World Library and Information Congress:
70th IFLA General Conference and Council

22-27 August 2004
Buenos Aires, Argentina

Programme: http://www.ifla.org/IV/ifla70/prog04.htm

Code Number: 055-E
Meeting: 155. Information Technology
Simultaneous Interpretation: -

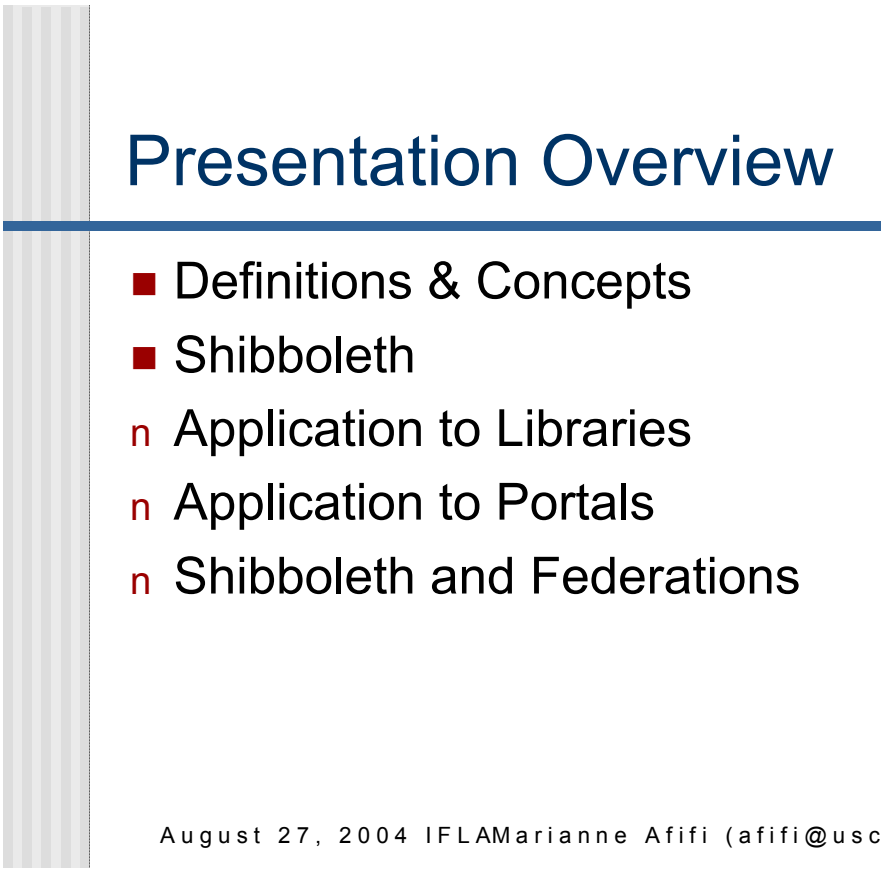
Shibboleth: An Open-Source Solution To Resource Sharing

Marianne Afifi
Director of Electronic Resources & Special Projects Development
Information Services Division, LVL 113C
University of Southern California
650 W 35th Street
Los Angeles, CA 90089-2571
Tel: 213.740.8817
Fax: 213.740.7713
Email: afifi@usc.edu
Web: http://isd.usc.edu/~afifi

ABSTRACT:
Shibboleth is a project of the National Science Foundation Internet2 Middleware Initiative. Its aim is to "develop an open, standards-based solution to the needs for organizations to exchange information about their users in a secure, and privacy-preserving manner."[i] Libraries and academic institutions see an increasing need for providing protected and validated access to online resources they produce, license, purchase, or share. Vendors and publishers are also interested in providing models of access that are different from the traditional IP-based, proxy server and username/password authentication. Researchers and teachers who use the web extensively have rapidly growing requirements for the exchange of information that gives participants certain rights as well as privacy.

Shibboleth was developed to enable resource-sharing interactions for all these constituencies in a simple, standards-based manner.

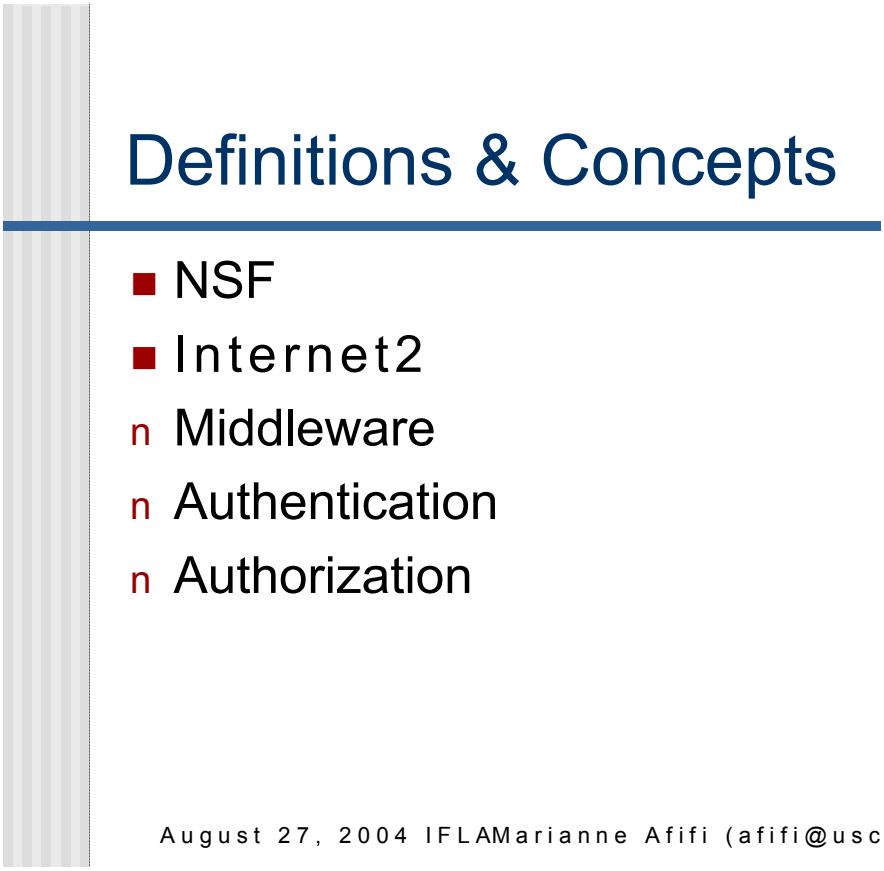
In my presentation I will summarize the origins of Shibboleth, give an overview of its development and explain how it works. I will also explore the reasons why it is an important tool for resource sharing in libraries and federated environments.



I am pleased to speak to you today about an interesting project that has relevance to the theme of "Authentication and Authorization related to Library Services" which have been put forward by IFLA's Information Technology Section as goals for this conference. The project is called Shibboleth. First, I wish to provide first some definitions of terms that I will be using in my presentation. Then, I will give an overview of why Shibboleth was created, how it works, and speak about its relevance to libraries, especially multi-search portal projects, such as the Scholars Portal.

The development of Shibboleth has been supported by many different organizations. Some of you may not be familiar with them and therefore I am listing them here. The National Science Foundation's Internet2

consortium is the sponsor of the project. Middleware exploration is one of the goals put forward by Internet2.



National Science Foundation
was "established by the National Science Foundation Act of 1950, as amended, and related legislation, 42 U.S.C. 1861 et seq., and was given additional authority by the Science and Engineering Equal Opportunities Act (42 U.S.C. 1885), and Title I of the Education for Economic Security Act (20 U.S.C. 3911 to 3922). …"

"The Act established the NSF's mission:
To promote the progress of science; to advance the national health, prosperity, and welfare; and to secure the national defense."[ii]
http://www.nsf.gov/home/about/creation.htm

Internet2
is a consortium being of 206 universities working in partnership with industry and government to develop and deploy advanced network applications and technologies, accelerating the creation of tomorrow's Internet. Internet2 is recreating the partnership among academia, industry

and government that fostered today's Internet in its infancy. The primary goals of Internet2 are to:

- Create a leading edge network capability for the national research community
- Enable revolutionary Internet applications
- Ensure the rapid transfer of new network services and applications to the broader Internet community.

http://www.internet2.edu/about/aboutinternet2.html[iii]

Middleware
or "glue", is a layer of software between the network and the applications. This software provides services such as identification, authentication, authorization, directories, and security. In today's Internet, applications usually have to provide these services themselves, which leads to competing and incompatible standards. By promoting standardization and interoperability, middleware will make advanced network applications much easier to use. The **Internet2 Middleware Initiative (I2-MI)** is working toward the deployment of core middleware services at Internet2 universities. http://middleware.internet2.edu/[iv]

Two definitions that I think are essential to understanding this presentation. They are authentication and authorization. I have selected the definitions used by Clifford Lynch from the Coalition for Networked Information (CNI):

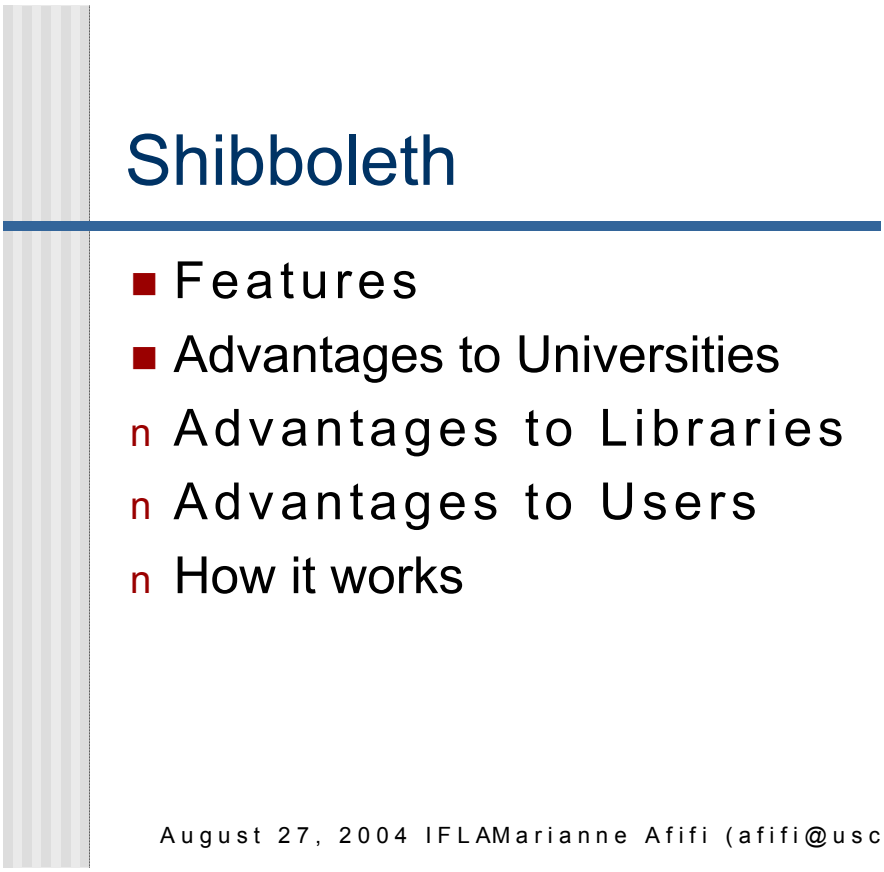
- **Authentication**

"Authentication is the process where a network user establishes a right to an identity -- in essence, the right to use a name."[v]

- **Authorization**

"Authorization is the process of determining whether an identity (plus a set of attributes associated with that identity) is permitted to perform some action, such as accessing a resource."[vi]

These terms are often confused but are important to be distinguished, although currently in library settings, they are often the same. For example, when accessing a resource where access is based on IP validation, users are authenticated by virtue of their IP address, which automatically means that they are also authorized. In a Shibboleth environment users may be authenticated as part of a certain group but may not be authorized to use a specific resource.



Shibboleth
The Oxford English Dictionary, 2nd edition, 1989 gives several definitions of the word Shibboleth, two of which I have selected. They apply to the project we are discussing today.
"1. The Hebrew word used by Jephthah as a test-word by which to distinguish the fleeing Ephraimites (who could not pronounce the *sh)* from his own men the Gileadites (Judges xii. 4-6).
3. *fig.* A catchword or formula adopted by a party or sect, by which their adherents or followers may be discerned, or those not their followers may be excluded."

Thus the word Shibboleth is a fitting name for a process of access management.

The Shibboleth software project arose out of a need for Internet2 participants, and other organizations that have similar structures and requirements, to find a better way to communicate information about themselves and their users. Shibboleth was developed to facilitate communication and information exchange among interested parties such as universities, libraries, government agencies and commercial entities. Because of the diverse nature of Internet2 participants the requirements for

Shibboleth were that it be based on standards, that it be open-source, that it should preserve privacy and that it allows and encourages federation among institutions. We will see later how the federation aspects are especially important in resource sharing. Current communication structures in the arenas of Internet 2 institutions rely on technologies that tend to be complex and labor intensive to administer. Shibboleth enables participants to exchange information without having to use multiple, non-standard, and difficult-to-use authentication and authorization methods. When implemented fully, Shibboleth will enable information flow that is based on standards and preserves privacy for the members of participating institutions. While full implementation among members and in related communities is some time away, many institutions have begun projects to test Shibboleth. In addition, organizations have also set in motion large-scale efforts to create information architectures and structures within their organizations to enable Shibboleth to work with them.

Universities are likely to implement Shibboleth for communication among researchers, in libraries, for portals, and for course management systems (CMS). The connections and the interoperability among all of these systems are very desirable. The importance of Shibboleth for libraries will be apparent to anyone who has worked in areas of electronic resource management. Publisher, vendor and aggregator sites typically require lists of IP addresses or provide username and password access that needs to be managed by the library. Some libraries are required to deploy proxy servers to identify users. Others use VPN services to enable their users to log into the campus network. In addition, a few vendors place rigorous restrictions on resources, for example a librarian in a physical location in a library must enter a username and password for a user. In the 21st century these types of access methods put a burden on the library, institution and users and are increasingly becoming outdated.

As typical university user populations are becoming more diverse, so are library users. Users are no longer only faculty, students, and staff on site. Distributed education, alumni relations, community involvement, and collaborative university activities have created many different classes of users who have a variety of library privileges in a variety of locations but are all connected to the university in some way. It is now difficult to provide and manage access to library and other resources for these populations.

Shibboleth can ease the management of access to resources, but to do so institutions must keep management of user information in a standards-based directory service. The development of such directory services has generally not been based on standards, however efforts have been made towards that goal. One of these efforts is EduPerson, which is "an auxiliary

object class for campus directories designed to facilitate communication among higher education institutions. It consists of a set of data elements or attributes about individuals within higher education, along with recommendations on the syntax and semantics of the data that may be assigned to those attributes"[vii]. Once attributes about users are clearly defined and a standards-based directory service is created and managed over time, access can be managed much more easily. On the user's side revealing attributes about themselves in a Shibboleth environment is more privacy preserving than other current methods of authentication.

The advantage to the library is that resources can be targeted to specific populations, which could potentially be saving costs. For example, access to an expensive journal could be limited to members of a certain research project rather than the whole university population, thus freeing up money for other resources. In addition, the federated aspects of implementing Shibboleth could be advantageous to library consortia, multi-campus universities, universities within a state, collaborative ventures within a federation, and to vendors. From the perspective of vendors using Shibboleth, it allows them to eventually unify the variety of access methods they now have in place. Although the process will take some time, once vendors have implemented Shibboleth, adding new customers that are also Shibboleth enabled will be very easy. Shibboleth will also allow vendors to provide more segmented and targeted customer services and will likely also need to develop different business models.

With the Shibboleth software there is quite a change in the way resources are accessed. Under old models, users are logged into a resource owner's system, for example a publisher's site or OCLC FirstSearch, and must be authenticated there. Shibboleth is designed such that the users' home institution verifies their existence in the institution's database and provides necessary attributes about them to the resource owner. This process is accomplished is as follows (see Figure 1):

A user goes to a resource owner's web site, let's say JSTOR, a scholarly journal archive provider.

1. JSTOR does not know where the user comes from.
2. JSTOR's SHIRE (Shibboleth Indexical Reference Establisher) asks the WAYF (Where Are You From) service about where the user is from.
3. The WAYF asks the user where she is from.
4. The user redirects the question to her home organization's (origin's) Shibboleth Handle Service (HS).
5. The Handle Service looks in the users home organization's directory database to find out whether the user exists in the database.

6. After verification, the Handle Service then sends back a request to the user to send her credentials.
7. The user sends her credentials to the Handle Service
8. The Handle Service sends back a handle (or alias) to JSTOR's SHIRE indicating that the user is part of the origin site.
9. The SHIRE then forwards the handle to the SHAR (Shibboleth Attribute Requester), which asks for attributes about the user and sends a query to the Attribute Authority (AA) located at the origin site.
10. The Attribute Authority releases the user's attributes to the SHAR.
11. The SHAR releases the user attributes to JSTOR's resource manager (RM) that then allows the user the appropriate access to the resource.

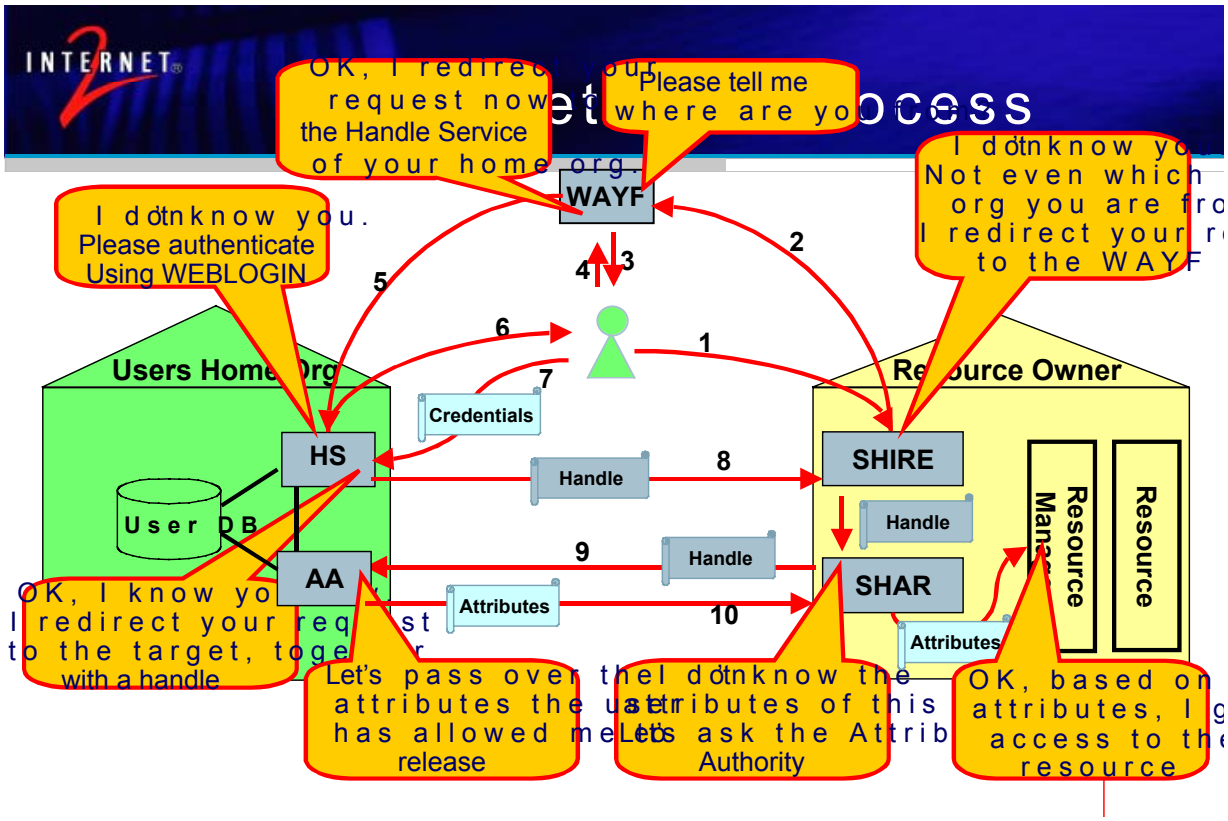


Figure 1. Diagram courtesy of Dr. Ken Klingenstein, Internet2 from a presentation given at the Coalition of Networked Information (CNI) Task Force Meeting in April 2004 based on a SWITCH (Swiss Education and Research Network) demo.

Although the description of what goes on behind the scenes may seem complex, the interaction between the various services is all provided within the web environment and provides speedy and seamless access to the resources.

Application to Libraries & Portals

- Resource management
- Attribute management
- User segmentation
- Portals as targets
- Portals as users
- Scholars Portal

August 27, 2004 IFLAMarianne Afifi (afifi@usc.edu) 4

Two key concepts emerge from the diagram: the origin and the target. These concepts are important because, as we will see portals can play either role. The origin is where the Handles Server and Attribute Authority reside and the target is the resource. Both the Handle Server and the Attribute Authority rely on data that is stored in the origin's (the institution's) user database that manages and controls the information about the user. This information is targeted to what the resource needs to know without revealing information it does not need to know, as is now the case with other authentication methods. The interaction preserves the privacy of the user since the resource manager only needs to know certain attributes about the user, whereas with current authentication methods more information may be revealed about the user than is necessary.

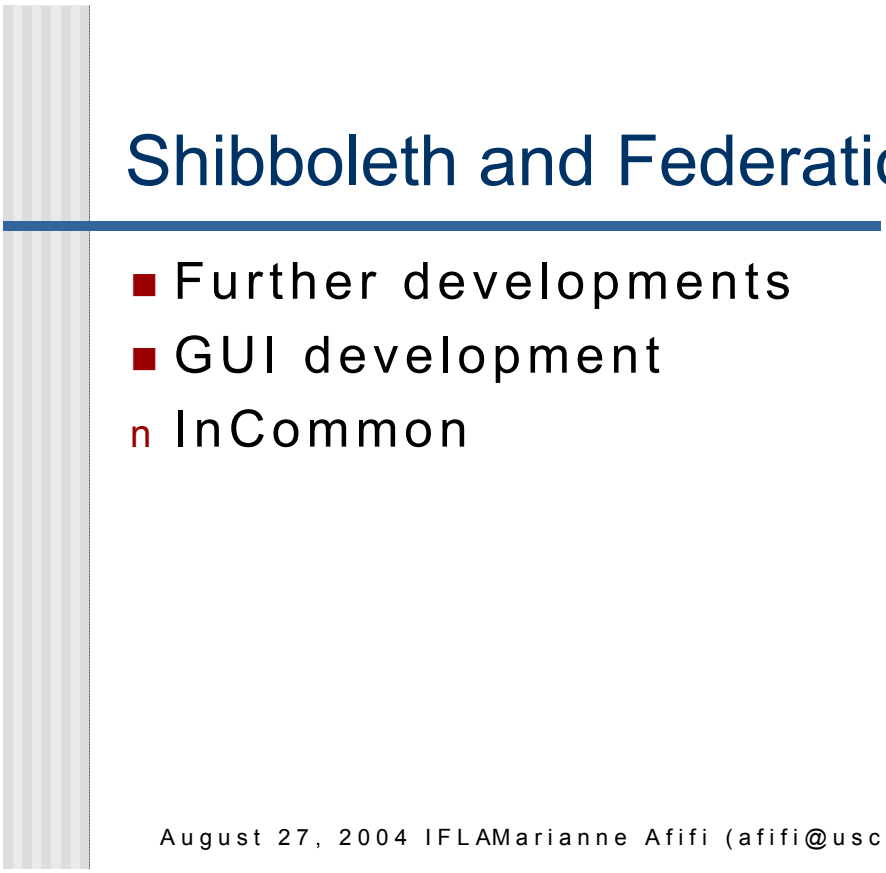
Shibboleth and Portals

The characteristics of a portal are that information and data are retrieved from various systems and sources and combined in a web environment such that information particular to a user group or an individual user is presented either automatically or on demand. In addition, users are being offered services such as e-mail, chat, and news by the portal. Libraries have increasingly implemented these portals to provide one-stop shopping for

library resources and services such as subject based searches, metasearches, and online reference services.

Don Gourley from the Washington Research Library Consortium [viii] has recently written an article about Shibboleth and library portals and presented some possible scenarios and roles for such portals. In his article he proposes the library portal as a Shibboleth target.

One example of a portal that is used as a target is the Scholars Portal. The Scholars Portal is an ARL project involving 7 U.S. universities developing a metasearch portal together with a vendor, Fretwell Downing[ix]. Of the seven universities none used the same way of authenticating and authorizing its users. Fretwell Downing has been working with the various institutions to accommodate their authentication needs for the portal users. The University of Southern California (USC), a member of Internet2 and an implementer of Shibboleth, was interested in using it for the Scholars Portal. Implementation is currently underway, specifically USC requires the Scholars Portal (ZPortal by Fretwell Downing) to be a target, which means that ZPortal must accommodate the SHIRE, SHAR and RM and be in contact with USC's origin service. Therefore users who want to use the Scholars Portal's personalization features will use Shibboleth to do so. However, resource access from Scholars Portal will still be managed the same way as before (largely IP authentication with off-campus access through a Virtual Private Network (VPN)). Only when Zportal itself becomes the origin, and all or most of USC's vendors will be able to implement target services, will we be able to access the majority of our resources in the way Shibboleth intends.



Shibboleth and Federation
The Shibboleth project began in 2000 and Shibboleth v.1.1 was released in 2003 and is currently being implemented in an ever-increasing number of universities and vendors. Work is being continued on the next version, which is due to be released this year. Future upgrades will include portal support in addition to planned other improvements. Currently, a Shibboleth GUI Development Focus Group is participating in providing requirements and feedback about the creation of a Shibboleth Attribute Release Policy (ARP) Editor (SHARPE). This editor will enable libraries, within a federation, to manage attributes of resources and services they currently manage.

To take advantage of Shibboleth, federated environments must be created among participating institutions that establish a trust infrastructure so that the management and access to resources is greatly improved. To this end InCommon was formed and is also a project of Internet2. "InCommon is a formal federation of organizations focused on creating a common framework for trust in support of research and education. The primary purpose of the federation is to facilitate collaboration through the sharing of protected resources, by means of an agreed common trust fabric."[x] While there are only few extant federations and the process of forming them is still in its

infancy, we hope that more and more of them will be created since this model of trust promises to facilitate access and preserve privacy for users.

Shibboleth will change how various communities interact with each other on the network and how new forms of resources can be accessed. It is easy to imagine that new patterns of communication among institutions will emerge in a private and secure environment.

More Information

- Shibboleth
  - http://shibboleth.internet2.edu//
- SWITCH Authentication and Authorization Infrastructure(AAI)
  - http://www.switch.ch/aai/
- Scholars Portal—
  - http://www.arl.org/access/scholarsportal/
- InCommon—
  - http://www.incommonfederation.org/index.cfm
- Contact: Marianne Afifi—afifi@usc.edu

August 27, 2004 IFLAMarianne Afifi (afifi@usc.edu) 6

[i] Shibboleth Frequently Asked Questions (FAQ)
http://shibboleth.internet2.edu/shib-faq.html#001
Nate Klingenstein
4 May, 2002
Retrieved 2/10/04
[ii] NSF Creation and Mission
http://www.nsf.gov/home/about/creation.htm
Retrieved 4/15/04
[iii] About Internet2®
http://www.internet2.edu/about/aboutinternet2.html
21 April 2004, Retrieved April 21, 2004
[iv] Internet2 web site

http://middleware.internet2.edu/
Retrieved March 30, 2004
[v] Cross-organizational Use of Networked Information Resources
Clifford Lynch, editor (cliff@cni.org)
Coalition for Networked Information
Revised Discussion Draft of April 14, 1998
http://www.cni.org/projects/authentication/authentication-wp.html
Retrieved April 21, 2004
[vi] Ibid.
[vii] EduPerson Specification (200312)
http://www.nmi-edit.org/eduPerson/internet2-mace-dir-eduperson-200312.html
Retrieved May 21, 2004
[viii] Library Portal Roles in a Shibboleth Federation Don Gourley, Washington Research Library Consortium (WRLC), 30 October 2003
http://shibboleth.internet2.edu/docs/gourley-shibboleth-library-portals-200310.html
Retrieved April 21, 2004
[ix] The Scholars Portal Project. http://www.arl.org/access/scholarsportal/
Retrieved April 21, 2004
[x] InCommon Home Page
http://www.incommonfederation.org/index.cfm
Retrieved April 15, 2004

Back to the Programme: http://www.ifla.org/IV/ifla70/prog04.htm