapter 2

Authentication, Authorization, and the Appropriate Copy Problem

Some Basic Concepts for Access Management of Library Resources

Before we start to discuss problems and solutions for access to library resources, it is useful to know some basic concepts regarding authentication and authorization for access management.

The issue that authentication and authorization attempt to solve boils down to the following question: When a user lands on a content owner's platform, such as a journal platform, should the content owner allow the user to access paywalled content?

Another related problem common to delivery of library resources, particularly for journal articles, is the "appropriate copy problem."[1] The appropriate copy problem arises from the fact that content such as journal articles can reside in multiple locations online. For example, a journal article can be available at a publisher site (such as Wiley), an aggregator or a reseller site (such as EBSCOhost platform), and open-access repositories (such as institutional repositories), and the most appropriate copy varies depending on the entitlements of the user making the request. (For example, what institution do they belong to, and given their position, what are they allowed to access?) This issue comes up particularly for discovery systems such as Google Scholar and citation indexes, which do not carry the full-text content.[2] Because these are the first point of reference that many students look to when researching online, these researchers may not know that they would have full-text access through their library's database. As a result, answering the appropriate copy problem here involves determining where to direct users to get the most suitable copy. See figure 2.1 for example scenarios of these issues.

In this chapter, we will focus on the fundamental concepts necessary for understanding authentication issues and the solutions that traditionally have been used to answer them. We'll also look at some common problems that occur with these traditional solutions.

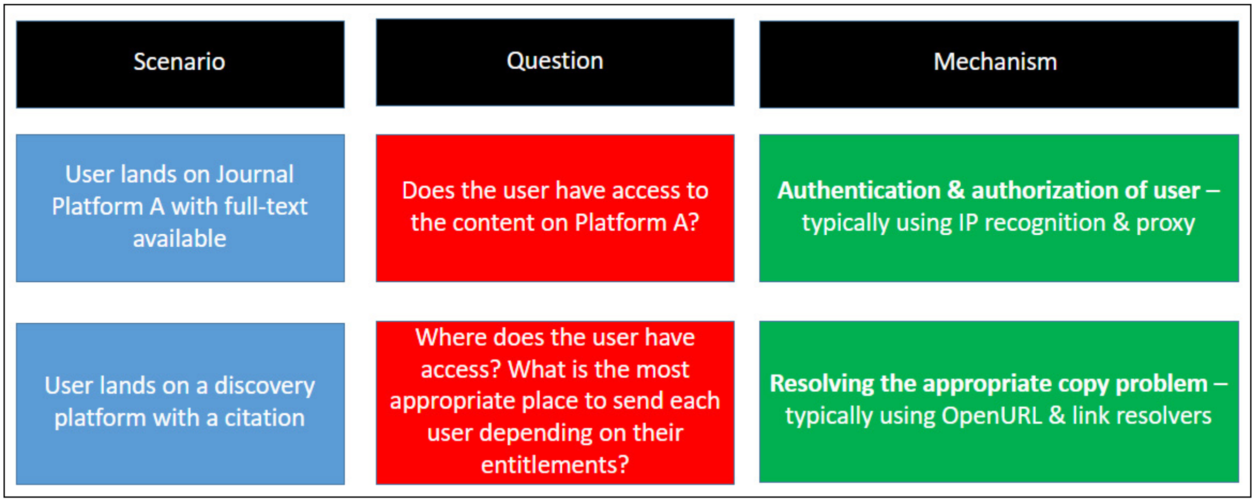
Authentication and Authorization

Understanding how access management works can be technical; however, a very good resource targeted at librarians exists—Kristina Botyriute's *Access to Online Resources: A Guide for the Modern Librarian* provides the essentials needed for a library worker to quickly get up to speed with the issues.[3] I encourage you to refer to that source for more detail.

From a technical point of view, when someone logs in to access a resource, they go through two distinct but related processes:

1. The process of *authentication* confirms that users are who they say they are.
2. Once users are authenticated, the process of *authorization* ensures users are given the right permissions to access resources.

Take this simple example: an undergraduate student from the school of social science may log in to your system with a username and password. After the system authenticates the student, it looks up what access rights they have and grants those rights to them. This is the process of *authorization.*



Scenario	Question	Mechanism
User lands on Journal Platform A with full-text available	Does the user have access to the content on Platform A?	**Authentication & authorization of user –** typically using IP recognition & proxy
User lands on a discovery platform with a citation	Where does the user have access? What is the most appropriate place to send each user depending on their entitlements?	**Resolving the appropriate copy problem –** typically using OpenURL & link resolvers

Figure 2.1
Two common delivery issues for library resources—authentication and authorization, and the appropriate copy problem.

While the two processes are related, they are distinct. For example, a user trying to log in to the JSTOR database to access a journal article can be successfully authenticated as a current student at Institution X (we will discuss how later), but they may not be authorized to access that particular article in JSTOR. Similarly, two users from the same institution but different departments may have different access rights. For example, a medical researcher might have access to Embase, a specialized medical database, that another researcher from a different department in the same institution might not have. In the next section, we will discuss the three major ways academic libraries provide authentication and authorization today: (1) individual account passwords, (2) IP recognition, and (3) SAML-based SSO methods.

Providing Access with Individual Usernames and Passwords

Imagine a scenario where you are the electronic resource librarian at an institution that has successfully negotiated a subscription with access to a bundle of journal titles on the JSTOR platform. Great! Now, how does the publisher ensure that only authorized people (users from your institution) are allowed access to the full-text articles in these journals on JSTOR? One obvious but very uncommon way (particularly in this scenario) is to issue individual usernames and passwords to everyone. Each user enters their own username and password to authenticate themselves and access the resource. From the user's point of view, registering and remembering a separate set of user credentials for each library resource is inconvenient. For many students, any access barrier is likely to push them into the arms of free web services and content. Another issue is how to handle turnover when users join and leave your institution. Surely the publishers expect you to ensure that only current students and staff have access, which requires quite a bit of maintenance. Now multiply this effort by the number of resources you subscribe to. Clearly, doing this manually is not sustainable except for a small select number of resources with low usage.

Box 2.1

Rarely Used Work-Arounds for Passwords

One attempt to work around the problem with passwords is to provide a single shared username and password for all members of your institution. Typically, this works by making users of your institution sign in and authenticate themselves first on a web page before displaying the username and password they can use directly on the resource.

The problem with this solution is that it is not very user-friendly because the user needs to authenticate twice (once with the institution and once with the publisher) to gain access. In addition, there might be concerns on whether the account will be shared with unauthorized users. It is difficult to track who is actually using the account if there is a need for this information. Still, this might be the only solution for publishers that do not support IP authentication methods.

There have been other work-arounds, such as making users log in to virtual environments or embedding passwords and tokens in EZproxy sign-ins, but such work-arounds are quite frail and can easily break.

For some other work-arounds to address these issues, see box 2.1.

In this example, JSTOR is a popular database highly used by students and researchers. Therefore, manually maintaining individual accounts and passwords is

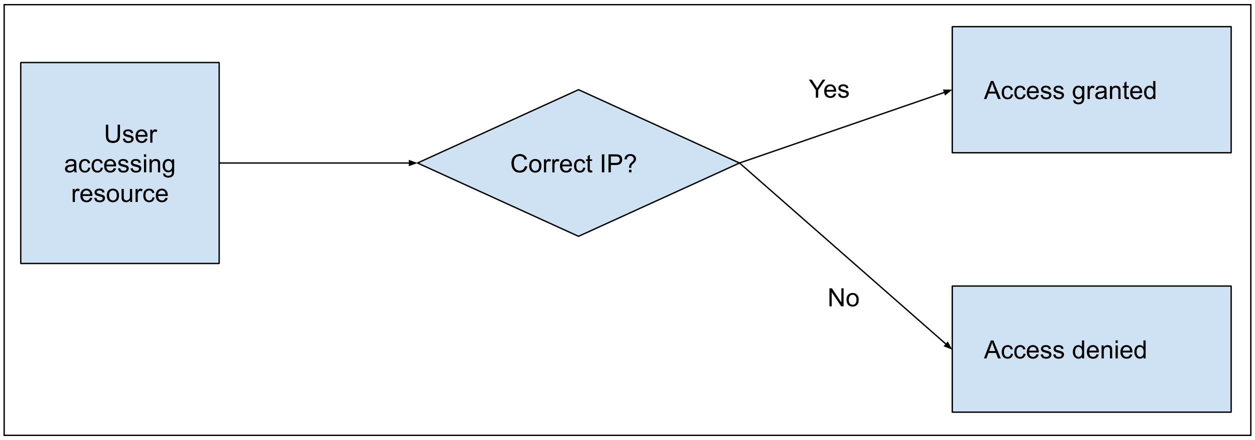


Figure 2.2
Accessing library resources via IP authentication (correct IP vs. wrong IP)

most certainly not viable. In fact, academic libraries seldom provide access by individual usernames and passwords because the effort is too great. Instead, IP recognition is far more commonly used today.

Providing Access through IP Recognition

As we have seen, issuing individual passwords is not sustainable. Thankfully, this is not the main way institutions provide access for users to most of their resources. Today, access to most online resources subscribed to by libraries is provided via IP recognition. The idea is simple. When electronic resource librarians subscribe to an online resource, all they need to do is provide a list of IP addresses (the IP range) that are used by users of your community to access the resources. Typically, this would be the IP range of your users when they are on campus using the campus Wi-Fi. The publisher of the resource will set up its server to allow access whenever it receives a request coming from these IP addresses. Put in another way, we create a whitelist of IP addresses where requesters from those IPs are allowed access (see figure 2.2).

From the users' point of view, access is seamless because they do not need to do anything, not even explicitly sign in, as long as they are on campus and in the campus Wi-Fi range. Arguably, access might be *too seamless*, as users may not even know that they are accessing the institution's subscriptions if they miss the sometimes-subtle signs on the publisher platforms that recognize them via IP.

THE OFF-CAMPUS PROBLEM: IP RECOGNITION AND PROXY SERVERS

So far, we have seen that when libraries use IP recognition to provide access, users get a very seamless experience as long as they are requesting the resource via the right IP address (i.e., they are on campus using campus Wi-Fi). However, in our global and post-COVID-19 world, expecting our users to access resources only on campus seems unrealistic. So how do we provide access with IP recognition when users are off campus? There are two main methods: (1) proxy servers and (2) VPNs. Both methods make the user's request appear to be from the right IP address, but proxy methods are far more popular today, so let us discuss them.

A proxy server, in simplified terms, is a piece of software that sits between you, the user, and the online resource you are trying to access; it sends and retrieves content on your behalf. Today, the most popular proxy server used for this purpose in libraries is OCLC's EZproxy, but others exist. Let's see how this works. Again, let us take the example of a user trying to access a journal that is available only behind a paywall. If the user is off campus and tries to directly access the resource, they will be denied access because their IP address is not recognized (see figure 2.3).

One way around the problem is through a proxy server, which requests the resource on behalf of the user and retrieves the content on their behalf (see figure 2.4).

But how does the user get the proxy server to make the request on their behalf? They will need to use a specially treated link to do so—one that is set up to direct user requests through the proxy. This type of link is informally called a *proxied link*. Here is an example of such a proxied link from my institution. This link, when clicked, directs the user's request to access the JSTOR database (in bold type: http://www.jstor.org) via the proxy server.

> http://libproxy.smu.edu.sg/login?url=http://libproxy.smu.edu.sg/login?url=**http://www.jstor.org**[4]

But how does the user find such a link? One way is for the user to go to the library home page, look for

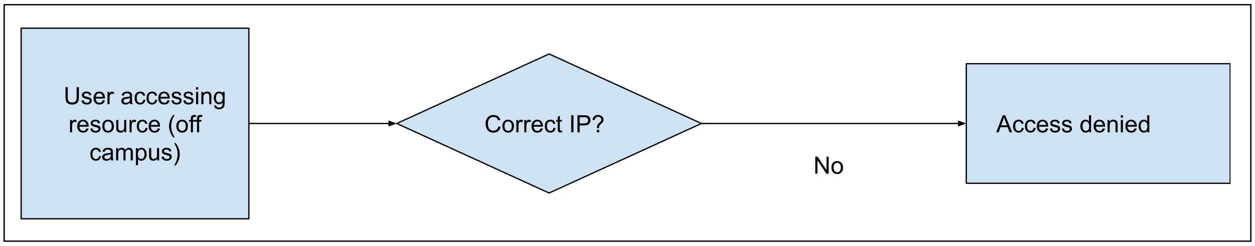


Figure 2.3
Prevented from accessing library resources when off campus

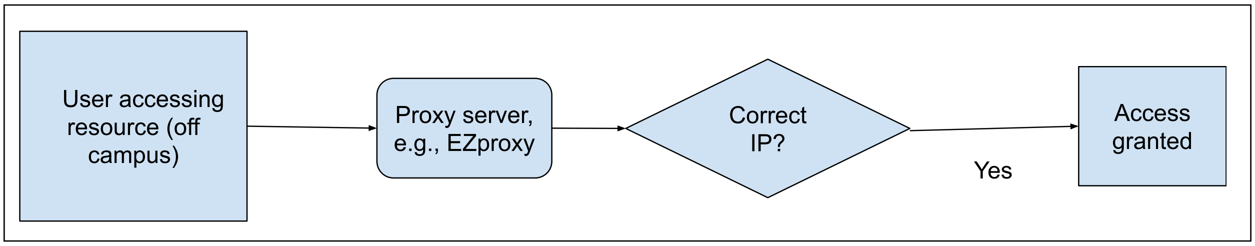

User accessing resource (off campus)
Proxy server, e.g., EZproxy
Correct IP?
Yes
Access granted

Figure 2.4
Granted access via proxy when accessing library resources off campus

the desired online resource (e.g., via the library search engine or database A–Z list), and click on the link provided; access is granted. How do we then ensure that unauthorized users cannot use this method by using the previous link? Simple: whenever someone tries to access a resource via the proxy server, they will need to authenticate themselves with a sign-on.

Assuming the proxy server configuration is set correctly for each online resource that the library is licensed to access, users need only to use the same sign-on (which typically is their institutional sign-on credentials) each time regardless of the online resource they are trying to access via the proxy. For more details on proxy servers and configuration settings for libraries, please refer to the documentation of the proxy server you are using.

Overall, despite any drawbacks of this method (see discussion in the next section), IP recognition is currently the dominant way access is provided. For a typical library, access to 70 to 80 percent of resources will be provided this way, though SAML-based methods may be rising in popularity.

Single Sign-On with SAML

As noted in chapter 1, in the section Delivery and Access Library Solutions Are Not Seamless Enough, one of the current major situations that cause access to be less seamless is when users are off campus. To benefit from IP authentication and proxy solutions when off campus, they will need to start from library-controlled pages with proxied links.

Unfortunately, we know that most of our users do not start their research from our library home pages. Assuming they are off campus when they land on a resource, they will not be able to benefit from IP recognition, nor use the proxy, unless they have installed a software solution, such as an access broker browser extension like Lean Library, that helps them with access (see chapter 3).

However, some of these web pages have a log-in button or even a Log in with Your Institution button, and some users are able to obtain access through that method. Other times they may see strange jargon like “Log in with Shibboleth” or “Sign in with OpenAthens” and try to log in with those (see figure 2.5).

Both Shibboleth and OpenAthens employ SAML (Security Assertion Markup Language) technology. At their best, such solutions will be intuitive and user-friendly enough that users can easily select their institution (a process known as Where-Are-You-From, or WAYF, which will we discuss further in chapter 4) and then sign in immediately with their existing university user credentials without the need to create new accounts and passwords. The experience is akin to options like “Sign in with Google” or “Sign in with Facebook” that you may have used to sign on to other platforms, except that instead of using your social network credential account, you use your institutional account.[5] This type of sign-in process is known as *single sign-on* (SSO) and can be implemented in a few ways. In the academic library space, SAML-based technologies are usually employed and are the major alternative to IP recognition.

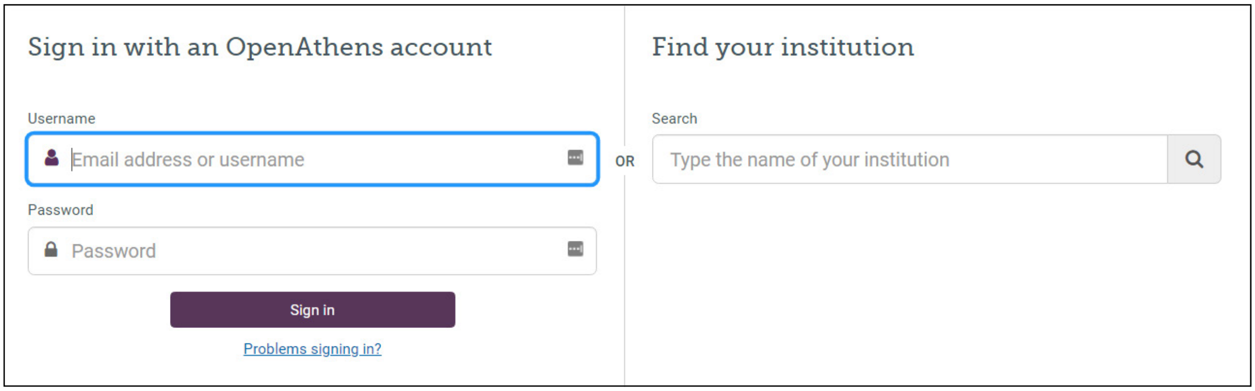

Sign in with an OpenAthens account
Username
Email address or username
Password
Password
Sign in
Problems signing in?
OR
Find your institution
Search
Type the name of your institution

Figure 2.5
Example of SAML sign-in options

SAML IN DEPTH

SAML is an open standard used for identity management by allowing different parties to exchange authentication and authorization data. The standard, which was first created in 2003 and was updated to 2.0 in 2005, underlies both Shibboleth and OpenAthens log-ins, which are commonly used in the academic library space.

See box 2.2 for information on differences between Shibboleth and OpenAthens.

SAML SSO methods improve on simple account password systems in two ways: (1) the user does not need to register and create user accounts in advance, and (2) the user does not need to create and remember new usernames and passwords for each SAML-enabled service. Instead, they may just need to use the institutional credentials that they have no doubt memorized by using them for accessing common university services, such as e-mail, university Wi-Fi, and so on. At worse, they may just need to remember one more common password for access to all library electronic resources. All access is controlled centrally, so access to all these services, including SAML-enabled services, can be revoked when the user leaves the institution.

So how does this work under the hood? Whenever a user tries to log in via Shibboleth- or OpenAthens-enabled resource, they select the institution they claim to be from. The service, which is termed a *service provider* (SP) in SAML speak, doesn't take this at face value but redirects the user back to an identity provider (IdP) to verify that they really are from the institution selected. The IdP may then verify the user. Typically, the user might sign in with their institutional password, and once the user is verified, the IdP will redirect them back to the original SP and assert that the user is indeed verified as being from the institution they claim to be from. The SP will use this information to provide access (see figure 2.6).

For example, a user clicks to sign in to JSTOR via OpenAthens and indicates they are from your institution. JSTOR, which is the SP, redirects the user to your university's IdP, which is usually a server where you sign on to access various university-related services such as e-mail. The user then signs in as normal, and if authentication is successful, the IdP will redirect the user back to the SP (JSTOR) with an assertion confirming that the user is indeed from your institution.

In chapter 4, we will discuss in more detail what assertions are eligible to be sent back to the SP, but for now let's just say that the IdP asserts to the SP that that user is a valid user from your institution. The SP can now be sure that the user is a valid member of your institution and can provide the appropriate level of rights (authorization). If most library resources are enabled to support SAML in the same way, this means all the user needs to remember is one set of account

Box 2.2

What Is the Difference between Shibboleth and OpenAthens?

Both Shibboleth and OpenAthens support SSO infrastructure via SAML.

Shibboleth is open-source software and can be difficult to install and manage for libraries with little experience. A typical library would need to work with the institution's campus IT department to setup Shibboleth use for library resources. Using OpenAthens is less complex than using Shibboleth because it is a cloud-based solution for libraries looking to go down the SAML route. Among other advantages, OpenAthens provides easy-to-use analytics and support (governed by a service level agreement) for setting up access to different resources, troubleshooting, and more. It is essentially an easy way for libraries with no expertise in SAML to set up an identity server.

For more information on OpenAthens, refer to its website, https://www.openathens.net.

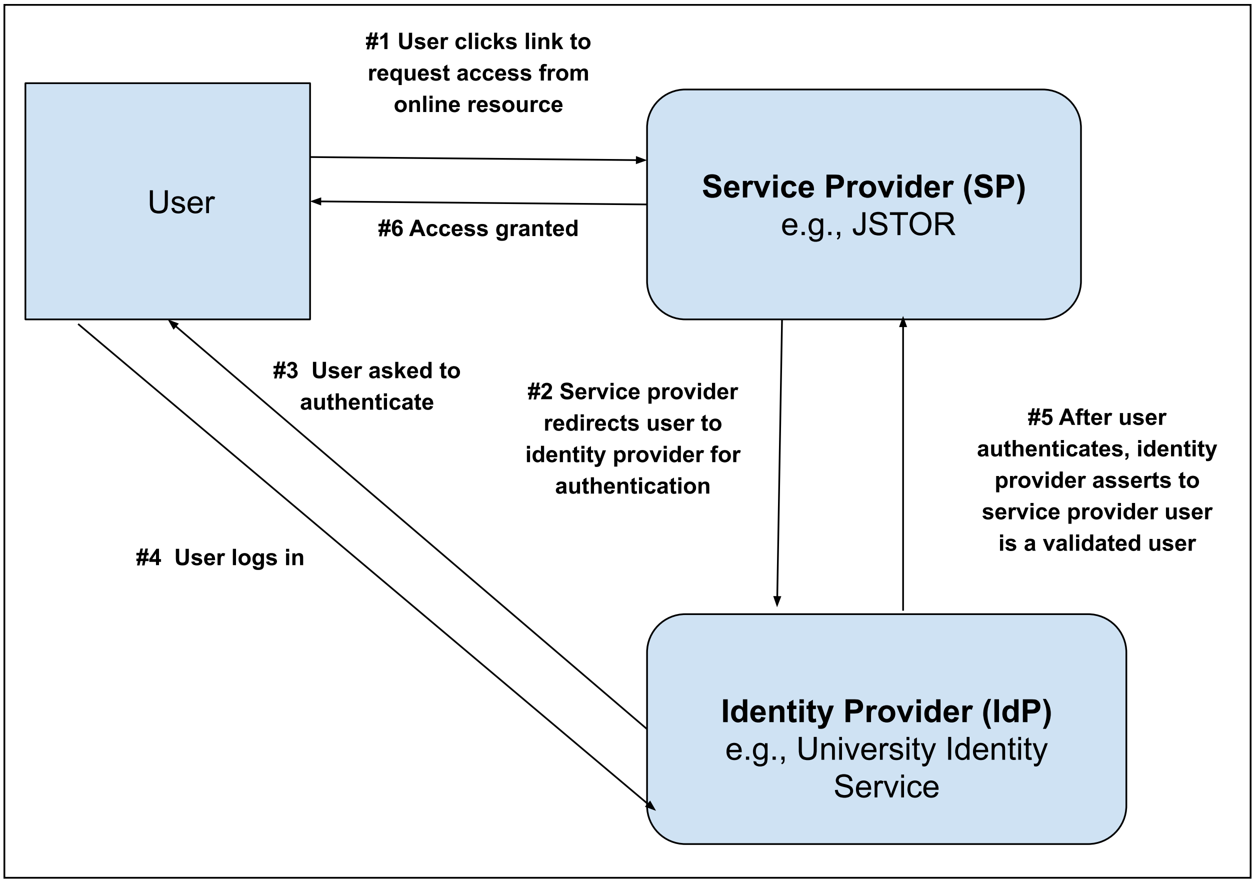

#1 User clicks link to request access from online resource
User
#6 Access granted
Service Provider (SP) e.g., JSTOR
#3 User asked to authenticate
#2 Service provider redirects user to identity provider for authentication
#5 After user authenticates, identity provider asserts to service provider user is a validated user
#4 User logs in
Identity Provider (IdP) e.g., University Identity Service

Figure 2.6
Diagram of the SAML SSO process

passwords to access all the services provided by their institution, from e-mail and Wi-Fi to learning management resource, library account, and, yes, online resources like databases.

SAML AND FEDERATION

In the previous section, we've seen the following steps:

1. A user tries to sign in to access a resource from an SP by indicating which institution they are from.
2. The SP redirects the user to the appropriate IdP based on their selected institution.
3. The user signs in with the IdP.
4. The IdP checks whether the sign-in is correct and then redirects the user back to the SP with a trusted assertation that the user is verified.
5. The SP grants access.

But how does the SP at step 2 know the location of the appropriate IdP? The simplest answer is that the SP and the IdP have an agreement in advance, and, in practice, this type of one-to-one relationship is often employed. In a scenario where there is only one SP and one IdP, knowing to which IdP to send users is a simple matter. However, a service like JSTOR may have thousands of customers from all around the world, so maintaining lists of customers and their IdPs can get unwieldy. Similarly, the library and the institution may want to enable SAML with hundreds, if not thousands, of services. It is important to note that SAML can be used to authenticate all sorts of online resources, not just library resources.

This is where the idea of *federations* comes into play. Rather than SPs contracting directly with individual institutions and IdPs, they join federations or more precisely identity federations. SPs also joining those same identity federations results in data and standards that can be trusted by both sides without the need for individual arrangements. At a very basic level, identity federations are trusted registries where SPs and IdPs can do lookups to find metadata of institutions and organizations as well as agreed-on protocols for completing the SAML process. There are dozens of identity federations out there, including the following:

- UK Access Management Federation for Education and Research

- InCommon
- Australian Access Federation (AAF)

They are often at the national level, but global ones like OpenAthens do exist.

IS SAML A PERFECT SOLUTION?

So far, SAML SSO, with its promise of single sign-on even when the user is off campus, seems to be a better solution than IP recognition. To recap, all users have to do on any SAML SSO-supported resource site is

1. Click on the Sign In or Log In button.
2. Sign in with their standard institutional password.

Then access is granted. There is no need to struggle with proxied links or remember unique passwords.

However, there are a couple of issues with this solution. First, not every online resource supports SAML-based authentication. While this is also true for IP authentication, SAML support is still less common, particularly among smaller publishers and content owners. Second, not all libraries have experience with SAML technology, and often expertise on identity federation and identity management resides at the institutional campus IT level. This is particularly true in terms of management of the IdP server. See box 2.3 for information about implementing SAML technology.

Third, depending on how the IdP is set up and the agreements in place, SAML authentication can lead to less privacy for users compared to IP recognition methods. We will discuss this further in chapter 4. Lastly, traditionally, library databases and providers have not been very consistent in the way they signal to users that they support SAML-based authentication. Using jargon like the names OpenAthens and Shibboleth on their web pages, coupled with poor user interface experiences, tends to lead to poor user experience and low usage rates.

As we will see in chapter 4, a sign-in process where you select your institution is known as a Where-Are-You-From (WAYF) process. The WAYF process has always been a stumbling block for users. A new initiative, RA21, has risen to tackle this issue by systematically studying the problem and helping to set consistent standards.

Box 2.3

New to SAML and Federated Access?

Libraries' experience and expertise with SAML varies across regions. Traditionally, UK and to some extent US academic libraries have had the longest experience with such technologies, but not all academic libraries are equally familiar with the technology. For libraries new to SAML technology, considering a switch to this mode of access can be daunting.

Here are some general considerations when thinking of moving in this direction and things to find out. Do you have in-house expertise from people who know and understand the following?

- the basic concepts of service provider (SP), identity provider (IdP), and federated identity
- what attributes are and how they can affect privacy (See chapter 4 for details.)
- what existing identity management servers are used by the larger parent organization
- what identity federations the parent organization and prospective SPs are in

In many institutions, the library itself may have limited experience with SAML access. It may have to consult the larger parent organization, typically the university's central IT unit, which may be managing the IdP, and work closely with it on the possibility of SAML support of library resources.

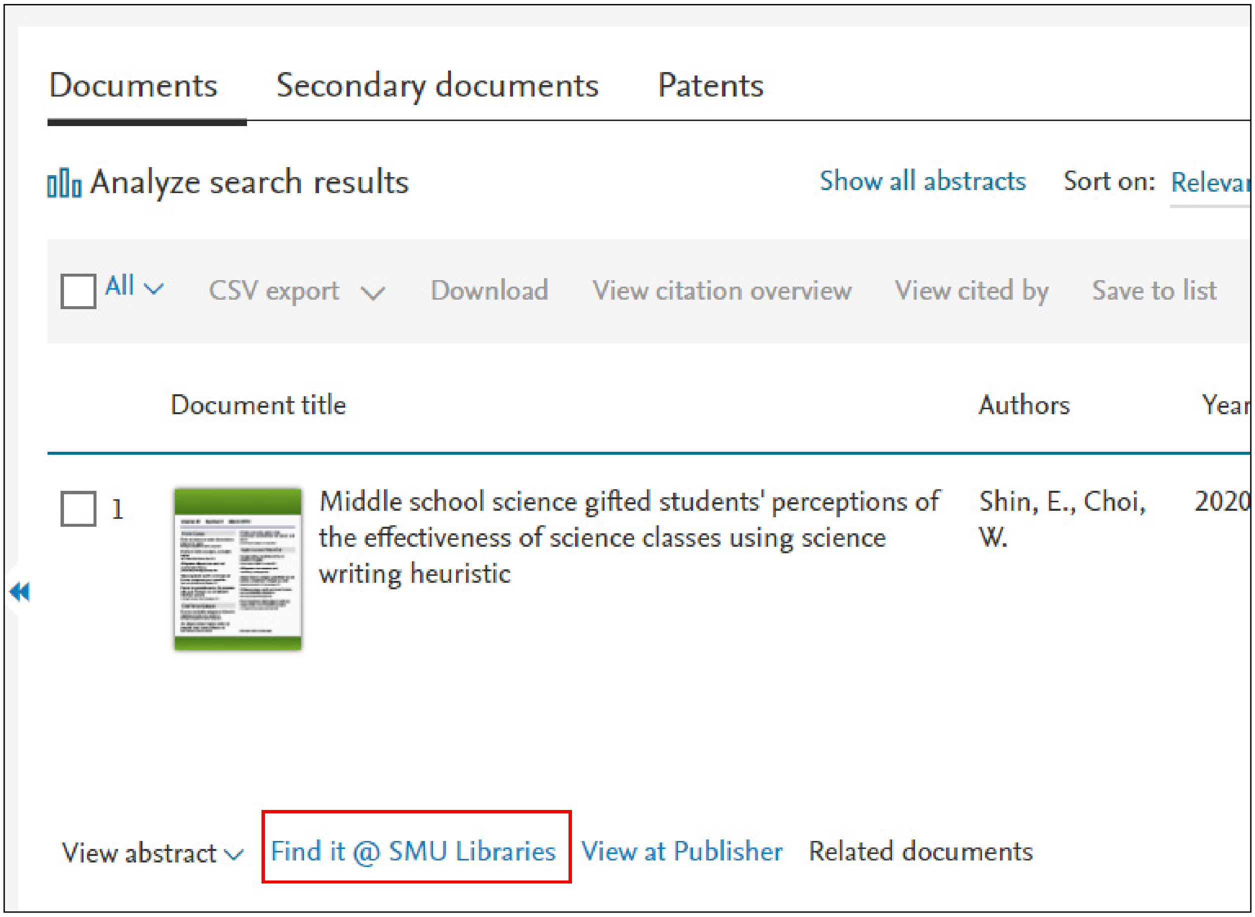
Alternatively, the library can consider running its own IdP by opting for a service such as OpenAthens.

The Appropriate Copy Problem Explained

We began this chapter with the scenario of a user landing on an article landing page in the JSTOR database and discussed how JSTOR could authenticate or authorize the user appropriately through their institution and allow the user to gain access to the full text on JSTOR past the paywalls. As discussed earlier, this is not a trivial problem if you want access to be as seamless as possible.

Even if this issue is resolved and the user is authenticated, there are further complications. Thus far, we have assumed that each requested journal article is available in only one location—the location the user is at—and all we need to do is to figure out a way to authenticate the user to determine access past the paywall. However, things can be further complicated if multiple valid copies that are appropriate for different users to access reside at multiple sites rather than just one site.

For example, while a journal article might be available on JSTOR, it might also be available on aggregator sites such as EBSCO or ProQuest or publisher sites such as Wiley. Also, open-access copies might exist in repositories. With all these options, to which copy is it appropriate to send the user?[6] How would a third-party abstract and indexing site such as, say, Web of Science



Middle school science gifted students' perceptions of the effectiveness of science classes using science writing heuristic — Shin, E., Choi, W. — 2020

Figure 2.7
Example of a link-resolver button on the SCOPUS platform

or an academic search engine such as Google Scholar know the answer of where to send users?[7] *The appropriate copy problem* was the term coined over twenty years ago to describe this issue. Given an online citation to a journal article, how should systems direct users who have different access and entitlements to the appropriate copy?[8] The solution that libraries and technologists settled upon was the OpenURL standard, which works together with identifiers such as DOIs in library link resolvers to direct users to the appropriate copy.

Today many academic platforms—including popular citation indexes, databases, and academic search engines such as Scopus, Web of Science, Google Scholar, and JSTOR—all support OpenURL and link resolvers, which provide buttons that users can click to be redirected to the appropriate copy wherever that copy may be. See figure 2.7 for an example of such a link-resolver button (in this case labeled Find it @ SMU Libraries) on the Scopus platform. Of course, sometimes no appropriate copy may be available for the user, in which case the typical academic library will display some other service, such as a document delivery service.

OpenURL Briefly Explained

A full discussion of OpenURL is beyond the scope of this text; however, it is useful to be aware of roughly how OpenURL, which is a NISO Standard (Z39.83), works. Let's assume the user has signed on to the platform via either IP authentication or SAML-based methods and the platform knows the user's institution. The idea behind platforms and databases that support OpenURL is that when a user clicks on an OpenURL request link (see figure 2.7), the request link will send information (metadata) about the item the user is requesting back to the user's institutional link resolver. The institution's link resolver will then do the work and figure out where to send the user (see figure 2.8).

More specifically, the OpenURL request typically consists of two parts. The first is the base URL, which contains the address of the user's institutional link server. This base URL may be automatically set when the user authenticates or in some cases may be selected manually by the user (e.g., in Google Scholar). The second part is the OpenURL request itself, which consists of a query, which can be understood as

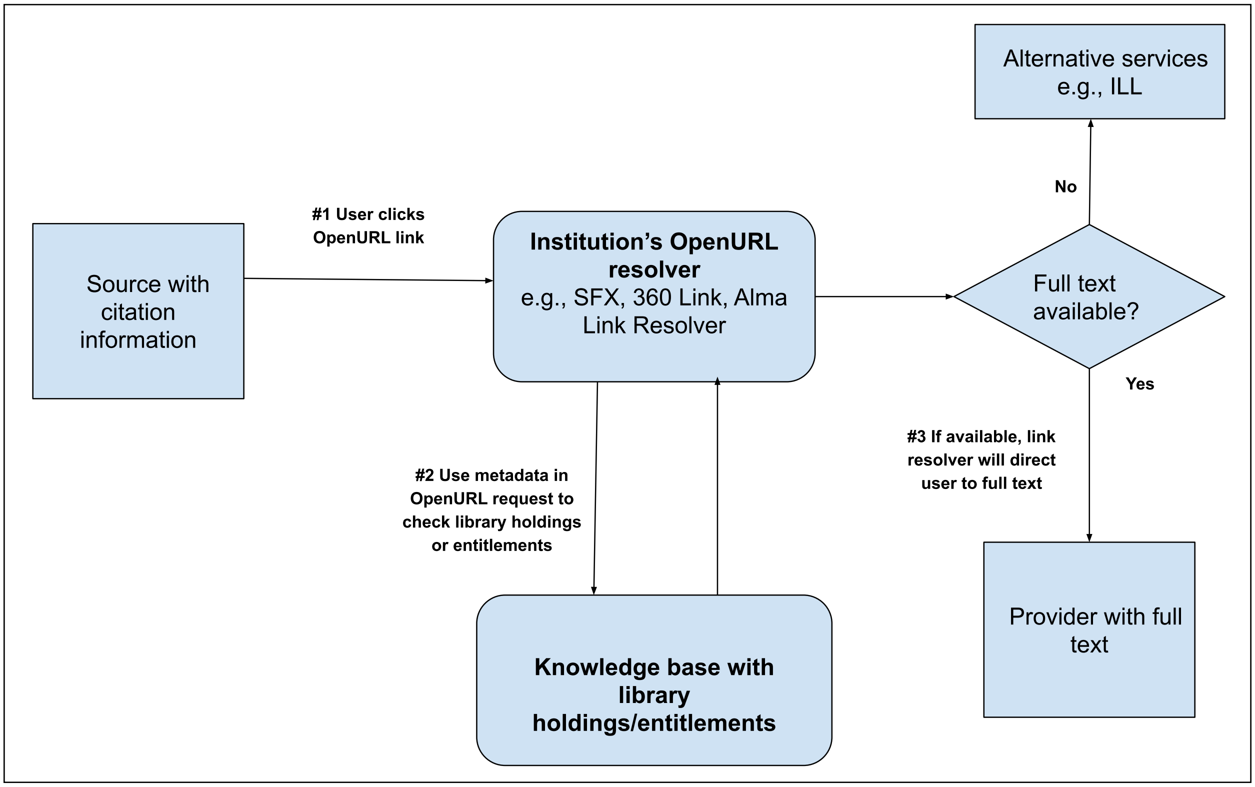

Alternative services e.g., ILL
No
#1 User clicks OpenURL link
Source with citation information
Institution's OpenURL resolver e.g., SFX, 360 Link, Alma Link Resolver
Full text available?
Yes
#3 If available, link resolver will direct user to full text
#2 Use metadata in OpenURL request to check library holdings or entitlements
Provider with full text
Knowledge base with library holdings/entitlements

Figure 2.8
Diagram of the process of using an OpenURL link

something that describes information or metadata about the requested resource. This is typically a journal article, but it can be a book, a patent, or any other resource. Here's an example of an OpenURL request for an article in the *Journal of the American Society for Information Science and Technology*:

> **https://search.library.smu.edu.sg/discovery/openurl?institution=65SMU_INST&vid=65SMU_INST:SMU_NUI**&volume=59&date=2008&aulast=Luyt&issue=2&issn=1532-2882&spage=318&id=doi:10.1002%2Fasi.20755&auinit=B&title=Journal%20of%20the%20American%20Society%20for%20Information%20Science%20and%20Technology.&atitle=Improving%20Wikipedia%27s%20accuracy:%20Is%20edit%20age%20a%20solution%3F&sid=google

The part in bold is the base URL, which sends the user to the right institutional server to check for sources. The remaining part is the metadata describing the requested resource. You may be able to make out from the OpenURL that the requested resource is something that

- is in *Journal of the American Society for Information Science and Technology*
- is in volume 59, issue 2, published in 2008
- has an author with the last name Luyt
- has the article title "Improving Wikipedia's Accuracy: Is Edit Age a Solution?"

The OpenURL standard provides standards on what metadata fields can be used in the OpenURL request: for example, ISSN, volume, issue, starting page, and so on.

Once the user is directed to the appropriate institutional link resolver, the link resolver will use the metadata of the requested item to check the institution's knowledge base (e.g., Alma) to figure out whether the institution has access to that resource and if so try to figure out where to send the user. Using the metadata provided in the OpenURL request, the institutional link resolver will construct a link to the resource. This resolved link could be a link to the publisher, the aggregator, or an open-access copy.

It is important to note that such a process is not magic. For the link resolver to work reliably, the knowledge bases, which contain information on the entitlements of the institution, need to be updated faithfully. Erroneously leaving entitlements out of the knowledge base will lead to the link resolver wrongly indicating something is not available. Doing the opposite will mislead the user into thinking they have access, and they will get an Access Denied message when directed to the requested resource.

Box 2.4

Are Library Link Resolvers Using Only OpenURL?

OpenURL is a NISO Standard (Z39.83) that provides a way to describe a desired scholarly item in a URL using standard metadata fields such as ISSN, volume, issue, start page, and so on.

However, the OpenURL process consists of two processes. First, there is "inbound OpenURL," which uses OpenURL to provide information or metadata on the item the user is trying to access when they click on the link-resolver button on the database platform. Once this metadata is received by the link resolver and checked against the institution's knowledge base, assuming the institution has access to full text somewhere, OpenURL can be used to construct or create a link using a predefined link syntax (which needs to be regularly updated) to the appropriate full-text provider. This second process of creating the link to full text is known as "outbound OpenURL" linking. As mentioned in the text, the process of creating outbound OpenURL links often leads to unreliable linking due to issues like metadata inaccuracies and mismatches.

Therefore, modern library link resolvers incorporate other forms of linking that are not OpenURL-related (though the process can be triggered by inbound OpenURL requests). These other linking methods often rely on article or proprietary identifiers or specialized link syntaxes that are not OpenURL-supported. As a result of using such identifiers, these links are often more reliable than links generated from OpenURL, which requires less stable metadata such as title, journal volume, and page number to match. However, such linking methods are often available only for specific full-text providers as they require the link-resolver vendor to have special arrangements with individual full-text providers.

We will see alternative linking mechanisms such as LibKey in chapters 3 and 4.

Even if the knowledge bases are updated with the right entitlements, links provided via OpenURL might still break. There are many reasons for this, but a common reason is due to errors in the metadata provided in the OpenURL request.[9] For example, the page number or author in an OpenURL request from a platform might be slightly off and thus lead to a wrong link being generated.

It is important to note that, while traditional link resolvers use only OpenURL technology for linking, modern library link resolvers also use other methods to generate links. For more detail, refer to box 2.4.

Is OpenURL a Perfect Solution?

While OpenURL has been a standard in use for over two decades, there have been a variety of problems. Over two decades of research has shown that OpenURL linking tends to be fairly unreliable even for journal articles (which have the highest reliability of all types).[10] There are many reasons for unreliability, such as metadata mismatch inaccuracies, different granularity of linking at the target and source, and the already mentioned inaccuracy of entitlements or holdings data in the knowledge base. In future chapters, we will see how GetFTR and some access broker browser extensions such as LibKey Nomad claim to provide improvements to these issues. More recently, Bulock argued that researchers today work increasingly in open web contexts, which leads them to citations on web pages that either do not support OpenURL or where they are unable to indicate their institutional context.[11] Both access broker browser extensions (covered in chapter 3) and GetFTR (covered in chapter 4) provide some improvements to this problem.

Conclusion

In this chapter, we have presented a high-level view on the issues around providing access via authentication and authorization. We introduced the issues of giving individual accounts to users and outlined two main solutions to these issues: IP recognition and SAML SSO methods.

We also briefly described a further issue with delivery, the appropriate copy problem, and discussed how OpenURL and library link resolvers traditionally handle this problem. In the next chapter, we will discuss access broker browser extensions, which attempt to improve on some of the weaknesses presented here.

Notes

1. Oren Beit-Arie, Miriam Blake, Priscilla Caplan, Dale Flecker, Tim Ingoldsby, Laurence W. Lannom, William H. Mischo, et al., "Linking to the Appropriate Copy: Report of a DOI-Based Prototype," *D-Lib Magazine* 7, no. 9 (September 2001), https://doi.org/10.1045/september2001-caplan.
2. Platforms that attempt to help with the appropriate copy problem are not restricted to discovery systems. Some reference managers, academic social networks, and even some blogs help with this problem by supporting ContextObjects in Spans (COinS). Even

publishers and other platforms that carry content—e.g., JSTOR and EBSCO—also provide alternatives to the user by supporting link resolvers.

3. Kristina Botyriute, *Access to Online Resources: A Guide for the Modern Librarian* (Cham, Switzerland: Springer International Publishing, 2018), https://doi.org/10.1007/978-3-319-73990-8.
4. This is the proxied link for the general JSTOR home page. Specific URLs within JSTOR (e.g., URLs for journal articles) will need to be similarly proxied. In general, once the user is on a proxied URL, all other clicks to links on the same domain will continue to be proxied. It is also important to note that institutions' entitlements may be only a subset of articles on JSTOR and that not all content is fully accessible.
5. Such technologies where you sign in with social accounts like Google or Twitter are usually based on OAuth, which is a different protocol from SAML and differs in some functional ways. For example, while both support SSO, OAuth is a narrower standard that focuses only on authorization, not authentication. Currently, OAuth is not commonly used in the academic space.
6. Here we assume all the different copies are identical; in practice, open-access copies might be different versions of the article—e.g., an accepted manuscript, the version of record, or even a preprint—which further complicates things.
7. While the appropriate copy problem is most salient for discovery platforms and other third-party sites, content platforms such as JSTOR and Wiley do often provide solutions for users who may have access elsewhere.
8. Beit-Arie et al., "Linking to the Appropriate Copy"; Herbert Van de Sompel and Oren Beit-Arie, "Open Linking in the Scholarly Information Environment Using the OpenURL Framework," *D-Lib Magazine* 7, no. 3 (March 2001), https://doi.org/10.1045/march2001-vandesompel.
9. Ex Libris Knowledge Centre, "How Does Incorrect Metadata Break OpenURL Linking," November 1, 2019, https://knowledge.exlibrisgroup.com/Primo/Content_Corner/Primo_Central_Index/Knowledge_Articles/How_does_incorrect_metadata_break_OpenURL_linking.
10. Cindi Trainor and Jason Price, *Rethinking Library Linking: Breathing New Life into OpenURL* (Chicago: American Library Association, 2010).
11. Chris Bulock, "Get Full Text Research and the Search for Appropriate Copies," *Serials Review* 46, no. 2 (2020): 160–62, https://doi.org/10.1080/00987913.2020.1759361.