Technical Implementation of Shibboleth-based Access Management for the N-LIST Programme

Yatrik Patel[1] and Jagdish Arora[2]
[1]Scientist C, [2]Director, INFLIBNET Centre, Ahmadabad

Abstract

Information and Library Network Centre, an Inter University Centre of University Grants Commission, acts as a gateway to scholarly content for Indian academic and research community. Centre also provides access to electronic resources to universities and colleges across India through its UGC-INFONET Digital Library Consortium and N-LIST Programme. INFLIBNET Centre is implementing state of the art, standard based open source technological solutions towards for extending access to e-resource to users irrespective of their physical location. Shibboleth is such standards-based, open source software system that facilitates single sign-on web-based access to electronic resources across or within organizational boundaries. This paper discusses advantages, customization and implementation of Shibboleth-based Access Management Solution for its multiple national award winner N-LIST (National Library and Information Services Infrastructure for Scholarly Content) programme.

1. Introduction

The INFLIBNET Centre, as one of its core mandates, provides access to scholarly content to universities and colleges in India under the UGC–INFONET Digital Library Consortium and N-LIST Programme. The Centre has taken steps to optimize the utilization of e-resources so as to ensure better Returns on Investment (RoI) and greater benefits to the academic community. At present, access to e-resources in universities is IP-authenticated and, as such, access is restricted within the confine of a given university campus due to lack of proper authentication mechanism. The Centre is working towards deploying appropriate access management tools, enabling users to access e-resources from his / her campus, home or even while travelling. Implementation of such a solution requires setting-up of proper user authentication and access control mechanism ensuring trust relationship between publishers, identity providing and authenticating agency and the user institution. Centre is working towards implementation of Shibboleth for its all consortium resources. The Shibboleth is standard-based open source middleware software that provides Web-based single sign-on (SSO) access to subscribed e-resources across or within organizational boundaries so as to enable users to access e-resources from anywhere irrespective of his / her physical location.

The project entitled "National Library and Information Services Infrastructure for Scholarly Content (N-LIST)" being jointly executed by the UGC-INFONET Digital Library Consortium, INFLIBNET Centre and the INDEST-AICTE Consortium, IIT Delhi, provides for access to scholarly content to colleges, universities as well as centrally-funded technical institutions. The project provides for cross-subscription of e-resources wherein selected e-resources subscribed under the UGC-INFONET Digital Library Consortium are made accessible to technical institutions (IITs, IISc,

IISERs, NITs, etc.), likewise, selected e-resources subscribed under the INDEST-AICTE Consortium are made accessible to the universities. Besides, e-resources including more than 2,100 electronic journals and 51,000 electronic books are made accessible to Govt./ Govt.-aided colleges. The INFLIBNET Centre acts as National Monitoring Agency to manage and monitor access, and impart training to promote optimal usage of e-resources in colleges as well as to monitor all activities involved in the process of providing effective and efficient access to e-resources to colleges. The Centre is also responsible for developing and deploying appropriate software tools and techniques for authenticating authorized users so as to enable them to access e-resources from anywhere, anytime (INFLIBNET, 2010).

As on January 30th 2011, N-LIST programme, formally launched by the Honourable Union Minister of Human Resource Development, Shri Kapil Sibal on 4th May, 2010 at Shastri Bhawan, New Delhi, has registered a total number of 1,610 colleges including 1,087 eligible colleges that are already getting access to resources subscribed under the N-LIST programme (NLIST Web site, 2011). Log-in IDs and passwords have been issued to more than 98,500 faculty members, students and researchers after obtaining list of authorized users from these 1,087 registered colleges. Group login ID and passwords have also been issued to colleges awaiting list of actual users. Colleges that are not recognized under 2(F) and 12(B) Section of UGC Act are being advised to join the initiative as N-LIST Associates. Efforts are being made to enrol more colleges through advertisement in newspaper as well as by organizing training and orientation programmes through affiliating universities and regional offices of the UGC.

2. Reaching out to Govt. / Govt.-aided Colleges: Access Methodology

E-resources identified for cross subscription for universities and technical institutions are being made accessible to the beneficiary institutions on their respective IPs. Given the fact that majority of colleges that are registered under the N-LIST programme, do not have static IP addresses, access to e-resource based on IP filtering cannot be used for colleges. Moreover, the Government of India under its National Mission on Education through ICT has launched a programme wherein each college is being given five Mbps broadband Internet connectivity without static IP addresses. Further, colleges had to be allowed to enrol themselves for NLIST programme anytime during the year as well as during the project period.

With the background given above, most publishers were reluctant to handle authentication for users from more than 6,000 colleges. As such an innovative model had to be involved to reach out to all the targeted colleges. As such, the INFLIBNET Centre took the responsibility of providing access to e-resources to all the registered colleges through proxy servers deploying appropriate authentication and authorization mechanism. It had to be ensured that individuals (including students, researchers, staff and faculty) from colleges and other beneficiary institutions should have direct access to e-resources with facility to download articles from journals and chapters from books directly from publisher's website once they are duly authenticated as authorized users through the authentication mechanism deployed at the INFLIBNET.

3. EZProxy: Implementation and Functioning

While, the Centre is working towards implementation of Shibboleth-based access management system which is time-consuming process requiring intensive software configuration, cooperation and collaboration amongst participating institutions, publishers and recognition of INFLIBNET Centre as an Identity Provider (IDP) by the publishers as Service Provider (SP), the Centre has implemented and configured EZProxy from OCLC to facilitate access to e-resources to students, researchers and faculty in colleges.

EZproxy is a user-friendly, web-based commercial solution from OCLC (OCLC, 2010). It is used by a large number of education institutions across the world as a mechanism to authenticate authorised users, enabling them to access subscribed e-resources irrespective of their physical location. It works by rewriting the URL of a desired web page in such a way that it appears as though the request originates from server having a static IP addresses recognized by the publisher, even though the user is actually off-campus. One of the advantages of EZproxy is that it provides a simple mechanism that allows institutions to create one set of web resources that can be used by both on-site and off-site users to gain access to licensed databases.

3.1. Limitation of EZProxy

One of the main disadvantages of EZproxy is the maintenance required in terms of managing the database of URLs (i.e. websites of subscribed resources) that EZproxy is allowed to visit. This is particularly bothersome for libraries that subscribe to thousands of journal titles from a variety of publishers and vendors. Moreover, when a large number of users logs-on through the EZProxy, it leads to an overload on the proxy server slowing down of the Internet access for all proxy users. Furthermore, the proxy server generally does not work when users are behind firewalls, because firewalls block the port used by the proxy server by default (Simpson, M. et al.) and needs to be specifically configured to allow access via EZProxy.

4. Shibboleth-based User's Authentication: Proposed Model

Considering the limitations of EZ Proxy, it was decided to implement Shibboleth-based user's authentication system to authenticate authorized users from colleges registered under the N-LIST progamme. By default, Shibboleth works in a federated mode using organization's internal identity and access management system for authenticating users. However, in cases of colleges, the INFLIBNET Centre will have to take the responsibility of authenticating users from all 6,000 colleges, since neither these colleges nor their affiliating universities, are equipped or have technical capabilities to run their own Shibboleth-based authentication system. The INFLIBNET Centre would, therefore, be running Identity Provider Service (IDPs) that would virtually serve all 6,000 colleges. Since the Shibboleth model is designed to run Identity Provider Services for individual institutions, necessary changes will have to be made in the Shibboleth Identity Provider software in order to accommodate one Identity Provider Service virtually serving 6,000 identity providers, one each for 6,000 colleges (Patel, Y., 2010).

4.1. What is Shibboleth?

Shibboleth system (Shibboleth Web Site, 2011) is standards-based, open source software system that facilitates single sign-on web-based access to electronic resources across or within organizational boundaries. Shibboleth system, on one hand allows institutions to authenticate authorized users using organization's internal identity and access management system enabling them to access subscription-based electronic resources in a privacy-preserving manner. On the other hand, it allows protected or subscription-based Web sites to make informed decisions about users that are duly authenticated by their respective institutions, and enable access to its e-resources based on their entitlement as defined in the authentication attributes passed on by their institutions in the process of authentication. Using this technology, user can access designated electronic resources within the confine of their institute campuses as well as off campus.

The Shibboleth software implements widely used federated identity standards, principally OASIS' (Organization for the Advancement of Structured Information Standards) Security Assertion Markup Language (SAML), to provide a federated single sign-on and attribute exchange framework. Shibboleth also provides extended privacy functionality allowing the user and their home site to control the attributes released to each application. Using Shibboleth-enabled access simplifies management of identity and permissions for organizations supporting users and applications. Shibboleth is developed in an open and participatory environment and is freely available.

In addition to providing single sign-on functionality, Shibboleth can help control access to either campus-based or licensed resources. Working with identity management systems, Shibboleth releases the information for which service providers need to authorize actions or customize the user's experience. This reduces the need for developers to have access to the directory and instead provides fresh data, just-in-time. This can be implemented on and off-site.

Shibboleth can effectively address the challenges mentioned below that are encountered by content providers as well as institutions as identity providers:

- Requirement of multiple passwords for multiple applications: Shibboleth supports single-sign on functionality;

- Scaling the access management for multiple applications: Most e-resources are Shibboleth compliant. Shibboleth allows institutions to use their existing access management system to authenticate users. For example, an institute having a mail services, can use the same authentication system for Shibboleth that is deployed for users to access their e-mails;

- Security issues associated with accessing third-party services privacy: Shibboleth uses Security Assertion Markup Language (SAML) and digital certificates for transfer of data;

- Interoperability within and across organizational boundaries: Shibboleth is open source software;

- Enabling institutions to choose their authentication technology: Shibboleth facilitates use of authentication mechanism that already exists in an organization;

- Enabling service providers to control access to their resources: Shibboleth's Identity Provider interface at user's end interact with the Shibboleth's Service Provider's interface at the publisher's end;

- Overloading of intermediate server used for authentication: Unlike proxy server programme, Shibboleth authentication works by setting-up web cookies both at user's end as well as at publisher's end (SP). As such, role of intermediate server is over as soon as a user is authenticated; and

- Publisher need to have an updated database of authorized users for each subscribing institution: An individual uses his or her campus login and password to get authenticated from his / her own organization so as to access e-resources subscribed by the parent institution. University can use their default authentication mechanism. Shibboleth sits on top and provides the web single sign-on functionality.

4.2. Components of Shibboleth

Shibboleth consists of two major components, i.e. Identity Provider (IDP) and Service Provider (SP) that trusts each other.

i) **Identity Provider**: Identity Provider software is run by the institutions having database of its authorized users. Shibboleth leverages the organization's identity and access management system, so that the individual's relationship with the institution can be used to determine access rights to subscription based e-resources or services. In other words, different categories of users in an institution may have access to different sets of resources based on attributes assigned to them. In case of Shibboleth implementation at INFLIBNET, participating universities and colleges would register themselves for creation of an identity management system using college administrative interface. A trusted officer, nominated by the college authority, will be responsible for maintaining identity management system for a given college on the INFLIBNET Server.

ii) **Service Provider**: Service Provider software is run by the publisher of a subscription-based e-resource or services. The Service Provider receives a set of pre-defined attributes from the Identity Provider and provides access to subscribed e-resources or services to the user depending upon the attributes received.

The INFLIBNET Centre, as Shibboleth Federation, would manage the trust between all the parties and host database of authorized users at servers installed at the INFLIBNET Centre. As a result, when a user wants to access a subscribed resource, he / she would be directed to the Identity Provider Service (IDP) at the INFLIBNET Centre to get himself/herself authenticated. The IDP at the INFLIBNET Centre, in turn, would pass requisite attributes of a user to the Service Provider using associated user's database from colleges. In other words, the Service Provider will receive all necessary attributes for a user from the INFLIBNET's Identity Provider which it

trusts. Attributes passed on by the IDP at INFLIBNET will determine level of access a user gets from the Service Provider. Figure 1 given below illustrates the functioning of Shibboleth.

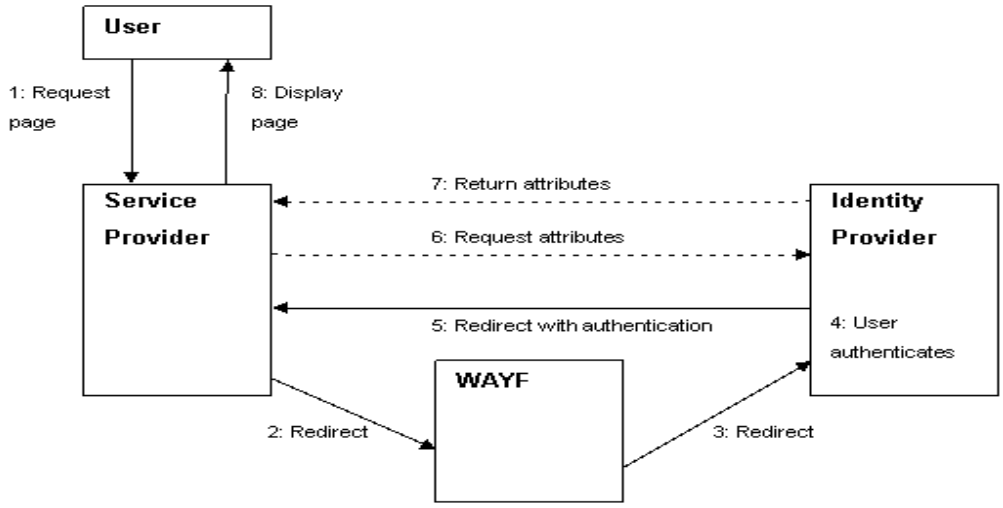

Fig. 1: Authentication of Users through Shibboleth
Image courtesy (http://www.jisc.ac.uk/)

4.3. Shibboleth: Features and Functionalities

Shibboleth serves as an access management tool for providing controlled access to subscription-based resources or licensed resources. It enables interaction between Identity Provider (IDP) and Service Provider (SP), wherein Service Provider (SP) seeks authentication information from IDP that it requires to authorize access to subscribed e-resource to a user. In effect, Shibboleth eliminates the requirement for developers and publishers of e-resources to develop maintain and access directory of authorized users. Instead, it facilitates use of authentication datasets available with the subscribing institutions to get up-to-date authentication data from the primary source as trusted partner. Important features and functionalities of Shibboleth system are described below.

4.3.1. Single sign-on

Most web-based applications have their own authentication system, wherein each user of that application is issued a Login ID and password to access that application. Likewise, owners of protected or subscription-based web sites are issued Log-in IDs and passwords to access these protected or subscription-based resources. As such, a typical user is likely to have multiple Login IDs and passwords to access Institute Intranet, local PC, Bank Accounts, E-mail Account, Virtual Learning resources, subscription-based academic journals, etc.

Multiplicity of Login IDs and passwords are cumbersome and difficult to manage for people. It confuses the users and creates problems both for users as well as for service providers. The purpose of a single sign-on system is twofold:

- To allow the same Log-in ID to be used to access multiple online resources; and

- To allow users to navigate from one resource to another without having to re-enter the Log-in ID and password.

The principle objective of Shibboleth is to allow an organisation to have a single set of Log-in IDs and passwords for each user to access multiple online resources, whether local or external, available to authorized users in an organisation. Shibboleth system allows organisation itself to authenticate its own user by deploying its own existing authentication mechanism. Shibboleth does not pre-ordain the Identity Provider to use any specific authentication method. The Shibboleth software implements widely used federated identity standards, principally OASIS' Security Assertion Markup Language (SAML), to provide a federated single sign-on and attribute exchange framework.

4.3.2. Attributes

An individual organisation, as a member of Shibboleth Access Management Federation, acts as an Identity Provider and provides requisite attributes for each authorized users to the publishers of subscribed resources. Attributes passed on to the publisher of a subscribed resources includes: department to which a user belongs, his / her role (i.e. student, faculty, staff), his / her entitlement to access a given resources, etc. The organisation, as an Identity Provider, also defines their Attribute Release Policy (ARP) so that the administrator can choose which attributes are to be released to which online resources. The Shibboleth IDP also provides an Attribute Authority (AA) which can be used to retrieve attributes from various sources, such as LDAP directories, databases and files.

4.3.3. Privacy of Individual Users

The architecture of Shibboleth enforces the concept of individual privacy, allowing users to have one-time session identifier and no persistent identity visible outside the organisation. Individual privacy is also enforced by the concept of the Attribute Release policy which is designed to allow the user to restrict the release of attributes to third parties. Interface to manage and enforce the Attribute Release Polices are not yet available but can be configured by doing appropriate configurations in xml files available in configuration folder. Use of Shibboleth-enabled access, simplifies management of identity and permissions for organizations supporting users and applications.

4.3.4. Federation

Shibboleth access management model is essentially designed to run in a federated mode wherein individual participating institutions are required to run their own Identity Provider Services for users in their respective institutions. A formal federation is required as trusted interface between the institutions and publishers / service providers. The federation is required to provide a list of participating organisations, with details of the registered Shibboleth components for that organisation to the publisher (service provider). This list is made available to users wishing to access resources registered with the federation, to allow them to navigate

to their home organisation and get themselves authenticated. This list is known as the Where Are You From (WAYF) service.

Major role of the federation includes development of federation, participating community, provide assistance to universities and colleges to create and maintain their identity management system, assigning responsibilities to the trusted officers of the organizations for maintenance of database of users and to manage their identity, processing of participant metadata, overseeing operations of Shibboleth Service platforms, dispute resolution, etc.

4.3.5. Shibboleth @ INFLIBNET

The Shibboleth working architecture described above requires each participating institute to set-up their own identity provider services. Looking at the present scenario, universities and colleges do not have requisite technical know-how and ICT infrastructure, INFLIBNET would, therefore, act as an IDP for all the institutions, including universities and colleges under its umbrella. Shibboleth implementation at INFLIBNET will be suitably amended as mentioned below:

i) Users from N-LIST or UGC-INFONET Digital Library Consortium visits the publisher's web site to access e-resources;

ii) The resource redirects the user to WAYF, so that he/she can select his / her home organisation. The service provider (publisher) would recognize INFLIBNET Centre as a trusted organization for authenticating user, and would give an option on Web site of their resource to select INFLIBNET as an Identity Provider Service. Since INFLIBNET will serve as an IDP for all its member universities and colleges, individual institutions would not be require to set-up their separate IDP and publisher would not be required to maintain separate link for each institution;

iii) When a user select INFLIBNET Federation for his / her authentication, he / she would be re-directed to IDP link at INFLIBNET Server;

iv) After verifying user's credentials, IDP at INFLIBNET will pass "user attributes" which may also contain his / her institute, department, role (faculty/student/researcher), and if agreed, whether he / she is having access to a given resource or not and / or any other attributes which are mutually agreed;

v) After successful authentication, a one-time handle or session identifier is generated for a session, and the user is returned to the resource;

vi) The resource uses the handle to request attribute information from the Identity Provider for this user. The organisation, as IDP, allows or denies the attribute information to be made available to this resource using the Attribute Release policy; and

vii) Based on attributes associated with the authenticated user, he / she would be allowed access to e-resources.

4.4. Customization of Shibboleth System

The Shibboleth working architecture requires that each participating institute should set-up their own Identity Provider Services (IDP). Given the fact that most universities and colleges do not have requisite technical know-how and ICT infrastructure, INFLIBNET would act as an IDP for all the institutions, including universities and colleges under its umbrella. Shibboleth implementation at INFLIBNET is being amended as mentioned below.

4.4.1. User Authentication

Native installation of Shibboleth supports LDAP or Kerberos authentication mechanism, keeping in mind the present scenario of infrastructure available at colleges, it was not possible to have LDAP / Kerberos database for all the users of N-LIST, It was, therefore, decided to re-write entire JAAS (Java Authentication and Authorization Service) module, to accept authentication for users and their MD5 encoded passwords from MySQL database that was created while implementing EZproxy. Entire JAAS module was written from scratch and complied as 'jaasjini.jar' and was placed in application's library. Moreover, for directing shibboleth IDPs login handler towards newly created authentication module, following changes were made in login.config:

```
ShibUserPassAuth {
com.jini.auth.DBLogin required debug=true /* Look at jaasjini.jar for
authentication */
dbDriver="com.mysql.jdbc.Driver" /* MySQL Driver */
dbURL="jdbc:mysql://localhost:3306/inflibnet" /*JDBC URL */
dbUser="databaseusername"
dbPassword="databasepassword"
userTable="college_login"
userColumn="UserName" /* Column that contains username */
passColumn="Password"; /* Column that contains password in MD5
*/
};
```

Careful examination of above code reveals that the JAAS module has been written in a way that it will pickup authentication information (username and password) from any table and column of MySQL Database. In code snippet above, 'inflibnet' is name of database which contains NLIST users,'UserName' and "Password" are columns of 'college_login' tables which contains user's email and MD5 encrypted password respectively.

4.4.2. Attribute Release

Native mechanism of shibboleth IDP supports only one institute which is hosting IDP server. The primary task for Shibboleth implementation at INFLIBNET was to enable Shibboleth in a way that it should release user's organization, Department and its affiliation, once user has been authenticated. As attributes will be dynamic in nature for each user, it was decided to use "Relational Database Connector" instead of "Static Connector" in "attribute-reslover.xml" of Shibboleth configuration, as these attributes were already available in RDBMS tables of NLIST database.

```
<resolver:DataConnector id="mySIS" xsi:type="RelationalDatabase"
xmlns="urn:mace:shibboleth:2.0:resolver:dc">
<ApplicationManagedConnection jdbcDriver="com.mysql.jdbc.Driver"
    jdbcURL="jdbc:mysql://localhost:3306/inflibnet"       jdbcUserName="user"
jdbcPassword="password" />
    <QueryTemplate>
  <![CDATA[SELECT * FROM  college_login  WHERE User_Name  =
'$requestContext.principalName'    ]]>
    </QueryTemplate>
      <Column columnName="User_Name" attributeID="uid" />
      <Column columnName="Email_Id" attributeID="email" />
      <Column columnName="Designation" attributeID = "eduPersonAffiliation" />
      <Column columnName="cn" attributeID = "eduPersonOrgDN" />
      <Column columnName="Department" attributeID = "eduPersonOrgUnitDN" />
      <Column columnName="First_Name" attributeID =
"eduPersonPrincipalName" />
 </resolver:DataConnector>
```

In above code, a Relational Database connector was created and named as "mySIS" which points to a database named 'inflibnet' through MySQL JDBC driver, and attributes are dynamically fetched and mapped to corresponding LDAP attribute for authenticated user ('$requestContext.principalName').

For each attribute definition, Shibboleth's attribute resolver need to be informed that attribute values should be obtained from RDBMS Connector, following code snippet shows mechanism to obtain Institute Name (Organization Name i.e. eduPersonOrgDN) from RDBMS connector we have described above.

```
<resolver:AttributeDefinition id="eduPersonOrgDN" xsi:type="Simple"
xmlns="urn:mace:shibboleth:2.0:resolver:ad"
sourceAttributeID="eduPersonOrgDN">
 <resolver:Dependency ref="mySIS" />
 <resolver:AttributeEncoder xsi:type="SAML1String"
xmlns="urn:mace:shibboleth:2.0:attribute:encoder"
         name="urn:mace:dir:attribute-def:eduPersonOrgDN" />
 <resolver:AttributeEncoder xsi:type="SAML2String"
xmlns="urn:mace:shibboleth:2.0:attribute:encoder"
         name="urn:oid:1.3.6.1.4.1.5923.1.1.1.3" friendlyName="eduPersonOrgDN"
/>
  </resolver:AttributeDefinition>
```

As above example is only showing customization made for a single attribute, i.e "Institute Name) (i.e. eduPersonOrgDN) similar exercise needs to be done for all the other attributes that needs to be fetched from the database.

5. Conclusion

Access management is necessary for commercial digital collections because their access is restricted to its subscribers or licensed users. Access management may also be necessary within an organization to ascertain safety and confidentiality of documents such as commercial secrets, police records and classified government information. Access and delivery of digital content does not means just onsite access, it also means allowing access to authorized users of subscribing organizations regardless of their physical location. Authentication of users basically means ascertaining credentials of a user that allow him or her to establish the right to use subscribed resources. Login and passwords and IP filtering are two methods that are commonly deployed for authenticating users, although there are a number of other mechanisms in vogue to authenticate a user before s/he is provided access to a digital collection.

Shibboleth is open source middleware software that allows sites to make informed authorization decisions for individuals and provide access to subscription-based electronic resources. Shibboleth make use of the campus identity and access management infrastructure to authenticate individuals and then sends information about them to the resource site, enabling the resource provider to make an informed authorization decision about authenticity and authorization of a user. Using Shibboleth-enabled access simplifies management of identity and access permissions for both Identity and Service Providers. It allows for cross-domain single sign-on and removes the need for content providers to maintain usernames and passwords. Implementation of Shibboleth system at INFLIBNET has been suitably altered to suite the specific requirement of users in N-LIST programme and UGC-INFONET Digital Library Consortium.

References

INFLIBNET Centre. National Library and Information Services Infrastructure for Scholarly Content (N-LIST) Brochure. Ahmedabad, INFLIBNET, 2010. 4 p.

N-LIST Web site (http://nlist.inflibnet.ac.in) (Last visited on 31st January, 2011)

OCLC. EZproxy: authentication and access software; The leading access and authentication solution. (http://www.oclc.org/ezproxy/)

Patel, Yatrik. (2009). Shibboleth-based access management for consortia. INFLIBNET Newsletter, 16 (4), 11-13.

Simpson, Matthew, Allcock, Amy, Chen, Benjamin, Dagnone, Eugene, and Maranda, Suzanne. Ubiquitous Access to Library e-Resources. Leaflet by Queen's University. (http://meds.queensu.ca/medtech/assets/poster_proxy.pdf)

Shibboleth Web Site (http://shibboleth.internet2.edu/). Last visited on 5th February, 2011.

JISC. (http://www.jisc.ac.uk/). Last visited on 5th February, 2011.