UC Libraries Shibboleth Task Force: Improving Online Access Management through Shibboleth

September 10, 2010 [Revised 9/20/2010; Accepted by Systemwide Operations and Planning Group 9/16/2010]

Task Force Members:

John Ober (Chair, CDL/LTAG), Declan Fleming (UCSD/LTAG), Ann Frenkel (UCR/HOPS), Julia Kochi (UCSF/SOPAG), Eric Scott (UCM), Dan Suchy (UCSD), Terry Toy (UCR/LTAG)

Contents

Revision history: 9/20/2010 – added UCR to list of campus IdPs who can immediately comply with attribute Release Policy preferred by HathiTrust in recommendation 4; added UCD responses from UCTrust survey to Appendix 5.

I. Introduction

In its charge to the UC Libraries Shibboleth Task Force, SOPAG requested two products, a cost benefit analysis and an action plan.[1] This report presents action plan recommendations and rationale, comprising policy, technology, and related "paths" for the adoption of Shibboleth within the UC library community and its partners. Together with the cost-benefit analysis (submitted in April 2010 as an interim report on progress and included here as appendix 1), this document serves as the final report of Task Force efforts. The reader is advised to familiarize him/herself with that first report as necessary background or prelude to the additions made here.

The Task Force believes that Shibboleth can and will yield significant service improvements and has the potential to create operational cost savings for the UC libraries at both the collective and separate/local levels. As detailed below, we recommend that the UC libraries endorse and adopt Shibboleth as the primary means of online authentication to library owned and sponsored content and services. We further recommend several specific steps and activities and, where possible, provide departure points for additional analysis and to guide implementation.

II. Recommendations

1. **Our overarching recommendation is that the UC libraries build the knowledge, expertise, policies, and practices to use Shibboleth authentication wherever and whenever possible** in the provision of library-related content and services where the authentication of a user is required.

   In the case of content and services not directly under library control this recommendation effectively implies a) that the UC libraries strongly encourage 3rd party service and content providers to adopt Shibboleth authentication, and b) that the UC libraries develop the appropriate expertise and relationships to sponsor 3rd parties as registered requestors and recipients of UC-campus identity information, i.e. that each library become able to sponsor or broker a 3rd party as a Service Provider (SP) recognized by each campus Identity Provider (IdP).

   In the case of content and services directly controlled by the libraries themselves, this recommendation effectively suggests a) that the libraries collectively and separately create or support Shibboleth authentication components for locally run applications, and b) that the libraries define themselves as SPs (or, more accurately, that each library-run application that requires authentication be registered as a Service Provider) recognized by the local or appropriate Identity Provider(s).

   A high-level description of the costs and benefits of the use of Shibboleth is provided in Appendix 1. In addition to that analysis, the task force has discovered and queried several U.S. research libraries that have already initiated wide-scale adoption of Shibboleth (see Appendix 2), and provides a list of content vendors supporting Shibboleth authentication (see Appendix 3; note that there are entrants for whom Shibboleth is the only method of authentication available).

2. **We recommend that the UC libraries collectively endorse or participate in the recruitment of more library and content vendors into the U.S.-based InCommon Federation and the use of Shibboleth**.

[1] http://libraries.universityofcalifornia.edu/sopag/ShibbolethTF/Shibboleth_TF.pdf

While the list in Appendix 3 already shows a critical mass of content vendors supporting Shibboleth, the InCommon Library Services Collaboration has actively recruited additional vendors as recently as Winter 2009-2010 (see https://spaces.internet2.edu/display/inclibrary/TargetResources) . Prompted by this task force, several UC libraries became signatories to recruitment letters in November 2009. Given recommendation #1 above, the task force recommends that at least this level of support from UC continue, presumably through the InCommon-Library effort.

The UC agent charged to manage the mechanisms through which support is offered could be a component of LTAG's Shibboleth role as proposed in section III.B.5 - Communication.

We further suggest that the CDL and campus libraries should add a Vendor's support/adoption of Shibboleth as a principle for negotiations for tiers 1, 2, and 3 content, and strengthen the preference for Shibboleth in the Tier 1 "Technical Guidelines for Vendors" materials maintained by the CDL.[2]

3. **We recommend that UC and the UC Libraries follow the InCommon-Library best practices when possible**. The InCommon Library Services Collaboration in December 2009 finalized a set of best practices for service providers and identity providers relevant to the academic library community.[3] The best practices are a necessary complement to the InCommon-sanctioned Shibboleth protocols and software. We believe that full compliance with the best practices - compliance by any one organization as well as adoption by the entire community – is likely to progress stepwise over time and that the best practices themselves may also evolve. As shown in Appendix 7 (UCTrust Survey questions 5-8) we confirmed that almost all reporting UC campus identity managers can currently comply with core aspects of the best practices and are able or willing to comply with the other components when a request to do so is made explicit.

4. **We recommend that the UC Libraries collectively declare a Proof of Concept pilot of Shibboleth authentication for the UC community's use of HathiTrust services**, ideally to be initiated during Fall term 2010. In practical terms this will require each UC Library to negotiate with their campus identity manager to register and sponsor the HathiTrust as a Service Provider. It may also encourage the UCTrust working group to consider library requirements collectively and, in partnership with the libraries, to create a set of more or less formal "UC IdP library conventions and best practices." As described below, the TF has prepared the ground for this test, and at least three campuses – UCLA, UCR, and UCSD – have Shibboleth identity management practices that could support it immediately.

5. We confirm the preliminary recommendation reached in our Interim Report (see Appendix 1) that WAYFless URLs are problematic and expensive to pursue as the default for the UC library community, except for those libraries that can also follow the "EZProxy" solution outlined in the InCommon-Library best practices. Therefore **we recommend that in implementing Shibboleth the UC Libraries minimize resource cataloging changes needed for so-called "WAYFless" URLs, and instead adjust to the two main methods through which Shibboleth authentication can establish the source of appropriate identity information when users are not authenticated via IP address filtering**:
    a. when attempting to reach a "Shibbolized" resource the user encounters a "Where Are You From?" dialog box, either based on the InCommon WAYF or a Service Provider's own WAYF (after following the SP's specially constructed resource URL);

[2] See "Technical Requirements for database vendors" and "Technical requirements for eJournal vendors" at http://www.cdlib.org/gateways/vendors/guidelines_technical.html.
[3] See https://spaces.internet2.edu/display/inclibrary/Best+Practices

b. based on an earlier interaction with the InCommon WAYF, a cookie exists with information about where the user is from that can be passed on to the SP and a second "Where Are You From?" interaction is avoided.

III. Implementation Considerations

Much of the following information is drawn from in-depth discussions between the TF and the UCTrust Working Group. In those discussions the TF discovered that the UC campus identity practices are highly variable. In UC's case the common practices that InCommon/Shibboleth encourages or enables have been adopted for a limited number of Systemwide services. The identity management and release policies required of those services do not comprehensively overlap with the policies and practice needed to comply with the InCommon-Library best practices. Therefore campus identity managers are not uniformly ready to meet the Identity Provider commitments emerging in the Library community. However, of the campus identity managers who responded to the TF-UCTrust survey, all but Santa Barbara are willing to develop the necessary functional requirements. UC Santa Barbara is not yet a member of InCommon.

A. Probable implementation "stages"

As described in the TF charge and in the Interim Report, Shibboleth adoption and implementation is underway throughout the higher education and academic library community. However, there will be a lot of localized and environmental complexity in the migration from a mostly working but rigid and inefficient authentication regime built on location (IP-based authentication) to Shibboleth's more flexible credential (identity)-based methods. The complexity resides in at least four places: 1) the adoption of new infrastructure and policies by Service Providers; 2) the adoption of new infrastructure and policies by libraries and their concomitant need for deeper interactions with campus identity managers; 3) the libraries' dual-role as direct Service Provider and indirect sponsor of other Service Providers; 4) the UC context and its more or less formal pursuit of system-wide versus local action and policy (often formal in the libraries case with CDL acting as a host or broker for central service provision, but informal in the UCTrust case as an advisory body to the IT Leadership Council).

The TF therefore envisions an evolution toward an environment completely based on Shibboleth authentication, rather than a well-bounded or one-time conversion. While it cannot predict the timing of that evolution, several stages, created by or subject to influence by the UC libraries, can be suggested. In roughly chronologic order they include:

Phase 1: The libraries endorse and begin a proof of concept pilot for Shibboleth-based authentication to use the HathiTrust "Collection Builder" service and to download a full public domain book (as opposed to page-at-a-time use. HathiTrust (HT) requires a relatively simple Shibboleth interaction between HathiTrust as the SP and each IdP. When each library successfully sponsors HT and the IdP development work, if any, is complete, the use of these exclusive features, for both UC and non-UC content held by HT will be an immediate benefit to the UC community. As a relatively new resource, the addition of Shibboleth authentication should provide a good test case for public service and education issues too.

Phase 2: Drawing on the HT pilot, library issues are adopted as a central component of UCTrust activities. The TF has begun this process but it is not complete; a UCTrust recommendation for collective endorsement/adoption of the InCommon-Library IdP best practices would be a marker of significant progress in this stage.

Phase 3: UC libraries migrate to Shibboleth authentication for Tier 1 resources supplied by vendors that are already InCommon members and are Shib-ready (see Appendix 3). This evolutionary stage will

require most of the implementation components listed below, particularly maturation of the libraries' "sponsorship" role wherein the library sponsors a Service Provider to receive identity attributes from campus IdPs.

Phase 4: CDL-hosted resources and services convert to Shibboleth authentication. The CDL has already begun analyzing the implementation of Shibboleth for its major services (the Consortial Borrowing Service, eScholarship, the Digital Preservation Repository, Melvyl, etc). This evolutionary stage will require most of the implementation components listed below, particularly maturation of the libraries' "sponsorship" role wherein the library sponsors a Service Provider – in this case the CDL service - to receive identity attributes from campus IdPs.

Phase 5: Libraries add Shibboleth authentication for locally-hosted library applications, presumably for their own campus community. Although we think it likely that some campuses will wait for a critical mass of 3rd party services/content to be Shibbolized before attempting to convert their own applications, if any, there is nothing to prevent them from doing so at any time. Simplified Shibboleth local installation guidelines are provided in Appendix 7.

Phase 6: Remaining content vendors migrate their services to Shibboleth followed by library sponsorship with campus IdPs. We anticipate a "long tail" of mostly smaller content vendors who will delay adoption of Shibboleth (and InCommon membership) until it is clear that there is a business advantage in adoption (or a disadvantage, i.e. the loss of customers, in delay). The TF believes that comprehensive adoption of Shibboleth in the library vendor community is likely but is unable to predict how long it may take. In part it will depend upon the success of lobbying vendors as described in recommendation #2.

Phase 7: Enhancement of Shibboleth to support fine-grained authorization. For most library services, especially tier 1 content, authentication – providing proof that you are a (faculty, student, or staff) member at a UC campus – is all that is needed for the vendor to authorize your use of the resource because the entire class of users is authorized through the license between the University and the vendor. Authorization of specific individuals to use specific functions can be facilitated by Shibboleth in at least two ways:

1. Authentication is performed in a way that a personal identifier (e.g. "EduPersonPrincipalName") is also released and the service/application checks a local authorization table to see what privileges or personalized services the authenticated individual is allowed to use;
2. Authentication is performed in a way that includes the release of one or more affiliation or entitlement values; services are made available to any authenticated user with that affiliation (e.g. member of Physics department) or entitlement (e.g. entitled to "interlibrary loan").

The InCommon-Library working group anticipated the need for finer-grained authorization by recommending a best practice in the use of the EduPersonEntitlement value and a generic value of "common-lib-terms." For example, if the library wanted to deny access to content to a student with excessive overdue materials, the library could provide data to the campus identity manager updating the entitlement attribute, removing "common-lib-terms" as the entitlement value. If the vendor followed the best practice of basing access on the presence of "common-lib-terms" in the entitlement field, and found "common-lib-terms" missing, then access would be denied. Similar fine-grained access could be provided, presumably, if an affiliation attribute could be established and released, for example to allow access to a Physics resource licensed only for those who were members of the Physics department, or to allow interlibrary loan librarians to change records in the Consortial Borrowing System.

While Shibboleth includes the mechanisms for these scenarios, our survey of the literature and of UCTrust principals suggests that additional practical experience using Shibboleth is required, and

negotiation conventions for attribute values and attribute release policies are needed before these sorts of Shibboleth-based authorizations are likely.

B. Implementation components

1. Sponsorship of a Service Provider

Campus identity providers take several steps to accept, approve, and implement the release of identity information to a Service Provider. The library may request that identity information be released to one of its own applications or to a 3rd party application such as a content vendor's website. According to our research, UC identity managers are beginning to formalize the process for requests to add additional service providers to whom identity information should be released (to date, only a few UC systemwide services, such as the "At Your Service" employment and benefits application, are common across UC). Sponsorship will be eased significantly if the vendor is already a member of InCommon; otherwise registration of certain metadata about the vendor within the Shibboleth Identity Provider software run by the campus identity manager is required (metadata for members of InCommon is regularly and automatically loaded into IdP databases).

When common or extant attributes are requested to be released, implementing a new SP will be straightforward (a principle motivation for the InCommon and InCommon-Library best practices). UCTrust principals estimated from 24 hours to 2 weeks to complete in such cases. When attributes are needed that don't yet exist in the campus identity management system, implementation can take much longer.

Question four in our UCTrust survey (see Appendix 5) reports the request process in place at each reporting campus as of this writing. The TF speculates that local conditions might suggest an introductory meeting between the identity management team and the library team to discuss the number, types, and circumstances of imminent SP sponsorship by the library.

2. Technical Considerations

Although "sponsorship" of 3rd party SPs is likely to be the bulk of the foreseeable work for library implementation of Shibboleth, the TF considered three technical issues that may arise, and provides the following information and guidance for them

1. Use of Shibboleth for locally hosted applications.

a. Locally developed software may be configured to include Shibboleth authentication. For this situation the TF has created Appendix 6 – "Simplified Shibboleth Local Installation Guidelines." As noted in the Appendix, modules for "Shibbolizing" a service/application are available for common operating systems and a rich set of implementation help is available from the Internet2 middleware initiative which sponsors Shibboleth development. The TF estimates one to two weeks of mid-level developer time would be sufficient for initial deployment (assuming identity attributes are available and have been successfully requested for release from campus identity managers).

b. Commercial software hosted by the library may have a Shibboleth component that can be "turned on." Alternatively, the library may be able to request Shibboleth authentication features for the product's development priorities. We noted in our Interim Report (Appendix 1) a generally favorable situation with regards to the current use of or compatibility with Shibboleth by the most heavily-used 3rd party library software vendors and products. In most cases Shibboleth compatibility is included in the current release of the software (e.g. SFX, ContentDM, VDX, Moodle, Blackboard). This is not true in the notable case of the Innovative Interfaces Inc.

Millennium ILS product which requires further investment in their Single Sign On package, a front end reverse proxy server, and significant configuration effort.

2. Shibboleth for walk-in or public access workstations.

Shibboleth 2 has implemented a way to handle location-based authentication which allows the release of pre-determined identity values for a known IP address or block of addresses. This is accomplished with the Shibboleth authentication "handler" which assigns a time-limited affiliation of "library-walk-in" for any authentication request from a session originating from one of the pre-registered IPs (see https://spaces.internet2.edu/display/SHIB2/IdPAuthIP). Most UCTrust members indicated that they would be able to support this feature if the library were able to supply a list or range of workstation addresses used for "walk-in" use.

3. Use of EZProxy.

The role of EZ Proxy as a way to seamlessly handle a mixed environment of IP-based and Shibboleth-based authentication is a large component of the InCommon-Library working group's charge and was discussed in our Interim Report (Appendix 1). At this point EZProxy integration with Shibboleth has been successfully tested or is in production at a number of libraries (including UCSD in early pilots; also see Appendix 2). For those UC libraries who currently have or plan to adopt EZProxy, Shibboleth integration should be straightforward at this point. However, as mentioned in our Interim Report, proxy use has its own complications, including a requirement for users to be directed through a library-maintained list of resource URLs that can be rewritten for IP-based or Shibboleth-based authentication. Therefore the TF chose not to recommend the adoption of EZProxy for libraries/campuses that do not currently use it.

3. Policy

The TF speculates that an overarching policy statement could prove useful to strengthen, unify, and promulgate the UC lIbraries' commitment to Shibboleth, should that commitment be endorsed by appropriate UC Library principals and groups. An overarching policy statement could also be referenced in specific domains, e.g. in licensing guidelines, public service declarations, etc. Accordingly the TF offers the following draft policy statement:

> *Libraries of the University of California have adopted Shibboleth as the primary authentication standard for access to our research resources and services. In addition to providing a better experience for our users through the use of a single username and password, Shibboleth simplifies secure authentication management and builds stronger partnerships between UC, its vendors, and the wider academic community.*
>
> *Whenever possible and appropriate, the UC Libraries will implement Shibboleth for the following:*
>
> - *Resources licensed from external vendors*
> - *Resources created and hosted at the California Digital Library or an individual UC Library.*
> - *Internal systems (content management systems, staff portals, etc)*
>
> *To contact us and for more information concerning the UC Libraries' use of Shibboleth, please visit: http://someaddress.edu. For more information on Shibboleth and how to get involved, please visit http://shibboleth.internet2.edu/.*

Appendix 7 provides a sample of a domain policy statement that could reference or be drawn from such an overall policy, using the "CDL Technical Requirements for E-Journal Vendors."

4. Education/instruction/outreach

The libraries will face at least two dimensions of education and outreach in their separate or collective use of Shibboleth – for library patrons and for library staff (both in their public service roles and, eventually, as users themselves of Shibbolized library applications).

1. End-users

In an eventual "fully Shibbolized" world of university resources (including course management systems, library services, payroll systems, etc.) end-users will not only become accustomed to, but expect the use of their single-sign on ID and password to gain access to all services. When using those services from off-campus, or when using an externally-provided service at the beginning of any online session (and provided that the user is not directed through a library-maintained gateway page with proxy-authentication built-in), users will encounter a "Where Are You From?" challenge and will need to learn how to navigate to the appropriate choice that informs the Service Provider which institutional Identity Provider is able to authenticate the user and pass back the user's identity attributes to begin the session.

Until that time, and assuming a hybrid authentication world where access to library services is granted by location (i.e. IP-filtering for on-campus or VPN use), and/or proxy-service, and/or Shibboleth, the library will need to provide at least the following types of outreach and instruction:

1. Announcement that "single-sign on" and Shibboleth-based authentication is an additional access method for some or many resources;
2. Description of and instruction for responding to a "Where Are You From" challenge;
3. Shibboleth Authentication component of specific electronic resources (which may include or reference information from either or both above).

Fortunately, there are numerous models of all three types available, particularly from the UK whose academic libraries have developed materials to support the JISC-mandated move to Shibboleth authentication. Links to samples of all three types from U.S. and UK institutions are included in Appendix 8.

We note that the CDL maintains a list of UC-library "off-campus access" pages at http://www.cdlib.org/services/info_services/guides/off_campus_access.html and that HOPS is in the best position to advise further on these issues.

2. Library staff

Materials available to library IT staff for Shibboleth implementation are mentioned above. Other library staff, including those public service staff who will author end-user guides or answer end-user access questions, may want a foundational understanding of Shibboleth. We point to the description of Shibboleth architecture in Appendix 1 as one source for this information and also to the list of links for library staff we have assembled in Appendix 8.

5. Communication

In order to share expertise and adhere to established best practices, Shibboleth implementation will not only require close communication amongst the individual UC Libraries, but also between the UC Libraries, UCTrust, and InCommon (the InCommon-Library Task Force in particular).

The TF recommends that LTAG play a role in coordinating this communication. LTAG is an established representational group, and its members are well-positioned to communicate about common technology issues and to advise on systemwide technology approaches. We assume the exact

mechanism would be determined by LTAG, but one possible approach would be to have an LTAG member(s) designated as the Libraries' liaison(s) to UCTrust and the InCommon Library Task Force. Alternatively, the TF speculates that LTAG might want to propose a Shibboleth "Common Interest Group" which tracks and facilitates both inter-library communication and external library communication.

Appendix 1 – Interim Report Describing Costs and Benefits of Shibboleth Implementation

UC Libraries Shibboleth Task Force: Improving Online Access Management through Shibboleth

Interim Report
April 9, 2010

Contents

Task Force Members:

John Ober (Chair, CDL/LTAG), Declan Fleming (UCSD/LTAG), Ann Frenkel (UCR/HOPS), Julia Kochi (UCSF/SOPAG), Eric Scott (UCM), Dan Suchy (UCSD) Terry Toy (UCR/LTAG)

I. Introduction

In its charge to the UC Libraries Online Access Management Task Force, SOPAG requested two products, a cost benefit analysis and an action plan.[4] This report comprises the cost-benefit analysis, specifically "[An assessment of] the hypothesis that Shibboleth can/will yield service improvements and operational cost savings, and [a prioritization of] UC Library Shibboleth goals. [It also includes] an assessment of the "transition hurdles and costs," assuming a mixed authentication environment until most or all authentication is based upon Shibboleth."

II. Summary

Through bi-weekly meetings between 12/8/2009 and 4/5/2010, the task force (TF) has:

- Determined that there would be real costs and real benefits – both tangible and intangible - if the UC libraries were to pursue Shibboleth as the default authentication method for local, shared, and systemwide services that they host or broker for the UC community.
- Costs are primarily staff/labor costs (for technical implementation, end-user instructional materials, and, potentially, URL/link conversion; there will also be some direct costs associated with 3rd party software (notably III's ILS, and new EZproxy implementations).
- Benefits are bimodal: 1) improvements in end-user services related to single sign-on (these benefits are difficult to measure, i.e. are intangible); 2) staff/labor savings from a) an hypothesized reduction of staff responses to end-user access problems; and b) a savings in maintaining the infrastructure associated with current authorization methods.

The cost-benefit analysis is slightly different for three use cases, where shibboleth replaces current authorization and access management for:

1. Licensed materials or services;
2. Services based on library-hosted commercial software (such as ILS, course management, UC-eLinks, etc);
3. Services based on library developed and hosted software.

In the first case libraries act as a kind of broker and work with the campus Identity Provider (IdP) to register a commercial vendor as a "Service Provider" (SP) who is allowed to receive information (identity attributes) about the campus users of a service. The most basic form of identity information that is received is that the user is a legitimate member of the campus community (i.e. no personal information need be distributed).

In cases #2 and #3 the library or the CDL is a Service Provider itself, with responsibility for implementing authentication as well as registering the service with the campus Identity Provider (or separately with all 11 campus Identity Providers in the case of systemwide services).

Although a total conversion to Shibboleth-based authentication for library services (at both local and systemwide levels) will incur substantial conversion costs, the TF believes:

[4] http://libraries.universityofcalifornia.edu/sopag/ShibbolethTF/Shibboleth_TF.pdf

- Conversion is warranted and in many ways is inevitable, given current trends and expectations (see section IV.3 Benefits).
- Benefits can accrue quickly and the cost trajectory flattened and controlled by prioritizing conversion for the use cases in roughly the same order as above.
- The largest single cost will be the labor associated with eliminating the "Where Are You From?" step of Shibboleth-based authentication (see section III, Shibboleth architecture) because it requires a change to cataloging practices, cataloging records and/or databases that include persistent URLs. However, URL conversion appears to be optional if the UC libraries are willing to have users respond to a "Where Are You From" (WAYF) inquiry during authentication, rather than mask that step for them wherever and whenever possible. The TF action plan, forthcoming, will need to draw out this distinction and the related implications carefully.
- Even given a decision to pursue Shibboleth aggressively, library users at UC (and elsewhere) will continue to operate for some time in an environment with mixed authentication methods. This will, at the least, be due to content vendors who are slow to convert to Shibboleth and to the relative difficulty to convert legacy applications even as new software and updates to current vendor software include Shibboleth.

The following benefits and costs are explained in sections IV and V.

Benefits	
High (intangible)	Create a stable, predictable user experience (alleviate current user frustration and inconvenience)
Moderate(tangible)	Public service staff time saved by reducing the number and variety of user inquiries about online authentication.
Low to moderate (tangible)	IT staff time saved by reducing the number and variety of authentication methods needing local configuration.
Low to moderate (tangible)	Library staff time saved by reducing and eventually eliminating IP table maintenance and related communication with content vendors.
High (intangible)	Enhance UC's reputation through participation and leadership in deploying contemporary information services; leverage emerging higher ed authentication experience and policy decisions (to avoid costs related to independent actions/policies).
Costs	
Low	Increase the number of vendors adopting Shibboleth as the primary authentication method by participating in InC-Library (InCommon's Library Task Force) efforts.
High	($20-$30K per campus) Add Shibboleth/single-sign-on functionality to III ILS.
Low	Use Shibboleth components in most 3rd party library software (primarily staff/labor for configuration and maintenance and interaction with campus Identity Providers to register a new service).
Moderate	Integrate Shibboleth components into local software (primarily staff/labor for configuration and maintenance and interaction with campus Identity Providers to register a new service).

[If WAYFless URLs are deemed essential]	
Moderate	Purchase, configure, and maintain EZ Proxy if a decision was made to leverage its rewriting function to produce WAYFless URLs on the fly.
High	Construct and maintain WAYFless URLs and substitute them for current URLs in all discovery systems if EZproxy were not used to generate them on the fly.

III. Shibboleth Architecture

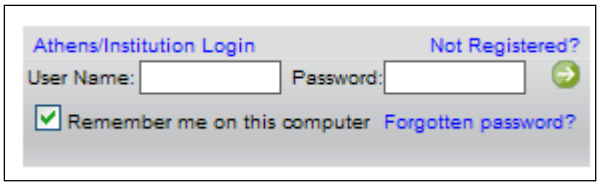
1. Overview

> Shibboleth is an Internet2 Middleware Initiative project that has created an architecture and open-source implementation for federated identity-based authentication and authorization infrastructure based on Security Assertion Markup Language (SAML). Federated identity allows for information about users in one security domain to be provided to other organizations in a federation. This allows for cross-domain single sign-on and removes the need for content providers to maintain user names and passwords. Identity providers (IdPs) supply user information, while service providers (SPs) consume this information and get access to secure content.[5]

That is the technical explanation. In practice, Shibboleth helps when a user wishes to access a resource. In the following example, we will use ScienceDirect, http://www.sciencedirect.com/science. We will assume the user is off campus, with no VPN or Proxy setup, and has navigated to ScienceDirect on her own, probably as the result of a regular Google search.

Note the text "You have Guest access to ScienceDirect" at the top of the page, indicating that the user is a generic visitor to the page with no special access. Also note the Login box at the top right of the screen:

[5] Wikipedia. Shibboleth (Internet2). http://en.wikipedia.org/wiki/Shibboleth_(Internet2)

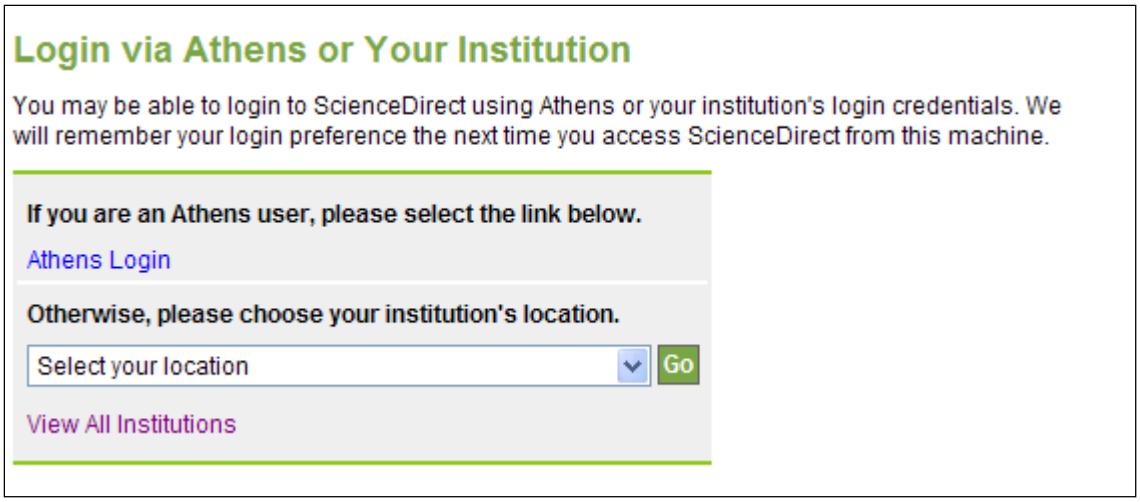


The “Athens/Institutional Login” link is what the user would select to gain access to the resource as if she were on campus.

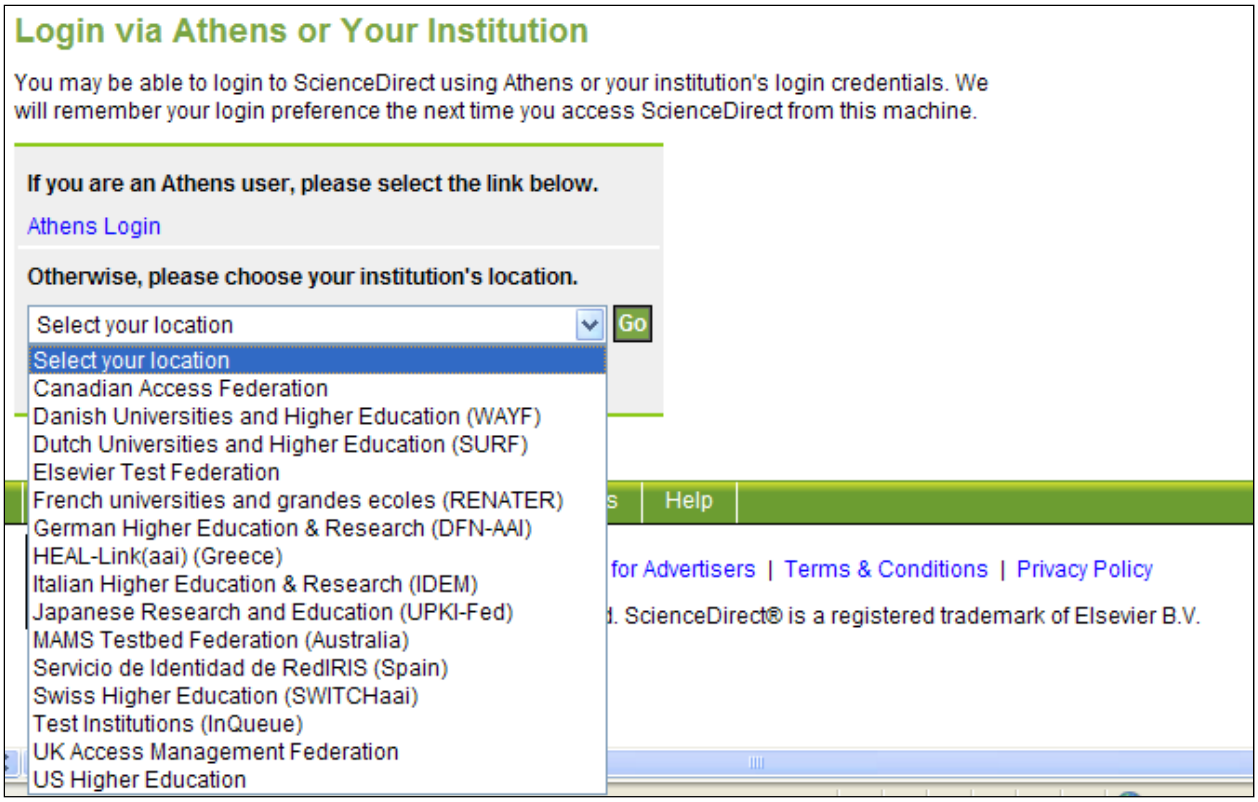
That last phrase, “as if she were on campus” deserves a little unpacking because it hides a few assumptions. VPNs and Proxies, another common way for off-campus users to get to resources, reroute a user’s internet traffic and change the originating IP address to fool the resource vendor’s server into believing that the user is on campus, coming from a range of IP addresses that have been previously blessed by previous arrangement between the vendor and each campus. Most VPNs and Proxies require a client side program and nontrivial configuration to work properly, hence the high rate of help calls to get them going.

A Shibbolized web site, on the other hand, makes no assumptions about where the user is coming from. It allows access from anywhere on the internet and directs the user to identify where she is from. Further, a Shibbolized web site needs no specific knowledge about the user, because once the user’s institution is selected from a list, the login and password management is handed off to her institution’s Single Sign On infrastructure, as shown in the four steps below.

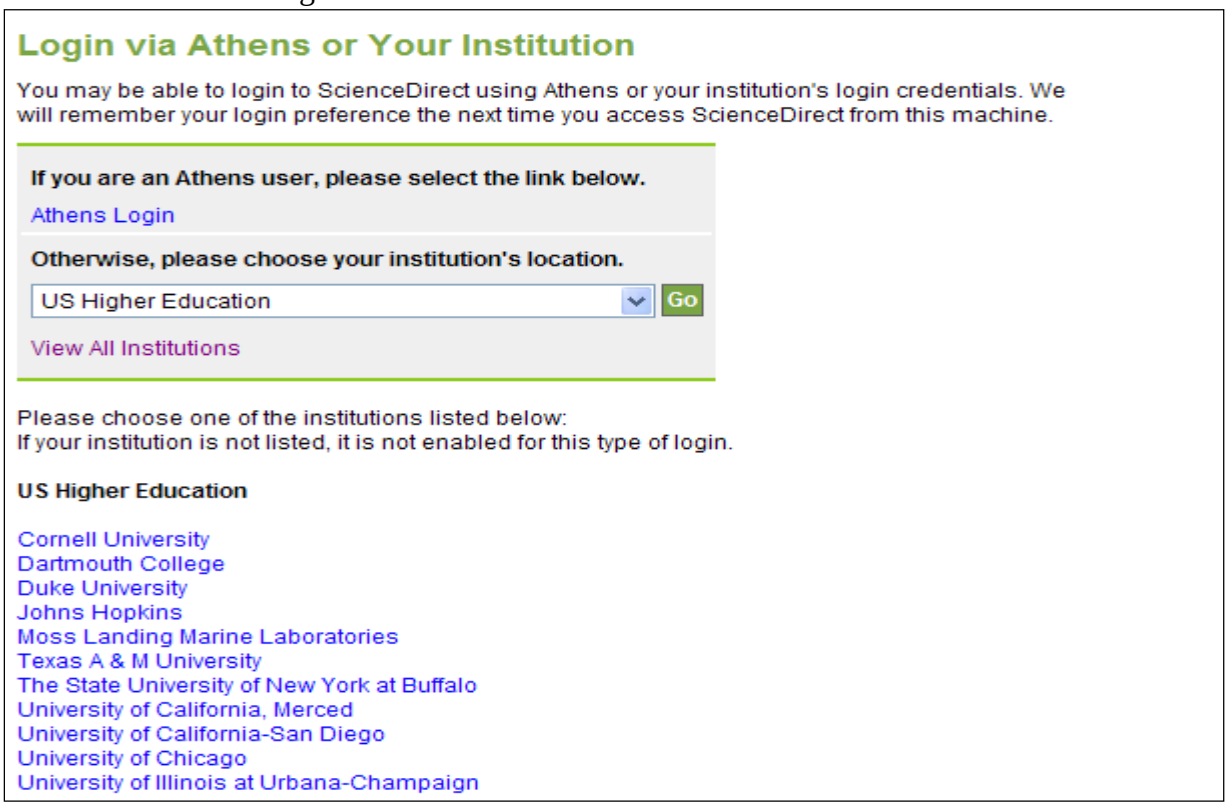
STEP 1: User clicks on Athens/Institution Login at the ScienceDirect Guest Access page.



STEP 2: User views the list of regions/federations



STEP 3: User selects US Higher Education



STEP 4: User selects the UCSD campus (which participates in a Shibboleth pilot sponsored by the US InCommon federation)

This last part about how the user gets from ScienceDirect to a local institution's Single Sign On infrastructure is the main power of Shibboleth. There is a lot of behind the scenes processing and pre-configuration that couple a random resource with a random institution.[6]

The reason this all works is because both the vendor and the institution are participants in a community of trust called the InCommon Federation (http://www.incommonfederation.org/participants/).

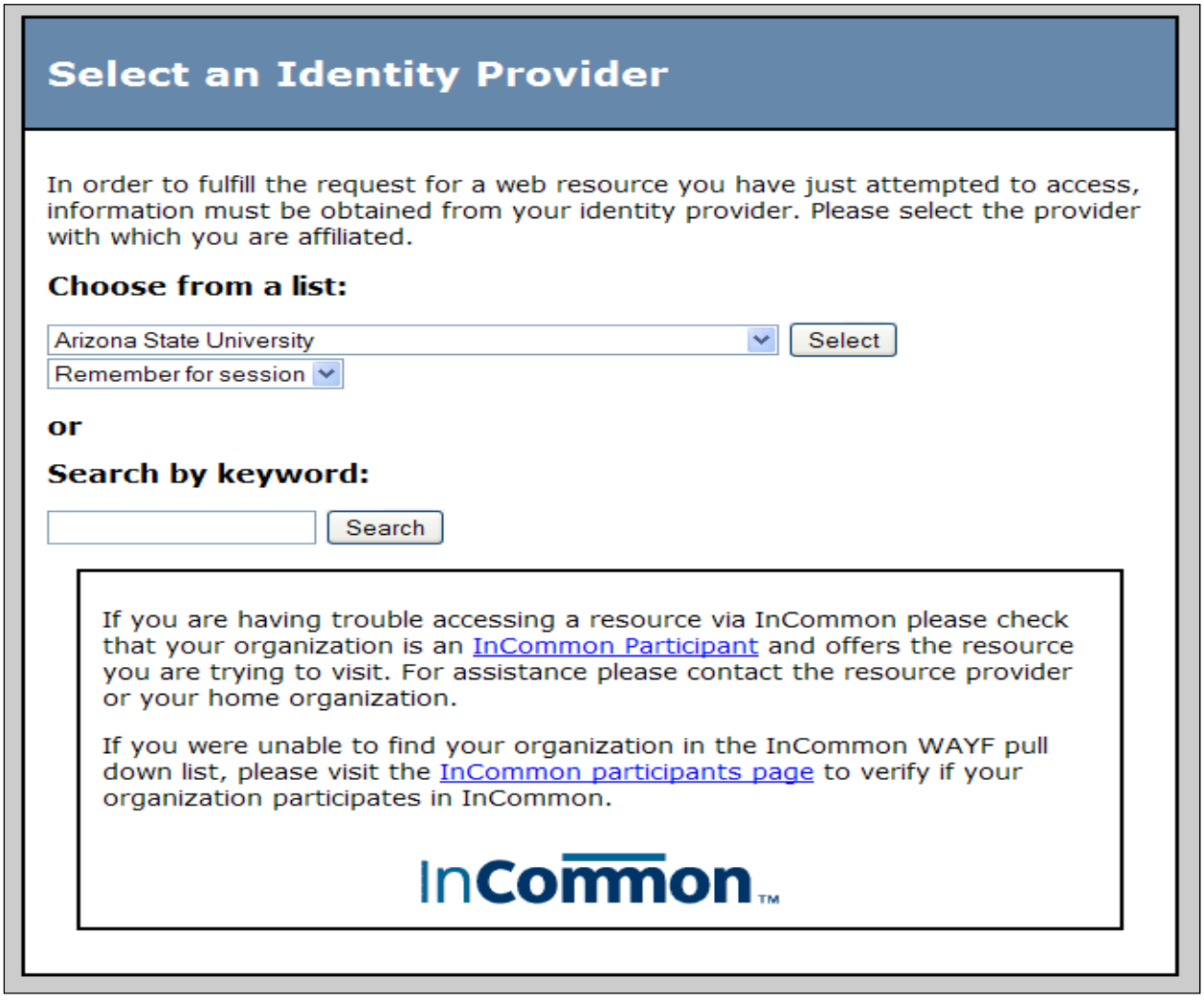
There have been agreements formed among the participants in the InCommon Federation that allow for institutional Single Sign On credentials to be used as valid entry into a vendor's resources. In a perfect world, where all campuses have a Shibboleth (or its underlying protocol, SAML) based Single Sign On mechanism, and all vendors have Shibbolized their web sites, the need for IP restrictions goes away, and thus the need for VPNs and Proxies for vendor resource access. The work involved in maintaining IP lists at the campus and vendors is replaced by membership in the InCommon Federation.

[6] In-depth technical details available at http://shibboleth.internet2.edu/docs/internet2-mace-shibboleth-arch-protocols-200509.pdf

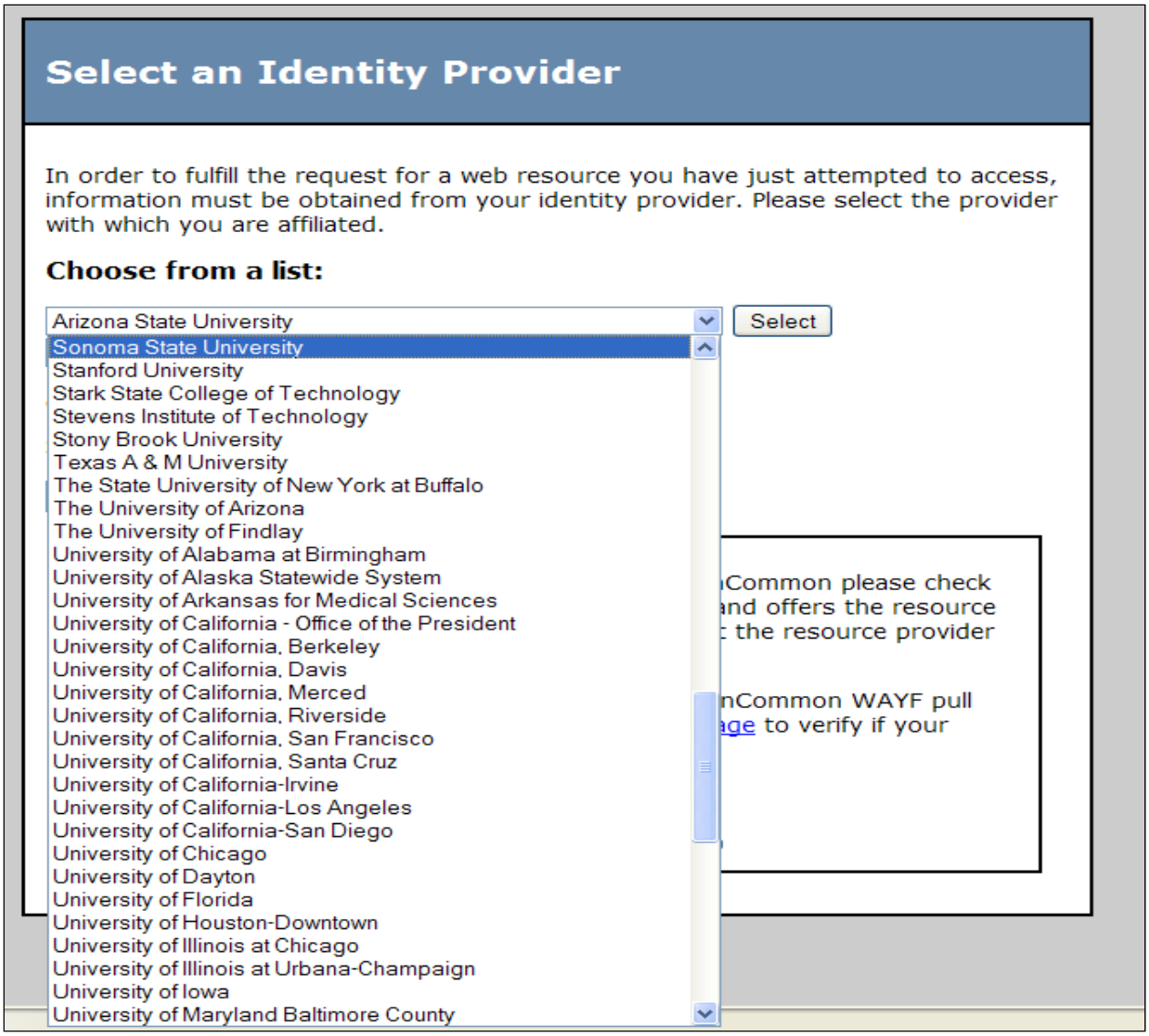
Of course, we are not in a perfect world. As we head in that direction, there will be a need for hybrid approaches that include all of the access mechanisms mentioned in this document, but an adoption of Shibboleth in the majority of institutions and vendors – a process that is well underway – will lead to a reduction in the number of and maintenance of authentication methods and more stability and uniformity in the user experience.

2. WAYFLESS URLs

The ScienceDirect example above shows the Vendor's own approach to asking the "Where Are You From" (WAYF). The InCommon Federation maintains a WAYF server that is available to all InCommon participating institutions and vendors. It looks like this:



And this next picture shows why some vendors choose not to use it:



This is just a small screen grab of the 155 Higher Ed Participants. The WAYF server, and subsequent list is seen as unwieldy by many. A user has to possibly scroll down many screens to get to her institution. However, many vendors support an alternative approach that removes the "Where Are You From" step and an institution's users can be directed to use specially constructed "WAYFless" URLs.

A WAYFless URL starts with the Service Provider's WAYFless server site and contains information about the user's institution (to which the user must authenticate) and the location of the specific resource that the user intends to reach. The generic form of a WAYFless URL is:

http://service-provider-site/session-initiator-url?entityID=IDENTITY-PROVIDER-ENTITYID&target=RESOURCE-LOCATION[7]

It creates some pretty ugly URLs, but because the institution's Single Sign On service is specified in the URL, the user does not have to pick from a long list. However, it is possible to create a short alias for this cumbersome URL.

Another reason WAYFless URLs are useful is for institutions whose users normally come to a library website to select their vendor resources. A common example of this is a Databases A-Z list. If the URLs behind the links in these lists are WAYFless, the user has one less step in getting to a resource.

As described further in the costs section below, consortially shared UC resources may have some challenges using WAYFless URLs as there is no centralized IdP for the University of California. Instead all 10 campuses and UCOP serve as separate IdPs for their constituents and presumably eleven WAYFless URLs would need to be constructed for each resource licensed for systemwide use.

IV. Benefits

1. Advantages of single sign on

The Task Force attempted to determine the extent of end-user difficulty caused by the various independent authentication methods at the ten UC campus libraries. We also wanted to demonstrate the benefits of Shibboleth implementation to the user. The Task Force looked at a variety of user information and data gathered from the University of California and beyond, and undertook a brief (non-scientific) survey of UC students regarding their use of authentication methods in the libraries.

User frustration with authentication to library resources in academic libraries has been a documented problem, ever since libraries started making online resources available remotely.[8] In a 2008 presentation documented in *The Serials Librarian*, Holly Eggleston from UC San Diego discussed some of the challenges in managing access to e-resources experience by users at a UC campus.[9] She enumerated the multiple and frustrating issues that users must deal with in the current campus environment(s) which include (but are not limited to) computer and browser configuration, browser compatibility, firewall issues, username and passwords for resources, username and passwords for proxy servers, and changing IP addresses. As documented in Appendix 1, currently there are 20

[7] More details about constructing WAYFless URLs in the UC context are available at https://spaces.ais.ucla.edu/display/uctrustwg/Invoking+UCTrust+and+InCommon+Applications+without+WAYF+Processing

[8] Furnell, S.M. "Authenticating ourselves: Will we ever escape the password?" *Network security volume*, no. 3 (2005): 8-13. Furnell, S.M., P.S. Dowland., H.M. Illingworth, P.L. Reynolds. "Authentication and supervision: A survey of user attitudes." *Computers and Security* 19, no. 6 (2000): 529-539.

[9] Eggleston, H. "Simplifying Licensed Resource Access Through Shibboleth.". *The Serials Librarian* 56 nos. 1-4 (2009: 209-14.

authentication implementations used in the 10 campus libraries and UCOP (some campuses have a proxy, web VPN and client VPN).

Access and authentication questions are commonly seen in digital reference queues. It is natural that users who are trying to connect to an online resource or service will choose an online method to ask for help. A 2005 survey, "Quantifying Cooperation: Collaborative Digital Reference Service in the Large Academic Library in College and Research Library" indicates that "database/technical access" questions are between 15 and 25% of questions asked via chat and email reference.[10]

The 20 authentication methods present challenges for the UC digital reference collaborative where librarians answer questions from users at all 10 campuses. In FY2007, the UC Digital Reference Common Interest Group analyzed the UC digital reference questions for five months during the academic year. They determined that 24% of the questions were coded as "Access" questions (defined as problems gaining access to electronic resources). Many of these questions were answered by supplying callers with information about the campus proxy servers or VPN services.

In 2009 an analysis of the UC digital reference questions asked in October and November shows that approximately 10% of the questions are coded by the librarians as "off-campus access." However, this does not account for access for problems with library services for on-campus users. A survey of UC Irvine digital reference transcripts during that same time, found that 20% of the questions were about off campus access. This difference is easily explained in that there are multiple ways to code questions, especially if the user was having problems authenticating to ILL, electronic reserves, CMS, etc. This data corresponds with a 2009 analysis of the CDL helpdesk in which a full 20% of the work tickets were categorized as problems with "access to electronic content." The UC digital reference averages 50 user questions per day, therefore decreasing the questions regarding authentication could reduce the staff time by five to ten questions per day.[11] Three quarters of digital reference questions take an average of ten minutes, therefore we could save from an hour and an hour and a half in digital reference staff time per day .[12]

To further investigate the extent of the concern regarding multiple authentication methods on UC students, the Task Force administered a brief survey to library student workers at three campuses (UCM, UCR, and UCSD). The survey asked students: how many secure online resources they regularly use for school or work; how many separate passwords and logins they currently use on a weekly basis;

---

[10] De Groote, Sandra L., Josephine L. Dorsch, Scott Collard, and Carol Scherrer. "Quantifying Cooperation: Collaborative Digital Reference Service in the Large Academic Library." *College & Research Libraries* 66, no. 5 (September 2005): 436-54.

[11] According to the 2009-2010 data collected by the UC Digital Reference Common Interest Group http://ucdigref.pbworks.com/f/09-10_UC_Dig-ref.xls (accessed 22 March 2010)

[12] Breitbach, William, Matthew Mallard and Robert Sage. "Using Meebo's embedded IM for academic reference services: A case study." Reference Services Review 37, no. 1 (2009): 83-98.

and to rate the value to them of a single sign-on for all online Library resources and services. Out of 52 students, 83% use three or more separate passwords on a weekly basis; 86% use three or more separate logins on a weekly basis. 72% rated the value of having a single sign-on for all Library resources and services as seven or above (where ten equals "great value").

The data and information we gathered demonstrates that the pursuit of Shibboleth as a single sign-on authentication method across the UC Library system is a step that should alleviate a great deal of frustration and inconvenience for library users. We note that this is in keeping with specific recommendations from other quarters, for example HOPS' 2008 "Big Idea to "Create a common user experience across all campuses through universal access to services and collections" and their specific recommendation to "Remove Technological and Administrative Roadblocks."[13] A complete adoption of Shibboleth-based single sign-on would also save significant staff time in responding to user questions.

- **High intangible benefit**: Create a stable, predictable user experience (alleviate current user frustration and inconvenience).
- **Moderate tangible benefit**: Public service staff time saved by reducing the number and variety of user inquiries about online authentication.

2. Reducing costs to maintain current authentication methods

A tangible benefit, but one that is difficult to quantify, are the costs saved or avoided upon a consolidation to a single Shibboleth-based authentication infrastructure. As mentioned above, the libraries collectively have three primary alternatives for authentication and twenty instances of high-level authentication structures. Library services are built or deployed to use one or more of these alternatives, and library IT staff are involved in local configurations and coordination in order to use them. Consolidation to a single authentication structure would create one-time conversion costs but long-term cost savings/avoidance in staff time devoted to maintenance of multiple authentication practices.

- **Low to moderate tangible benefit**: IT staff time saved by reducing the number and variety of authentication methods needing local configuration.
- **Low to moderate tangible benefit**: Library staff time saved by reducing and eventually eliminating IP table maintenance and related communication with content vendors.

3. Alignment with international, national, and local trends

As mentioned in the charge for the Task Force, the higher education community worldwide is reaching critical mass in support of Shibboleth and "Federated" authentication. The advantages for individual institutions, vendors, and regional federations derive primarily from the standardization of roles (identity providers, and service providers), of information and information exchanges (identity attributes and attribute release policies), and of trust relationships (based on intra-federation audits of compliance with required or best practices). Regional federations represent nearly every corner of the globe.

[13] As reported, eg. In *HOPS Activities Report & Goals 2008/2009* (http://libraries.universityofcalifornia.edu/hops/Activity_reports/HOPS_Activity_Report_0809.doc).

In some regions Shibboleth support is required as a condition for a service provider to provide services to the higher education members of the regional federation (e.g. JISC Collections licenses now require Shibboleth-based authentication).

In addition many institutions are specifying Shibboleth as the default authentication method for service provision within their institutional environment. At UC, while all campuses are members of InCommon, UCLA, UCSD, and Merced have also specified Shibboleth as the standard way to access campus web applications.

Finally, it appears that new partnerships and consortial arrangements are beginning to leverage Shibboleth. The UC libraries' participation in the HathiTrust is a case in point; HathiTrust end-user services will be provided through Shibboleth authentication, and collaborative service development is likely also to be managed with Shibboleth-based authentication of participants.

- **High intangible benefit**: Enhance reputation through participation and, where possible, leadership in deploying contemporary online information services; leverage emerging higher ed authentication experience and policy decisions (avoid costs related to independent actions/policies).

V. Costs

1. Licensed content

a. Increasing the pool of vendors who support Shibboleth-based authentication

In the U.S. the InCommon Federation was formed "to create and support a common framework for trustworthy shared management of access to on-line resources in support of education and research in the United States." Similar federations exist in the UK and Europe. The federations comprise both educational institutions and Service Providers such as publishers of scholarly material. Through the federations, collaborative decisions are reached on best practices for implementation of Shibboleth.

Since 2007, InCommon's library Task Force (InC-Library, at https://spaces.internet2.edu/display/inclibrary/) has concentrated on attracting scholarly publishers to the federation and developing best practices for the U.S. library community and its partners to embrace.

InC-Library maintains a registry of Resources (Publishers) that have already implemented Shibboleth or are planning to do so, (https://spaces.internet2.edu/display/inclibrary/RegistryOfResources) as well as a list of Resources/publishers belonging to the UK federation and under recruitment to InCommon (https://spaces.internet2.edu/display/inclibrary/TargetResources). Prompted by the Shibboleth Task Force, in the Winter of 2009-10 the UC libraries joined this recruitment effort in a modest way, with UC signatures added to recruitment letters sent to vendors.

The InC-Library's ongoing recruitment and lobbying efforts provide a low-cost and national-level forum through which UC can encourage content vendors and other third parties to adopt and improve Shibboleth-based authentication.

- **Low costs:** to increase the number of vendors adopting Shibboleth as the primary authentication method by participating in InC-Library efforts.

b. WAYFless URL construction

Recall from section II above that "WAYFless URLs are pointers to resources that allow for bypassing the Shibboleth Where Are You From (WAYF) step. For our purposes, this means providing a URL syntax such that a resource URL could be created to navigate the user through the authentication/SSO process without prompting the user to identify their institution."[14]

In order for a URL to be WAYFless compliant it minimally must include an IdentityProvider ID. This ID would be unique to each specific UC campus, regardless of whether it is a Tier 1, 2, or 3 resource. That means that the "IDENTITY-PROVIDER-ENTITYID" portion from this generic URL example would need to be a specific and consistent code identifying the location as UCLA, UCR, UCI, etc:

http://resource-provider-site/session-initiator-url?entityID=IDENTITY-PROVIDER-ENTITYID&target=RESOURCE-LOCATION

When WAYFless URLs are used in conjunction with Shibboleth the affiliation selection step is removed from the authentication process. Instead, the user clicks on a resource's URL and is taken directly to the resource. If WAYFless URLs are used they would presumably need to be made available in all discovery systems and databases, including OPAC bibliographic records, Database A-Z services, Course/Subject Guides, and wherever else the resource URLs appear in library online and print materials.

Working with EZproxy software, the InCommon Library task force has developed a shortcut to the construction of WAYFless URLs and is promulgating it as a "best practice." EX Proxy can be configured with the right logic – knowledge of the campus IdP and the Vendor's WAYFless server site - to construct WAYFless URLs on the fly. EZproxy also supports IP-based authentication, so assists in the mixed environment of Shibboleth and IP-based authentication.

Therefore EZproxy may simplify authentication management for those campuses that have it or are willing to adopt it. However, it does not solve the consortial challenge of multiple WAYFless URLs for each resource in a shared discovery tool, unless all campuses were to adopt it.

To summarize, the primary benefits of WAYFless URLs include:

1. One less step for users to access library resources and services;
2. The user does not need to know what institution they are affiliated with.

[10] From the section *Best practice #2: Implement WAYFless UR.* InC-Library Task Force..
https://spaces.internet2.edu/display/inclibrary/Best+Practices

The drawbacks of this approach, which requires that URLs for each resource would need to be different for each campus[15], include:

1. Potentially large workload to convert existing resource URLs to WAYFless URLS (for campuses without EZproxy) or to configure EZproxy to construct URLs with WAYFless format (for those campuses with EZproxy)
2. Additional workflow process for metadata and/or IT staff to convert new URLs to WAYFless standard.
3. Other projects, such as the proposed centralization of subject guides, would need to be designed to accommodate the construction of up to 11 variants of a resource's URL.

According to our research UCB, UCD, UCM, UCR, UCSD and UCSF employ a proxy service other than EZproxy or use VPN exclusively.

UCLA, UCSB and UCSC employ EZproxy and may be able to leverage its rewriting function to form WAYFless URLs on the fly. However, the proxy configuration file would need to be updated with the proper WAYFless construction logic.

If WAYFless URLs are employed there would be associated:

- **Moderate costs:** to purchase, configure, and maintain EZ Proxy if a decision was made to leverage its rewriting function to produce WAYFless URLs on the fly.
- **High conversion costs:** to construct and maintain WAYFless URLs and substitute them for current URLs in all discovery systems if EZproxy were not used to generate them on the fly.

2. Third party software costs

The UC libraries use commercial or open source software to provide many basic services. The services provided may require authentication for some or all of the service components, e.g. to initiate an interlibrary loan as a legitimate student, faculty or staff member, to take advantage of recall and hold options as a library catalog user, or to complete an interlibrary loan request as an interlibrary services librarian or staff member. In these cases the library is acting, in Shibboleth terms, as a Service Provider and must run the service- provider Shibboleth components and register with the campus IdP as a trusted requestor/recipient of information about campus users.

In order to scope the potential costs involved, the Task Force inventoried the types and numbers of library applications with an authentication component (the results are reported in Appendix 2). The inventory and follow up research revealed a generally favorable situation with regards to the current use of or compatibility with Shibboleth by the most heavily-used 3rd party software vendors and products. In most cases Shibboleth compatibility is included in the current release of the software (e.g. SFX, ContentDM, VDX, Moodle, Blackboard). This is not true in the notable case of Innovative Interfaces Inc. Millennium ILS product.

[15] It is conceivable that a "smart" utility could be built to construct the correct WAYFless URL on the fly. Indeed that possibility has been added to the ExLibris SFX development list, allowing SFX to leverage its knowledge of user affiliation and resource locations. The task force is unaware of an implementation of this approach.

- **High capital costs ($20-$30K per campus):** to add Shibboleth/single-sign-on functionality to III ILS.
- **Low costs:** to use Shibboleth components in most 3rd party library software (primarily staff/labor for configuration and maintenance and interaction with campus Identity Providers to register a new service).

3. Locally developed & hosted software

Many of the UC libraries, and the CDL, develop their own software to support internal and patron-facing services. The authentication component of these services, when required, is often custom-built as well, generally using a local database of accounts and passwords and occasionally leveraging campus single-sign-on infrastructure.

Fortunately, Shibboleth modules are open source themselves and are well-documented. Additionally, at least three libraries within UC have some experience incorporating Shibboleth into locally-built services (UCSD, UCLA, and, to a lesser extent, CDL). A preliminary analysis suggests moderate effort is required. The effort includes registering with the campus IdP and installing and integrating the Shibboleth service-provider software (normally as apache/web components). A rough estimate is two weeks of programmer/analyst time. A more thorough description of the tasks involved will be included in the action planning/implementation phase of the Task Force work.

- **Moderate costs:** to integrate Shibboleth components into local software (primarily staff/labor for configuration and maintenance and interaction with campus Identity Providers to register a new service).

VI. Next Steps

Although the two high value benefits – improved user experience and staying current with trends - are intangible, the TF believes that, in combination with the low and moderate benefits we have identified, that there is promise in their pursuit.

Of the two high cost tasks – integration of WAYFless URLs, and purchase of Shibboleth functionality for III Millennium systems – the integration of WAYFless URLs appears to be optional, depending upon how aggressively the libraries would attempt to reduce end-user steps to access resources.

Accordingly, the Task Force recommends that the libraries proceed collectively by continuing with the "action plan" component of the Task Force work. While we believe that the benefits accrue at both local and systemwide levels, we recommend that the action plan focus on the tasks and planning necessary for developing Shibboleth authentication for tier 1 and tier 2 resources (including HathiTrust) and for systemwide services (UCeLinks, eScholarship, etc).

We would intend in the action plan phase of our work to document decisions and steps necessary for using Shibboleth for tier 3 resources, and locally developed or hosted services, but it seems likely that those pursuits will be undertaken through separate decision processes at each campus.

Several actions taken for systemwide resources will pave the way for local action as well, for example a conventional way to request the addition of services and the release of attributes to service providers from the campus IdPs. Following on this intersection between systemwide and local action and experience, the Task Force is likely to recommend an ongoing mechanism for sharing Shibboleth experience and expertise across the Libraries.

Appendix 1 – Current Off Campus Access Methods

Campus	Proxy	Web VPN	Client VPN	Notes
Berkeley	X		X	
Davis	X	X		
Irvine		X	X	
Los Angeles	X		X	BioMed Library has a separate VPN system
Merced			X	
Riverside		X	X	
San Diego	X	X	X	
San Francisco		X	X	The description of the "Network Connect" option is similar to the client VPN.
Santa Barbara	X		X	Log in to proxy from home page
Santa Cruz	X			
UCOP			X	

Appendix 2: Inventory of Library Services Requiring Authentication

Note: these are the separate types of services reported by Task Force members and verified by the Library Technology Advisory Group (LTAG); we did not attempt to exhaustively inventory which libraries offered which services across UC. Arranged in descending order by size of likely/known authenticated user base.

Service	Service scope	# of users	Authentication for whom? (faculty, lib staff, etc.)
Tier 1 content	Systemwide	100,000s	UC Community
EReserves	Campus	10,000s	students and instructors
III Patron Record	Campus (but multiple campuses have the same system)	10,000s	faculty, students, staff, public
EReserves	Campus	10,000s	students and instructors
Moodle/Blackboard (course management)	Campus	10,000s	faculty, students, staff
eScholarship	world	1,000s	public for personal svcs
Melvyl catalog	world	1,000s	public for personal svcs
HathiTrust	Systemwide/HT partners	1,000s	UC users for personal svcs
Online Archive of California	Systemwide & other CA orgs	1,000s	statewide contributors
eScholarship editions	Systemwide	1,000s	UC Community for deposit/personal svcs
Next Generation Melvyl	Systemwide	1,000s	public for personal svcs
Request (PIR)	Systemwide	1,000s	UC Community
CBS/VDX	Systemwide	1,000s	UC ILL users
Wiki (e.g. Confluence)	Campus	1,000s	students and personnel, visiting guests
Library workstation login (both staff and users)	Campus	1,000s	students and personnel; library staff

Wireless access (for campus or for library only)	Campus	1,000s	students and personnel; library staff
Staff portal	Campus	1,000s	campus staff
Student portal	Campus	1,000s	students
Webmail	Campus	1,000s	faculty, staff, students
eScholarship	world	100s	depositors
CBS/VDX	Systemwide	100s	UC ILL staff
III Staff Interface	Campus (but multiple campuses have the same system)	100s	library staff
Public printing	Campus	100s	students and personnel, public
Web content management	Campus	100s	campus web editors
Digital Asset Management System	Campus, Public	100s for login, 1,000s for access	library staff
Web archiving services	Systemwide/world academic community	100s	archive builders + curators
Mass digitization database	Systemwide (potential)	10s	contributors (Google, OCA)
Melvyl catalog record uploads	Systemwide	10s	Lib. cataloging liaisons? Machine processes
UCeLinks (SFX)	Systemwide	10s	SFX liaisons
CDL & UCLibraries website	Systemwide	10s	web content editors
Digital Signage Content Manager	Library (+ other depts)	10s	library staff plus others
Numura Footprints (help desk software)	Library	10s	library staff
Bug tracking systems	Libraries	10s	library staff
Code/source control systems	Libraries	10s	library IT staff
Network Storage	Library	10s – 100s	library staff

Web content management	Library	10s – 100s	library staff
Curation micro services	Systemwide/world academic community	10s -100s	Service users
Digital Preservation Repository	Systemwide	10s-100s	Depositors

Appendix 2 - Research Library Experience with Shibboleth

Note: the following is based on brief interviews conducted via phone or email in April and May 2010.

Library	Info Source	Using shib with which content providers?	Library apps using shib	Following InC-lib Best practices?	WAYFless URLs	Rely on EZProxy?	Notes
Duke	Direct (email)	Refworks; ScienceDirect; Jstor; Ebscohost; Wilsonweb; Project Muse	catalog (aleph); federated search (metalib); proxy server (ezproxy) ILL (ILLiad);digital collections management (Trident/Fedora); link resolver - SFX basically gets some shib treatment, by way of ezproxy; our intranet, wikis, trac instances some home-grown interfaces	yes, and is using ezproxy to create WAYFless URLs	Yes; if vendor supports	Yes for off-campus users	"pretty much everything in the library that requires auth, we try to use shib; We have worked out with [campus IT to release] eduPersonEntitlement ; we do sometimes get specific attributes released to our internal applications."
Johns Hopkins University	Direct (email)	~ 10 vendors that support wayfless URLs, encouraging other vendors to support shib	unknown	yes, and is using ezproxy to create WAYFless URLs	Yes; for ~10 vendors that support	Yes for off-campus users; otherwise IP	"JHU has a very good shib support from our central IT group that using federated access to ADP."
MIT	Indirect (InC-Lib case study)						Our overall aim is to implement Shibboleth SSO as widely as possible so we can return authentication where it belongs; with central computing where credentials are managed
Univ. of	Direct	paused in efforts to	Planning to shibbolize	yes, especially	Yes: for	Yes for	"Many units on campus current

Chicago	(email)	move content providers from EZproxy to direct access via Shib; hope to resume efforts to add content SPs to our shib config during [summer]	SFX	with regard to use of EZProxy	services provided by Atlas Systems: Aeon, ARes, and ILLiad	off-campus users; otherwise IP	rely on the IdP for authN, some run their own SPs. [but] Procedures for adding SPs to our shib configuration have not been finalized, so it is unclear how much of that will reside in the Library."
UNC							
Note: the following libraries participated in 2008-2009 InCOmmon Library pilot: Cornell University, Penn State, UC-San Diego, The University of Chicago, University of Maryland, University of Washington							

Appendix 3 – Content Vendors Shibboleth Authentication Status

Note: info. From TF member interactions with vendors and per the InCommon Library Working Group's "Registry of Resources" page at https://spaces.internet2.edu/display/inclibrary/RegistryOfResources (consulted on 8/11/2010; source last updated on 7/22/2010).

Shibboleth in production now:
EbscoHost
Electronic Book Library [note: ONLY Shibboleth auth is available]
Elsevier
Gale
HathiTrust – [note: ONLY Shiboleth auth is available for several services]
Jstor
OCLC (FirstSearch)
Project Muse
ProQuest (Classic; Chadwyk-Healey; CSA)
Refworks
Thomson-Reuters (Web of Science)
H.W. Wilson

Considering but not yet in production:
ACS
BioOne
IEEE
LexisNexis
Nature Publishing Group
Oxford University Press journals
Ovid
Sage
Springer
Wiley

Appendix 4 – Sample Recruitment Letter sent to Content Vendor

-----Original Message-----
From: David Kennedy [mailto:david.kennedy@duke.edu]
Sent: Monday, January 25, 2010 9:29 AM
To: Lengwenat, Ulrike, Springer DE; heather.staines@springer.com; beth.mayes@springer.com
Cc: Andy Ingham; Adam Chandler; John Ober; varnum@umich.edu; jkiser@upenn.edu; parker@mlml.calstate.edu; ccarr1@vt.edu; paul-soderdahl@uiowa.edu; tod@uchicago.edu; t-howell@northwestern.edu; Anne_Nolan@brown.edu; julia.kochi@ucsf.edu; theodora.toy@ucr.edu; elin@ucmerced.edu; jdooley@ucmerced.edu; Dave Kennedy; inc-lib-vendor@incommonfederation.org
Subject: Shibboleth/InCommon authentication for Springer

Dear Ulrike,

I am writing you on behalf of Duke University and the other universities listed at the end of this email. We are all committed to the concept of federated identity management and federated access to licensed resources. Hence, we have implemented, or are implementing, the Shibboleth technology at our respective institutions (which have all also joined the InCommon federation).

We have noticed that you have implemented Shibboleth/SAML within your product line, and are using the technology to provide federated access to your services within the UK Access Management Federation. We, collectively, would like to invite you to also join the InCommon federation so that we, as well, will be able to federate with you. You provide valuable services to our user communities, and we would like to begin integrating those services with the single sign on infrastructures we have invested in at our respective institutions.

Duke University would like to be your sponsor for membership into InCommon and help facilitate the process. We can provide guidance in terms of federation policies and best practices. We also have a willing group of universities that should be able to help with testing and implementation, and we can help coordinate that as well.

If you are interested, we can set up a conference call to discuss some of the details and initiate the process.

Thank you for your time and consideration
Dave

David Kennedy, Duke University

Andy Ingham, University of North Carolina, Chapel Hill
Adam Chandler, Cornell University
John Ober, California Digital Library
Ken Varnum, University of Michigan
John Kiser, University of Pennsylvania
Joan Parker, Cal State, Moss Landing Marine Labs
Curtis Carr, Virginia Tech
Paul Soderdahl, University of Iowa
Tod Olson, University of Chicago
Thomas Howell, Northwestern University
Anne Nolan, Brown University
Julia Kochi, University of California, San Francisco
Theodora Toy, University of California, Riverside
Emily Lin, Jim Dooley, University of California, Merced

Appendix 5 - UCTrust Library Issues Survey Responses

The following survey was created by an ad hoc working group comprising members from UCTrust (David Walker and Surya Narayana) and the Libraries' Shibboleth Task Force (Declan Fleming and John Ober). The following responses were available as of 9/20/2010 (with last contribution from UCTrust members on 9/13/2010)

As discussed at the 6/21/2010 UCTrust Conference Call, UC Trust principal contacts are requested to answer questions that will inform the support of UC library Shibboleth planning. Refer to UC Trust - UC Libraries ad hoc working group -- Request for Info from UCTrust for context.

Background Info

1. What is your campus' single sign on solution?

UCB: CAS
UCD: CAS
UCI: Home grown WebAuth software on top of Kerberos.
UCM:
UCR: CAS
UCLA: We use Shibboleth natively as our SSO solution.
UCSD: We use Shibboleth natively as our SSO solution.
UCSF: Shibboleth
UCSB: Currently, the WAM is Netpoint. It will become OpenSSO this fall. There is no current plan to use intra-campus Shibb technology.
UCSC: We are using Shibboleth natively as our SSO solution.

2. At your campus, should the library use the contact listed in http://www.ucop.edu/irc/itlc/uctrust/contacts.html for IdP questions?
a. Is there a campus mailing list for service issues and questions?

UCB: Yes, or calnet-idm@lists.berkeley.edu
UCD: Yes, ldapadmin at ucdavis dot edu
UCI: Contact OIT@UCI.EDU since contacting a single person may be unfortunate if they or unavailable.
UCM:
UCR: Yes
UCLA: Yes, or contact iamucla@ucla.edu.
UCSD:Yes. a) we use shibsupport@ucsd.edu.
UCSF: Yes, we have a mailing list. The contact for UCSF should be Surya Narayana (surya.narayana@ucsf.edu)
UCSB:Yes.Identity problems are directed to directoryhelp@isc.ucsb.edu. This is not a list.
UCSC:Yes (there are alternates if necessary). We generally intake questions through help@ucsc.edu, which populates a trouble ticket that should be escalated to our group.

3. What attributes does your campus commonly release via Shibboleth?

UCB: At present, the following: ucnetid, uctrustcampusidshort, edupersonscopedaffiliation, edupersonprincipalname, givenname, displayname, mail. In the future, we will add uctrustassurance, targetedID, and others by request.
UCD: ePPN,displayName,mail,ucNetID,ucBasicAssurance,ucCampusShortID,cn,sn,givenName. Others available, depending on needs.
UCI: ePPN commonly, other local attributes on request
UCM:
UCR: In general, ePPN plus whatever the service requires.
UCLA: UCLA by default releases targetedID. Everything else is subject to data steward approval. We can assert name, basic contact info, eduPersonAffiliation, UC Trust attributes, and by agreement, eduPersonEntitlement vallues.
UCSD: targetedID, affiliation and scopedAffiliation, mail, name attributes. We have a lot of ucsd specific attributes we release internally as well, which can sometimes be mapped to standard attributes.
UCSF:ePPN, sn, givenname, mail
UCSB:None yet. The design is for what I believe are the minimums mandated: ucnetid, uctrustcampusidshort, uctrustassurance, ucemployeeid, edupersonscopedaffiliation, edupersonprincipalname, sn, givenname, displayname, mail.
UCSC:We have on occasion released ePAffiliation (unscoped), ePPN, sn, givenname, mail, uctrustcampusidshort. We have other attributes available, but notable ones we do NOT have are: targetedID, entitlement management.

4. At your campus, how do we add new vendors to the Shibboleth list?
a. Is there a formal process to request a new SP be registered as the recipient of ID information via shibboleth?
i. Is there a form, and if so what is its location/URL?
ii. Does the process accommodate both campus service providers and external service providers (e.g. that the library sponsors or brokers)?
iii. How does the process accommodate requests for attributes not otherwise/previously released (e.g. the InCommon -Lib recommended use of eduPersonEntitlement, see below)

UCB: At present, these requests are handled ad hoc. The CalNet team is the main point of contact and coordination (calnet-idm@lists.berkeley.edu) for InCommon registration, metadata updates, getting approval from data proprietors, etc. There is a formal process for requesting release of information not public in our directory. The vendor must respond to the same set of questions we use for privileged LDAP binds (see http://wikihub.berkeley.edu/x/WIJ)
UCD: Requests are generally received via email, or through meetings and phone calls from/with sponsors (the desire for a more formalized process has been expressed by some). Requests involving new attributes may be sent through administrative review for prioritization and scheduling, and sensitive/protected attributes may require assurance of ongoing protection.
UCI: There is a process being formalized. External SPs need campus sponsors. Additional attributes depend on their existence, ease of relay, and who is in charge of making decisions about them.
UCM:
UCR: Requests are currently handled on an ad hoc basis. A process is currently being defined and proposed.
UCLA: A vendor must be sponsored by a campus department. If the vendor is a member of InCommon, it's a matter of us releasing attributes. If the vendor isn't an InCommon vendor, the IAMUCLA team needs to register its metadata in our IDP. To get started, contact iamucla@ucla.edu.
UCSD:An email to shibsupport@ucsd.edu will get things started. Local SPs can use a form at

https://a4.ucsd.edu/shibreg/docs. If we decide that we can and should release a new attribute we take the time to implement it.
UCSF: We don't have a form yet. We have a process of getting approval for attribute release from the data owners; same goes for new attributes as well.
UCSB: There is no plan for local SPs. Special attribute requests will be handled by a committee.
UCSC: There is a form, but a request though our trouble ticket system is sufficient to initiate the process (help@ucsc.edu). Yes, we can accommodate external SPs; requires approval. Release of otherwise unreleased attributes is the same as release of existing, request is made, approval granted or not granted. Complexity of approval depends on sensitivity of the data and how it will be used.

b. How much lead time is needed to add a new SP - if only currently available attributes are needed? if new attributes are needed?

UCB: Usually 1-2 weeks for available attributes that do not need separate approval from data proprietors. If approval is needed, it can take up to a month.
UCD: If the SP is already configured, IdP mods can be done in as little as a couple of hours with currently available attributes. Registration with InCommon takes a day or two, depending how much back and forth is required with the SP owner. New attributes usually require a degree of research and policy validation. If they can be computed from existing attributes, it might take a day or more to prototype and test. If an attribute is completely de novo, it may take multiple weeks to a few months, depending on required infrastructure changes, priority and workload.
UCI: In an ideal case, lead time is two weeks. However, this is really completely dependent on the level of involvement from the SP as much as it is our own workloads, difficulties in setting up additional attributes and other Layer 8 issues.
UCM:
UCR: One week or less, after organizational approval. If new attributes are needed, it could be much longer.
UCLA: We recommend allowing for 2 weeks to 30 days for new SP registration. Attribute release aside, we have found that handshake testing takes time.
UCSD:Generally less than 24 hours if no special attribute requirements. Otherwise it could take weeks to implement a new attribute.
UCSF: A week and three weeks respectively.
UCSB: N/A. We assume all SPs will come to us from a UC system wide perspective rather than a local one.
UCSC: Generally a week or two to load a new SP with currently released attributes. Developing new attributes (not just unreleased ones, but ones we don't currently populate) depends entirely upon the requirements and prioritization.

Content vendors and InCommon-Library Best Practices

(assumes that the campus library or the CDL is the sponsoring agent for content vendors as 3rd party Service Providers)

5. Can you currently support an Attribute Release Policy that includes eduPersonScopedAffiliation?

UCB: Yes.
UCD: Yes.

UCI: Yes.
UCM:
UCR: Yes.
UCLA: Yes.
UCSD: Yes, although we don't currently assign anyone student@ucsd.edu. member, staff, and employee work.
UCSF: Yes. We currently only have "staff" "student" and "affiliate" only.
UCSB: Yes when it goes live.
UCSC: Yes, though we do not currently populate the "member" value.

6. If your library provides a list of IP addresses for terminals/workstations available for "walk-in" use, can you implement the IPAddress authentication "handler" and assign a time-limited affiliation of "library-walk-in" for any authentication request from that terminal (see https://spaces.internet2.edu/display/SHIB2/IdPAuthIP)?

UCB:
UCD: This handler is not currently supported, so would take a bit of work to prototype/implement in the context of existing priorities. We'd also want an SLA cf. who maintains CIDR blocks, expected response time to changes, etc.
UCI: Possible, but would take some work on both technical and political ends of that.
UCM:
UCR: Not without a substantial amount of work.
UCLA: Not today, but we'd like to engage in that conversation to make it happen.
UCSD: Yes, we could probably implement this.
UCSF: Will require some work to get this going, but yes.
UCSB: We currently use IP address verification through a proxy server. It is assumed that the proxy server will remain active until all Library vendors have converted to federating processes.
UCSC: We are planning to implement this in support of a local library application, so this should be fine. - update: we need to do some shib maintenance before we are able to support this function, and that maintenance can't take place before September. So we still expect to support this (hopefully by year's end), but are unable to at this time.

7. Can you currently support an Attribute Release Policy that includes eduPersonEntitlement with a value of urn:mace:dir:entitlement:common-lib-terms for all faculty, staff, students, and library-walk-ins?
a. If not, please note the timing and conditions necessary in order to support eduPersonEntitlement.
b. Are you able to assert eduPersonEntitlement selectively for individuals or groups of individuals?

UCB:
UCD: Similar as item 6 above. If entitlement can be computed from existing attributes, it should not be difficult to implement. If entitlement is required to be fine-grained, e.g. asserted for arbitrary individuals, we would have to design/implement a fair chunk of infrastructure to support it.
UCI: We could, but will have to work out the issues with walk-ins.
UCM:
UCR: We do not currently support an ARP that include eduPersonEntitlement. We could possibly implement this by the end of the calendar year (given support of upper management).
UCLA: We could, but we'd have to work out how to assert the value for walk-in's.

UCSD: We could for everyone except library walk ins at the moment. If we implement the previously mentioned IP address authentication, then walk ins would be okay.
UCSF: We could, once we have the "walk ins" configured.
UCSB: We do not have edupersonentitlement implemented. It would imply large scale deployment of custom processes for the delegated implementation that is appropriate to such a set of values. We have no funding for inventing those.
UCSC: We do not have such a value. We could probably implement the catch-all circumstance (all faculty, staff, students and library-walk-ins) using shib filters or some other process. We plan to support selective entitlement management in the future, but nothing is likely before the end of the year.

8. What information would you need about a content vendor or other 3rd party SP beyond what would be available in their InCommon certification?

UCB: As mentioned above, they would need to respond to specific questions if they are requesting release of attributes not publicly available.
UCD: Depends on application at hand cf. attribute sensitivity, and how they address (de)provisioning, SLA.
UCI: It depends on the specifics of the request, but we may like some other contact information.
UCM:
UCR: We would want assurances from the vendor on data security and data use practices.
UCLA: We may at some point ask for information on the vendor's security and data use practices, particularly as it pertains to the data we release to the vendor.
UCSD: It depends on what attributes they want. If they want attributes we don't feel giving out to verdors, I'm not sure what sort of information or contract we would expect of them. If they want basic stuff, we might not ask for much at all, especially if the company is well known.
UCSF: Contact information
UCSB: Unknown.
UCSC: As UCSD. We would expect a campus (or UC) sponsor to make the request. We would probably be happier if the vendor had signed Appendix DS.

HathiTrust as a test case (refer tohttp://www.hathitrust.org/shibboleth)

9. Can you currently support HathiTrust's Attribute Release Policy that includes eduPersonScopedAffiliation, eduPersonTargetedID, and optionally, DisplayName?

UCB: eduPersonTargetedID is not currently implemented at UCB.
UCD: eduPersonTargetedID not currently implemented at UCD, eduPersonScopedAffiliation is.
UCI: Affiliation and DisplayName, yes, but we do not currently support eduPersonTargetedID; that would take more lead time
UCM:
UCR: Yes.
UCLA:Yes
UCSD:Yes.
UCSF: Requires some work, but yes.
UCSB:No plan to at the moment. I would like to see the technical specs for a targetedid implementation and that every UC campus used the same spec. If not, we will probably avoid

it.
UCSC: eduPersonTargetedID is not currently implemented at UCSC. Assuming release of ePPN was approved (would take some time to get the okay to do this) I believe we would be able to support the service.

10. As a specific case of question #6 above, what would you need from your campus library or the CDL to register HathiTrust as a SP?

UCB: Same as UCLA.
UCD: Ditto cf. UCLA.
UCI: To register the SP we would need Campus Contact information, or more formal signatures if we've gotten along farther in our formalization process. For the IP Address configuration, we would need to set up a system for that.
UCM:
UCR: A request from the vendor and the CDL would be enough to get the ball rolling.
UCLA: We need official contact/endorsement from the UCLA library or CDL. We'll need the sponsoring party to submit a request for data release (which the IDM team can facilitate). Of course, technical contact info from {nl:HathiTrust} so we can coordinate handshake testing efforts.
UCSD:Probably just need to know that they want it. An email would probably suffice.
UCSF: Once we get over the "atttribute release" hurdle, we should be fine with just having contact info.
UCSB: Again. I see SPs as a UC issue, not a UCSB issue.
UCSC: A request from a UCSC/UC sponsor would be best.

11. Is there anything about the implementation of HathiTrust as a SP that you would find useful to track in order to inform future requests from library-sponsored SPs?

UCB: No.
UCD: No.
UCI: Probably not.
UCM:
UCR: Nope.
UCLA: not particularly.
UCSD:Probably not.
UCSF: HathiTrust implementation should point the way to streamline other library-sponsored SPs/implementations from a process perspective.
UCSB:No. I want other than UCSB to take responsibility for SPs.
UCSC:Not immediately.

Appendix 6 - Simplified Shibboleth Local Installation Guidelines

These instructions have been simplified from a UCSD specific document created by ACT's Security team. When you need more general instructions visit the Shibboleth wiki.[16] [We note also that InCommon has a set of implementation workshops for both Service Providers and Identity Providers at http://www.incommon.org/educate/shibboleth/program.html.]

Shibboleth is a web-based Single Sign-On infrastructure. It is based on SAML, a standard for web authentication through SOAP. Shibboleth has been adopted by the University of California as the asis for federated Single Sign-On between campuses (see UCTrust at http://www.ucop.edu/irc/itlc/uctrust/).

Using Shibboleth has several security and operational benefits over going direct to one of the authentication mechanisms.

1. Your server never handles the passwords so anything that goes wrong can't compromise the credentials.
2. The Shibboleth service can have additional account misuse and fraud detection capabilities.
3. The Shibboleth service can have a Logging infrastructure that meets campus requirements.
4. Future proof: with Shib you aren't binding yourself to a specific mechanism of authentication instead you are binding to a piece of middleware that allows you to pick from the many authentication mechanisms.

Terminology

Understanding Shibboleth and SAML is much easier after learning some terminology. A successful deployment of Shibboleth involves two critical software components:

Identity Provider (IdP) - This is the server that handles authentication of users.

Service Provider (SP) - An IdP is pointless without Service Providers. Service Providers are web applications, resources, or other services which require authentication. The Shibboleth SP software allows most web servers (namely Apache and IIS) to integrate with an IdP or a number of IdPs.

Service Provider Software

The SP software consists of several components:

ISAPI Filter

This is only used for Windows Server IIS deployments. It intercepts requests to IIS and redirects users to an IdP or WAYF. After the user authenticates it also handles the callback which tells your SP that the user has authenticated. During handling of this callback the ISAPI Filter collects attributes which describe the authenticated user. The filter is configured through the shibboleth2.xml configuration file.

mod_shib

[16] Note that Shibboleth v.2 has been released but not uniformly adopted across UC or the library vendor community. Therefore reference may need to be made to the now deprecated SHIB v1.3 materials at https://spaces.internet2.edu/display/SHIB/WebHome or to the SHIB v2.0 materials at https://spaces.internet2.edu/display/SHIB2/Home

This is only used for Apache deployments. It is essentially the same as the ISAPI Filter but for the Apache web server. In addition to shibboleth2.xml, some configuration is required via httpd.conf or shibd.conf.

shibd

This is a service (Windows) or daemon (UNIX) which handles attribute request queries from the SP to the IdP. Shibboleth attribute requests are part of the SAML standard and are made via a back channel SOAP call to the IdP (Usually on port 8443). In order to receive user attributes, this service must be running.

Most configuration is done via shibboleth2.xml. The following configuration areas are either included in or referenced by shibboleth2.xml:

Attribute Map

Attributes in Shibboleth are named with URNs. In order to easily access the attributes from within your application, they need to be mapped to environment variables or HTTP headers. The attribute-map.xml file defines these mappings.

Metadata

Shibboleth SPs and IdPs communicate with each other securely using X509 certificates. The SP uses metadata files to define the IdPs that it may interact with and the relevant URLs and certificates that each IdP will use. The IdP uses metadata to define the same information for SPs that it may interact with. This means that an SP and IdP must exchange their metadata before they can interact. Federations such as InCommon may assist with maintenance and distribution of metadata.

Installation (Apache or IIS)

1. Install the service provider software (http://shibboleth.internet2.edu/downloads.html). The Shibboleth wiki has installation instructions at http://spaces.internet2.edu/display/SHIB2/Installation - follow the instructions for your operating system.
2. Download shibboleth2.xml.
3. Download attribute-map.xml.
4. Extract the archives and move the extracted files to etc\shibboleth in your Shibboleth install folder.
5. Open shibboleth2.xml and change all occurrences of the install directory (/opt/shibboleth-sp/) to the path where you installed it. (If you used the default install directory then this step isn't necessary)
6. Change all occurrences of "changeme.ucsd.edu" to your server's hostname.
7. If using IIS, find the ISAPI section and change the Site ID attribute to the id of the site to protect. (You can find your site's ID in the management console)
8. By default Shibboleth protects the virtual path "/secure". If you would like to protect a different path you can change this in the RequestMapper.
9. Verify that the CredentialsProvider or CredentialsResolver has the correct file names for your certificate and private key. Shibboleth 2.x installs a keypair named sp-cert.pem and sp-key.pem.
10. Start the shibd service or daemon. In Windows, this will be in the Services control panel and will be called "Shibboleth 2.x Daemon." If shibd is already running restart it.
11. If you are using Apache, you must modify your httpd.conf as described at http://spaces.internet2.edu/display/SHIB/SPApacheConfiguration. (The instructions are for version

1.3 of the software but still work with version 2.x). You will also need to Include etc\shibboleth\apache22.config (or the correct file for your version of Apache) which is in the Shibboleth install folder. After doing so, edit it to protect the virtual path where your application resides. By default it only protects the /secure virtual path.

12. Restart Apache or IIS.
13. Test that when you try to access any file under /secure (or whatever path you protected) you are redirected to your campus' IdP. You may need to make a secure directory under htdocs or wwwroot and place a test file in there. You should receive an error message on your campus' IdP saying, "The application that you've attempted to access is not an authorized Single Sign-on application."
14. Register your site by sending an email to your campus' IdP support staff with the URL of your application, the Shibboleth version you installed, the type of authentication you want to use (i.e. Business Systems, Student, Active Directory, Network Username), and your support contacts. Your campus may also require more registration information.
15. Test a page under /secure again to make sure you get the SSO login form you expect. Log in and check for any errors.
16. Join you campus' IdP support mailing list.

Installation (Java/J2EE)

It is strongly recommended that Java application servers be deployed behind Apache httpd using a connector such as mod_proxy_ajp or mod_jk. Refer to https://spaces.internet2.edu/display/SHIB2/NativeSPJavaInstall and follow the Apache installation instructions on this page.

Troubleshooting

Sometimes after following the installation steps Shibboleth doesn't appear to do anything. This page https://spaces.internet2.edu/display/SHIB/SPWontProtect is a good resource for dealing with the problem. Once you get Shibboleth running, you may run into some other common errors which are listed here https://spaces.internet2.edu/display/SHIB2/NativeSPTroubleshootingCommonErrors.

Production deployment issues

Logout: Shibboleth generates a local logout URL for you at /Shibboleth.sso/Logout.

Load balancer / reverse proxy: When setting up an SP behind another system which proxies requests, refer to https://spaces.internet2.edu/display/SHIB/SPReverseProxy.

Load balancer / hardware SSL: If your load balancer or other hardware handles SSL on behalf of your web server, refer to https://spaces.internet2.edu/display/SHIB/SPNoSSL for special configuration steps.

Shibboleth Attributes

In the U.S. the InCommon Federation guides most Shibboleth deployment in the higher education community. InCommon establishes a set of commonly used and supported attributes, drawn largely from the "EduPerson" namespace. An overview of those attributes is available at http://www.incommonfederation.org/attributesummary.html. When registering your service provider you may choose from this list, or negotiate with your campus identity manager the list of attributes

available and needed for your application. Campus identity managers can support many standard EduPerson and UCTrust attributes and often have campus specific attributes that can be released internally as well (which can sometimes be mapped to standard attributes).Then after a user authenticates with the campus identity provider, those attributes regarding that user are provided to the application. This can be useful for authorization decisions; for example, if you only want to allow users from a specific department, an application can check the user's HOME_DEPT_CODE attribute.

Appendix 7 - Sample Shibboleth-related policy – CDL Technical Requirements for E-Journal Vendors

(http://www.cdlib.org/gateways/vendors/guidelines_technical.html)

Note: This requirements document is currently under review. Both a current version which mentions Shibboleth, and a potential revised version which declares a commitment to and strong preference for Shibboleth, quoteing from the overarching statement, are shown below.

Current version:

[section 9] Access authentication

Access should be designed to allow our licensed user community to get to the resource from anywhere with a minimum of effort on our part or that of the user, and with minimal disclosure of identity information. The legacy method of authentication uses IP addresses. Vendor systems should be able to accommodate IP-access authentication via campus proxy servers or VPN. Special requirements for access via campus proxy servers should be well documented. It is important that we have the option to authenticate either once per session or at every document accessed.
The CDL will provide an initial list of IP addresses for the UC community, with quarterly updates. The list indicates which addresses represent proxy servers. We require that vendors notify us via our CDL Support team list when the IP addresses list has been activated or updated so that we can begin testing to ensure that access is working. We do not announce a new resource to our user community until this testing is complete. Delays and problems in activation or updates will be taken into account when UC makes decisions on new products or renewals. Because the IP method is labor intensive, error prone, and often frustrating for our users, we are actively seeking new solutions, particularly those that stress federated identity management and privacy protection. For more information, see Appendix I: Authentication below.
Links cited: CDL Support team list email address: cdlsupport-l@ucop.edu

Appendix I: Authentication

The future of authentication at UC

A flexible authentication mechanism is crucial to UC's plans to enhance library services, and the choice of vendors for licensed electronic resources must reflect this priority. Our current system employs a large network address space with multiple domains, managed in a distributed fashion; each campus Network Operation Center (NOC) registers its own address space with InterNIC and manages those addresses locally. In addition, some campus units choose to contract with external ISPs for specialized services such as modem pools for dial-up and proxy services. These services require the use of IP addresses that are not registered to UC in the InterNIC database, but are dedicated to use by UC faculty, staff, students, and library patrons. Not surprisingly, UC has found this method increasingly cumbersome, and we are now seeking alternative mechanisms for authenticating valid users, wherever they may be located.

UC's Preferred Solution for Authentication

UC is seeking to implement access control mechanisms that simplify the authentication protocols that we employ in support of our enterprise. Methods for authenticating users should facilitate access by authorized users no matter where they are physically located. Access to products should not require individual passwords or user IDs, and UC must be allowed to use proxy servers to allow remote access for authorized users when necessary. Vendors who persist in requiring individual passwords, user IDs, or IP address authentication to access their products will accordingly appear significantly less attractive as the university weighs its options.

The most promising alternative to IP-based authentication is the Shibboleth protocol. A project of Internet2, Shibboleth is developing architectures, policy structures, practical technologies, and an open source implementation to support inter-institutional sharing of web resources subject to access controls. In addition, Shibboleth is developing a policy framework called InCommon that will allow federated inter-operation amongst the higher education community and the vendors that market online services to the community. UC is committed to compliance with the access security policy framework promoted by the InCommon Federation, and would expect vendors to make a similar commitment, ensuring an appropriate level of security for licensed resources, and privacy for personal information about users.

The CDL preferentially contracts with vendors who are working to implement resource access through the use of Shibboleth. Many vendors have implementation plans in place, and significant community resources exist to facilitate the migration to Shibboleth. Shibboleth software is standards-based and open source, so there are no license fees, and the software itself is not difficult to install and maintain. For more information on UC's implementation of Shibboleth, UCTrust, see UCTrust: The University of California Identity Management Federation.

Shibboleth: Benefits to Vendors

Recent trends in telecommuting, distance education, and the globalization of scholarship suggest the university's need to accommodate remote users will grow exponentially in the coming years. The early adoption of Shibboleth may very well preempt large-scale access problems as your clientele--and the sophistication of its access needs--grows.

Shibboleth allows currently valid users to access resources regardless of their physical location. At the same time, the protocol provides vendors with a more authoritative and up-to-date assurance that the user is a verified member of the UC community, which in turn makes it easier to identify and exclude users whose status has lapsed.
Once ubiquitous, the use of Shibboleth will undoubtedly prove a more cost-effective and efficient means of validating users' status, and will relieve both parties of the need to maintain the extensive IP address tables.

URLs and links cited:
Shibboleth Project: http://shibboleth.internet2.edu/
Shibboleth target server: http://shibboleth.internet2.edu/guides/deploy-guide-target1.2.html
UCTrust: The University of California Identity Management Federation: http://www.ucop.edu/irc/itlc/uctrust/

Potential Revision:

Authentication

Access should be designed to allow our licensed user community to get to the resource from anywhere with a minimum of effort on our part or that of the user, and with minimal disclosure of identity information.

Consistent Access Mechanisms

Users should not be presented with a login/password screen when access is controlled by other means, e.g., IP address, or when when trusted authentication has taken place and can be passed on in a trusted authentication federation such as InCommon via Shibboleth.

IP-based access

The CDL currently provides access to most of our licensed resources using IP address authentication. For more information about how IP addresses are updated, please see <x> - <x>.Because the IP method is labor intensive, error prone, and often frustrating for our users, we are actively investigating new solutions, particularly those that stress federated identity management and privacy protection such as Shibboleth.

Shibboleth

Recent trends in telecommuting, distance education, and the globalization of scholarship suggest that the university's need to accommodate remote users will grow exponentially in the coming years. The adoption of Shibboleth may very well preempt large-scale access problems as your clientele -- and the sophistication of its access needs -- grows. ***Therefore the University of California libraries have adopted Shibboleth as the primary authentication standard for access to our research resources and services. In addition to providing a better experience for our users through the use of a single username and password, Shibboleth simplifies secure authentication management and builds stronger partnerships between UC, its vendors, and the wider academic community.***

Shibboleth allows currently valid users to access resources regardless of their physical location. At the same time, the protocol provides vendors with a more authoritative and up-to-date assurance that the user is a verified member of the UC community, which in turn makes it easier to identify and exclude users whose status has lapsed.

Once it is ubiquitous, the use of Shibboleth is expected to prove a more cost-effective and efficient means of validating users' status, and will relieve both parties of the need to maintain extensive IP address tables.

UC Trust is a collection of campus IT representatives that are currently working on issues related to systemwide implementation of Shibboleth. The UC Trust, in conjunction with the InCommon Library Shibboleth project, is looking at best practices for implementing best practices for using Shibboleth to enable a seamless user experience when accessing library resources.

The University of California campuses are members of InCommon and prefer to work with vendors that are also members of InCommon.

References

Shibboleth Project: http://shibboleth.internet2.edu/
Library Shibboleth Project: https://spaces.internet2.edu/display/inclibrary/InC-Library
The UK Access Management Foundation (JISC): http://www.ukfederation.org.uk/
UCTrust: The University of California Identity Management Federation: http://www.ucop.edu/irc/itlc/uctrust/

Remote Access

Vendor systems should be able to accommodate IP-access authentication via campus proxy servers (traditional or rewrite) or VPN client software. Special requirements for access via campus proxy servers should be well documented

References

http://www.cdlib.org/services/info_services/guides/off_campus_access.html

Rewrite Proxies / WebVPN

Rewrite proxies have gained a substantial following at campuses as a way to provide off-campus access to resources without requiring the user to install client software or make configuration changes, which makes it ideal for use in environments where the user has no authority to make configuration changes to the machine, and also reduces the likelihood of user errors made during configuration.

In a nutshell, rewrite proxies route activity through the proxy server by prepending additional information to the default URL. For most resources, this works well, however, resources that are heavily reliant on scripted functionality, contain a large number of separate objects per page or require installation of client software on the user's machine will have problems when used through a rewrite proxy.

At the University of California in 2010, 8 of the 10 campuses use some form of rewrite proxy as a remote authentication mechanism, and four campuses use it as their sole method for providing access to off-campus users, so resource compatibility with this type of software is essential.

At a minimum, resources should be tested and compatible with the EZProxy software and with Cisco WebVPN to assure compatibility and usability for off-campus users. ***To follow UC's commitement to and strong preference for Shibboleth authentication, resources that are Shibboleth enabled should also provide a WAYFless URL target so that UC's rewrite proxies can create WAYFless URLs for Shibboleth-based authentication per the InCommon LibraryBest Practices.***

References: https://spaces.internt2.edu/display/inclibrary/Best+Practices.

References

EZProxy: http://www.oclc.org/ezproxy/
OCLC overview of rewrite proxies: http://www.oclc.org/support/documentation/ezproxy/rewrite.htm

Appendix 8 – Education and Outreach Issues: Useful Links

Education and outreach for end users

- Shibboleth announced and explained via newsletter: http://www.cer.jhu.edu/e-news/enews02-08.html#4; http://www.tdl.org/shibboleth/
- Shibboleth described among the options for off-campus login to resources: http://www.dartmouth.edu/~library/home/help/off-campus.html#Shib; http://aberdeenuniversityvirtuallibrary.pbworks.com/How-to-set-up-off-campus-access-to-electronic-resources; http://www.cardiff.ac.uk/insrv/eresources/offcampus/index.html
- Shibboleth logins as part of a resource user guide: http://library.duke.edu/services/instruction/refworks/index.html; http://www.lib.uchicago.edu/e/using/bibtools/refworks/librarydatabases.html; http://www.lib.uchicago.edu/e/using/bibtools/refworks/proxy.html

Education and Information for Librarians and Library Staff

- Federated identity management video (from JISC): http://www.jisc.ac.uk/whatwedo/themes/accessmanagement/federation/animation
- Shibboleth demo (Quicktime movie – includes library resources): http://shibboleth.internet2.edu/demo/shib_demo.html
- Vendor descriptions of Shibboleth Access: http://muse.jhu.edu/about/muse/faq.html#access; http://www.hathitrust.org/shibboleth/; http://www.info.sciverse.com/scopus/scopus-services/athens-shibboleth/; http://www.oclc.org/support/documentation/ezproxy/usr/shibboleth.htm