Design and Implementation of Iraqi Virtual Library

Bahaa I. Kazem*   Mohammed Najm Abdulla   Jalal B. Raouf***
Jean-Noël Colin****   Barry Levine*******

* Center of Continuous Learning/ University of Baghdad
Email: drbahaa@gmail.com
** Department of Computer Engineering and Information Technology/ University of Technology
Email: mustafamna@yahoo.com
*** College of Administration/ University of Baghdad
Email: Jalal_bassim@hotmail.com
**** Faculty of Computer Science/ University of Namur
Email: jnc@info.fundp.ac.be
***** Department of Computer Science/ University of San Francisco State
Email: levine@sfsu.edu

Abstract

In developing countries, individual students and researchers are not able to afford the high price of the subscription to the international publishers, like JSTOR, ELSEVIER,…; therefore the governments and/or universities of those countries aim to purchase one global subscription to the international publishers to provide their educational resources at a cheaper price, or even freely, to all students and researchers of those institutions. For realizing this concept, we must build a system that sits between the publishers and the users (students or researchers) and act as a gatekeeper and a director of information: this system must register its users and must have an adequate security to ensure that only the affiliated students are able to access its services. It also needs to have security and trust mechanisms built-in for commercial partners to accept their connections. This paper describes the work done on the design and implementation of the IVL (Iraqi Virtual Library).

1. Introduction

The term "virtual library" may be interpreted very differently and in any study of the literature the reader must be aware of the myriad of interpretations that are applied to it. The literature abounds in terms that might have provided inspiration: virtual library (like the IVSL [1] and the IVL), library without walls, networked library, desktop library, logical library, information nerve center, information management center, digital library, and electronic library [2].

The Virtual library is an Internet website that acts as a gatekeeper and a director for information. It purchases the expensive licenses required to access the publishers' resources and then it enables its users to access the publishers through it freely or at marginal cost, therefore its users will not need to purchase the licenses from the publisher. It electronically manages user access and passes the appropriate credentials to the resource provider [3]. Therefore it must have a User Management functionality to manage user information, together with appropriate authentication and authorization mechanisms to

validate user credentials and access to resources. This system enables large organizations like universities and educational institutions to build a virtual library and provide its services only to the organization's members like students. The origin of the IVL code is from the e-science-library project by Dr.Jean-Noel Colin, With this project, we contributed to the development of the e-science-library by the use of Open Collaborative tools and we called it 'Iraqi Virtual Library' to distinguish it from the currently working Iraqi Virtual Science Library (IVSL) which is not an open source software and is not available for development, while the IVL is open source and is available for continuous development at http://e-science-library.dev.java.net. In 2007 the IVL project (https://e-science-library.dev.java.net/) was deployed as the AVSL project (http://avsl.aua.am/AVSL/) at American University of Armenia (AUA). Since the VSL system is based on extended java and web tools some difficulties were encountered by the system administrator due to the lack of experience in these technologies. With assistance from the java.net project administrators the system was successfully deployed. During this redeployment the project documentation was further enhanced for future ports. The AVSL project will be utilized first by AUA students and faculty. Subsequently, the project will be housed within the Armenian Academy of Sciences to support scientific research within the entire country. The IVL system primarily relies on Java technologies, EZProxy Server for the authenticaiton with commercial providers, relational database and LDAP for the user authentication. Figure (1) shows the main page of the IVL. The IVL sits between the publishers (that provide or sell access to the Digital Resources) and the users. Therefore, the Virtual library manages two types of connections: one with the publishers and the other with its users and it should provide security mechanisms for both of them. These security mechanisms make the interior structure of the Virtual Library very complicated and consist of various technologies. The publishers protect access to their information resources by different techniques therefore the Virtual Library must provide the suitable connection mechanism with each publisher.

For convenience purposes, the IVL System also offers a bulleting board functionality that allows authorized users to publish short news messages that are displayed on the main page of the website.

The rest of this paper is structured as follows: section II describes the security aspects of the IVL system; section III presents the user management functionalities. Section IV describes the news bulleting board included in the platform; Section V and VI give more details on the design and development process, and finally, some concluding remarks are provided in section VII and VIII.

2. The Proposed IVL System Operations

The designed IVL system is depicted by Figure 1, while Figure 2 shows the main page of this system, the system processes are performed as the following steps:-

A. The user accesses the IVL system.
B. The user tries to log in.
C. The IVL's Application Server checks the availability of the user account in the LDAP Realm (Directory server). It then authenticates the user, authorizes him to access the publishers, and/or authorizes him to the Admin or Editor role according to the group in which he belongs to.
D. The IVL system uses the Java Naming & Directory Interface (JNDI) to access the LDAP when the Admin user enacts the users' registration information.
E. The IVL system uses the JNDI to lockup for the MySQL server and uses the Java DataBase Connectivity (JDBC) to access it when the Editor user enacts the news database.
F. When the user requests one publisher, EZProxy forwards this request to the appropriate publisher and when it receipts the result, it returns the result back to the user's web browser. EZProxy will continuously rewrite the publisher URL's links to continuously keep itself between the user and the publisher. The publisher will see the request as coming from the IVL system's IP address which it knows in advance and permits it [4].
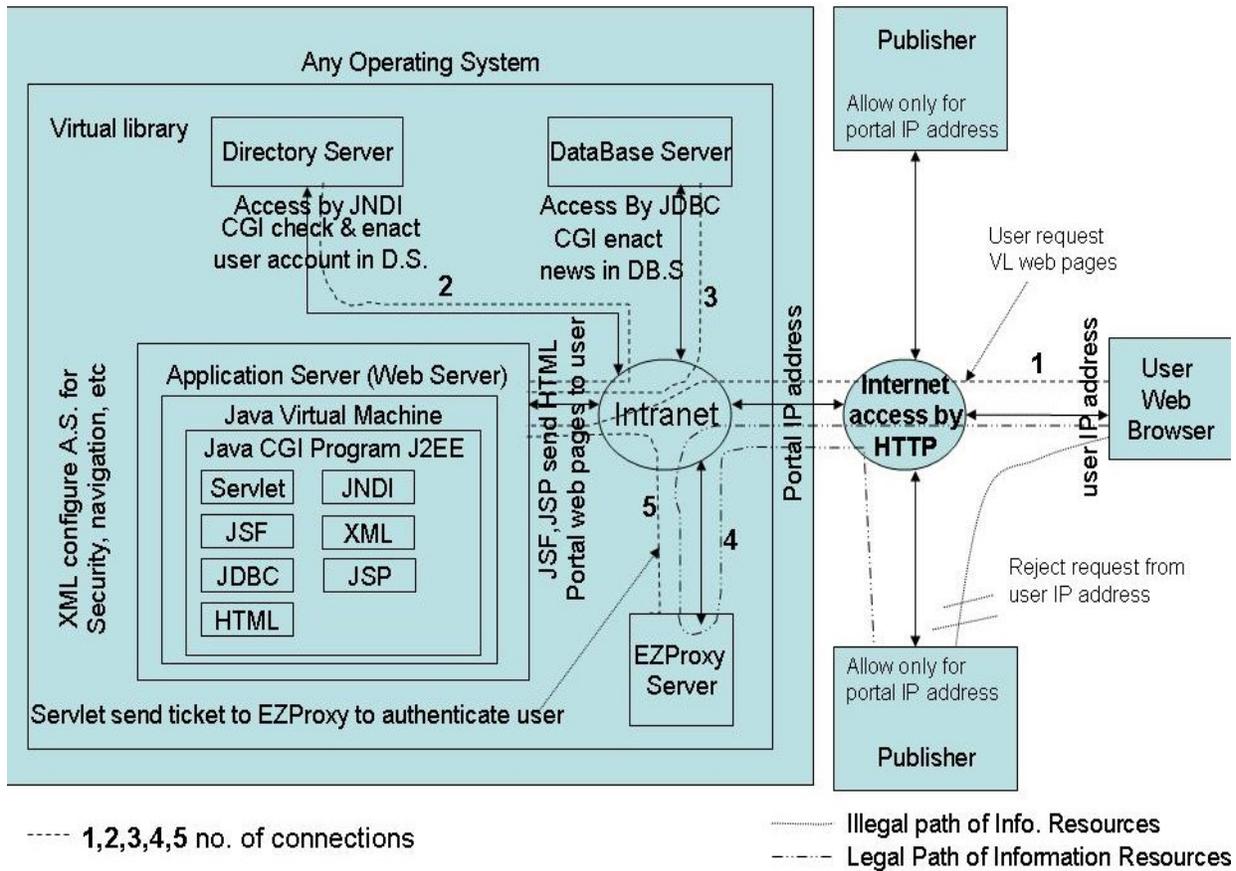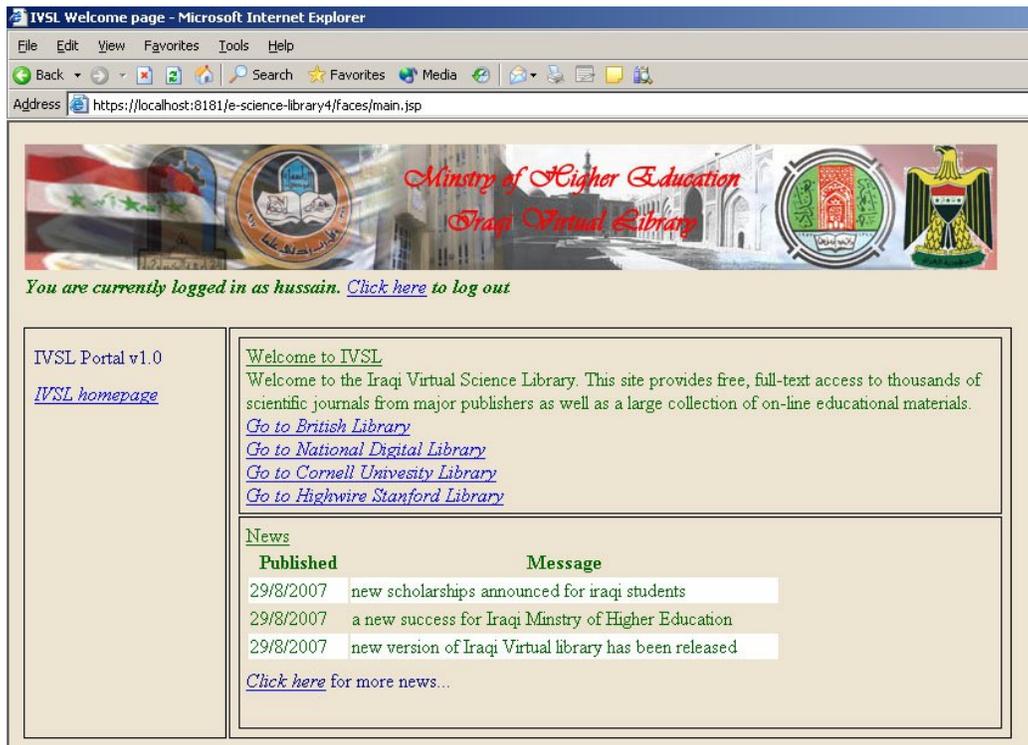
Fig.1. The IVL System Operations.

Fig.2. The Main Webpage of The IVL System.

3. The Security of the IVL System

In the IVL Distributed System, there are two types of connections, thus its security mechanisms are divided into two types accordingly:

- Securing the Virtual Library – User connection
- Securing the Virtual Library – publisher connection

How these two types of security have been implemented and what security technologies are used to secure them will be discussed in the following sections.

A. The Virtual Library – User Connection

In IVL distributed system, code representing business operations is hosted on the server. A client request acts as a trigger to execute server code that has the potential to perform critical operations that manipulate sensitive data. It is important to distinguish requests that can be trusted from those that cannot. The server must enforce security based on who is attempting to run the code, and that means being able to verify the identity of the caller. J2EE provides the Declarative and Programmatic security mechanisms which are used to enable the IVL server to authenticate the user. And the Secure Encrypted Communication is used to enable the user to authenticate the IVL server and encrypt his credentials before sending it to the IVL server [5].

1) Declarative Security

Because all applications share the following security operations and much of the work is similar regardless of the application being written, it is possible to abstract security away into a framework.

- Receives a request.
- Authenticates the caller.
- Checks the caller's authorization.
- Grant or deny access.

By using declarative security, an application can specify via parameters defined at deployment time (i.e., web.xml) the level of security that a given resource needs, and these security constraints are enforced by the Web Application container at runtime. This allows the developer to completely separate business and security concerns when designing his application. This declarative approach is reasonably flexible and easy to use. It is following the Role-Based Access Control (RBAC) model [] which is based on two main concepts: resources and roles.

- Resources are the entities that need to be protected;
- Roles define permissions to access and/or manipulate the resources. Roles are assigned to the users according to the level of permission they need to access the resources. Different users will be given different levels of access based on the role they are assigned for the application. In theory it would be possible to check every single access to the site based on the user's credentials, but for thousands of users, managing individual user access quickly becomes difficult. Instead, what happens is that users are assigned to a "role". Each resource in the site is then accessible only to certain roles [5].

In the IVL platform, we define three different roles (actors):

1. a 'regular' user is any user who reaches the IVL Portal, authenticated or not, he can do the following:

   a. Access the public IVL website
   b. Access to the publishers
   c. Login to the IVL website
   d. Logout from the IVL website

2. An 'Admin' user, who is allowed to perform the following users registration operation:

   a. Create user
   b. Modify user
   c. Delete user
   d. Suspend user
   e. Activate user
   f. Upload user data

3. an 'Editor' user, who is allowed to perform the following bulletin board management operations"

   a. Create news
   b. Modify news
   c. Delete news
   d. Archive news

Users may also have a status:

   1. an active user is known to the system and

authorized to perform any action granted by his role

2. A suspended user is known to the system, but all permissions have been suspended temporarily.

2) Programmatic Security

Programmatic Security is security mechanisms that are built in the application code itself. it is used to apply finer grain control in the application, and allow for object-level permissions. It allows the application to obtain information about the caller or the caller's role membership at runtime and apply different logic based on what is discovered [4].

3) Secure Encrypting Communication

If the user wants to send his credentials (user name, password), it has better be encrypted to be protected against eavesdropping. However, there is no point in encrypting the data if the endpoint (the IVL system) itself is not trusted. In other words, before the user sends private data, he should first ensure that he is talking to the IVL system. This is the opposite authentication that was discussed earlier, where the IVL authenticated the user. Now, the user has to authenticate the IVL server before sending the data.

To achieve this, we are implementing a solution based on X.509 certificates for server authentication. Several organizations can be used as trusted CAs to deliver and validate site certificates, like Verisign for instance. Certificates issued by Verisign are globally recognized, and Verisign's public key gets distributed with commercial client and server software such as browsers and Web servers. There is thus no need to propagate a custom CA certificate to all potential clients, which simplifies the implementation and key management.

We rely on the https protocol for the server authentication phase. When the client (the user) connects to the IVL server through his browser, the server sends its certificate, so the user can check and validate it. Further a session key can be negotiated securely. From that point on, data can be encrypted or signed using the shared session key (symmetric encryption). This is essentially what SSL does, although it's a little more complex than that [5].

B. The Virtual Library–Publisher Connection

Currently, many publishers manage access to electronic resources through the control based on clients IP addresses, an approach that makes sense when users are connecting from the university network. But relying upon IP addresses does not work for members of the community when they are at home, or when they are on leave, or traveling. Similarly, publishers need reliable authentication and authorization schemes that can distinguish legitimate users from others seeking unauthorized entry [6]. Therefore the IVL system uses the EZproxy® authentication and access software that acts as a secure gateway between the users and the publisher services. The IVL system easily provides a page containing appropriately tailored links for the publishers. After the user logs in the IVL system and clicks on a link for a publisher, the Java Servlet uses the Programmatic Security to get the user's name and then it computes the MD5 message digest of the user's name, the time of request, and the MD5 Password, then it sends this digest along with the user's name, time of request, and specified publisher's URL as a ticket to the EZProxy Server. After EZProxy Server receives the ticket, it recalculates the MD5 and authenticates the user only if that ticket is valid. After authentication, EZProxy Server forwards the user's request to the appropriate Publisher. Upon receipt of the information from the Publisher, the proxy server dynamically rewrites the information and provides it for the user, so that the publisher sees the request as coming from the EZProxy Server machine IP address. And because this machine exists within the university network as an authorized client for the publisher, access to the requested resource is granted. The most important advantage to EZProxy is that, no interaction or configuration is required from the end-user's perspective [7].

4. The User Registration System

In the IVL System, a Lightweight Directory Access Protocol (LDAP) server is used to store the users' information (user name, first name, last name, role type, e-mail, and password). A directory can ultimately be thought of as a very sophisticated address book in which entries can be associated with classes (or types) and where

entries are organized for efficient access. IVL system uses the LDAP protocol to query the server, look up users' names and validate their passwords. Typically, LDAP data are organized hierarchically, in the DIT (Directory Information Tree). The IVL directory has a single root node, with a second-level node for the university in which would be a list of the users of the IVL system. Below the single root node, there is a second-level node, one for each group of the Administrators, Editors, and users. Each user should be a member of one or more of those groups and will have its role. The members of the administrator group can modify the users list, the members of the Editors group can modify the news, and the members of the Users group can only access the publishers.

There is a one-to-one mapping between the roles defined in the declarative security described above and the group membership. In fact, being member of one of the three groups automatically assigns the corresponding role to the user. This is due to the use of a LDAP Realm in the declarative security mechanisms implemented: when the user tries to log in, the IVL system uses the Java declarative security with the LDAP Realm to check for the name, password, and the group of that user in the LDAP server and upon successfully authentication, assigns the user the Role corresponding to the group(s) the user belongs.

But when the Administrator user modifies the users' list, the IVL system uses the Java Naming and Directory Interface to access the LDAP and perform the required operations.

The reasons for having specialized directory services and not just use an existing data-storage mechanism such as a relational database management system are:-

- Directories are optimized for read-only access. You look up entries in a directory far more often than you add or update entries. Relational databases need to provide a balance between queries and update speed.
- Directories impose a structure on the data. All the entries have the same types of information in them. Relational databases have a more general-purpose structure than directories. Also, the restricted structure of a directory means that you don't have to worry about issues such as using first normal form, as you do in a database.
- Directories are often distributed. This makes administration easier and allows user load on the servers to be distributed as well. The simplified structure of a directory is easier to replicate than the structure of a relational database.
- Directories are sorted. However, only one sorting structure exists, and directories do not have to be very good at dealing with advanced queries. This allows them to be simpler, cheaper, and easier to administer [8].
- Relying on a LDAP server allows for a simpler integration in university infrastructure, since many of them already manage their users in such a directory.

5. The News Publishing Management System

The IVL also provides a bulleting board functionality that allows authorized users to publish news for the whole IVL community. The IVL system uses the relational database management system MySQL to store the news-related information and accesses it through the Java Database Connectivity (JDBC) protocol. MySQL server is designed to work in distributed environment, following the three-tier architecture (presentation, business, data)[9].

For independence between layers, the IVL system relies on the application server for the definition of the data source. It looks up the data source properties at runtime using JNDI (Java Naming and Directory Interface). It also delegates to the application server the whole connection management and pooling.

6. The Graphical User Interface

The IVL system uses the Java Server Faces (JSF) for building the server side Graphical User Interface of its WebPages. The Java Server Faces provides a standard framework for building presentation tiers for web applications [10]. JSF brings rapid user interface development to server side Java and provides much of the plumbing that JSP developers have to implement by hand [11]. JSF offers several advantages over traditional servlets and JSPs:

- It is a flexible and extensible framework, that supports the MVC paradigm, for better separation of interface and business logic development

- It offers a wide variety of graphic components, together with powerful primitives to handle events, validate user input, and define navigation rules among the pages.

The IVL system includes fifteen JSF pages [12]. The navigation between these pages is depicted by Figure 3.
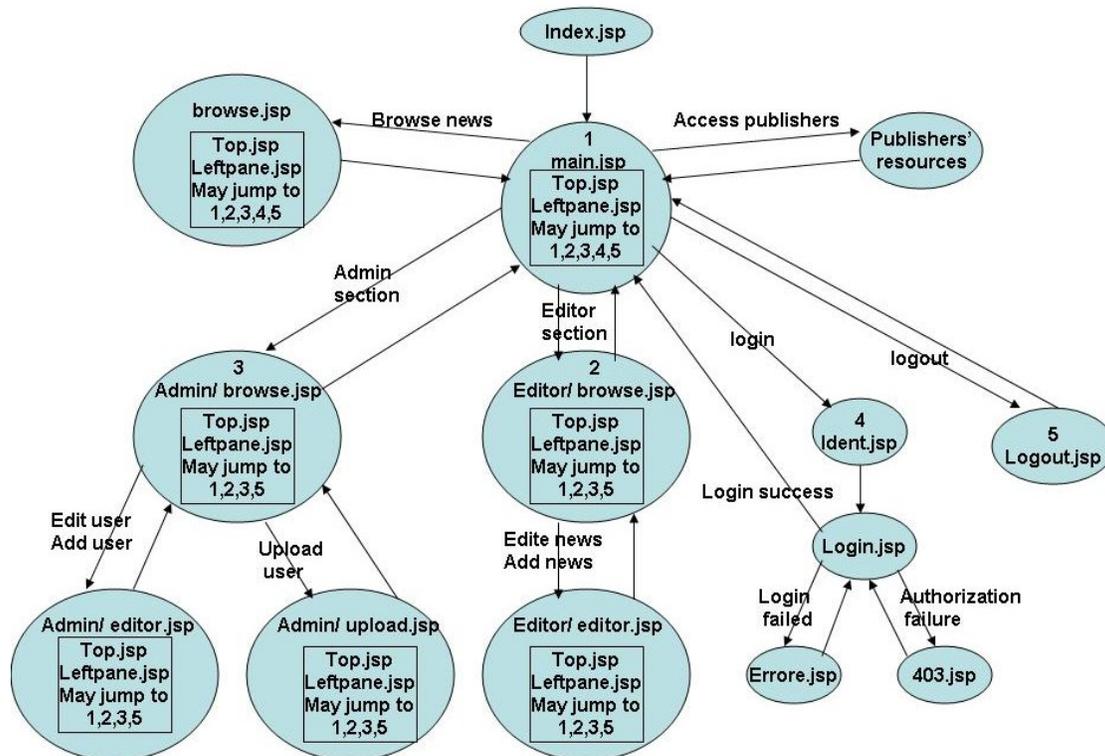

Fig.3. The Navigation between The IVL System WebPages.

7. The Development Environment

Because the codebase is publicly available, and because one of the goals of the IVL System is to contribute to the community, the Concurrent Version System (CVS) source code management have been used to contribute to the development of the IVL system at the java.net website http://e-science-library.dev.java.net. In addition to the source code repository, this web site provides free tools for supporting collaborative development by a distributed team of developers [13]. Direct communications between developers were managed via Skype and email, due to the fact that developers were split around the world. [14]. This Development mechanism provides high flexibility for the IVL and enabled the IVL to be developed,

implemented and tested by different countries including Iraq, Armenia…etc.

8. Design Comparison Between the IVL and IVSL Projects

Table 1 demonstrates the comparison between the IVL project and the IVSL project design.

9. Conclusions

A - In this paper, we have presented the design and implementation of the Iraqi Virtual Library, a solution that allows governments and institutions (mainly universities) to

legally connect to scientific publishers and provide controlled access to the resources offered by those publishers to authorized members of the staff and users.

B - The same User Registration System of the IVL system can be simultaneously adapted by the University to support services other than the access to the Publishers' resources. These services may include managing the admission to the university, recording the students' information, managing the Electronic Learning Management System.

C - The development of this system really enabled our institution to get access to a huge amount of scientific information that would otherwise be completely hidden and unreachable. This initiative definitely contributes to the spread of knowledge between countries.

D - From a technical perspective, this project has also been the occasion to learn and acquire in-depth knowledge on technologies that were not frequently used in our context. This acquired knowledge and skills will now benefit to the whole university community, being used in classes as well as the basis for further development of the IVL system.

E - Indeed, the Iraqi Virtual Library will now evolved and be integrated in a wider initiative aimed at proposing more and more integrated services through a single portal.

Table 1,
Comparison between The IVL Project and The IVSL Project Design.

	IVL	IVSL
Code	Open source code and will be in continuous development	Its code is not open for developers
Single Sign On process	Single Sign On process for accessing all publishers	Single Sign On process for accessing all publishers
Registration	The researchers at the supports institutions can apply to IVL by submitting an application by hand to the administrator, who verifies the applicant as a valid member of the institution and fills the application on-line. Approved users can use the password they wrote in their application to access the IVL	The researchers at the supported institutions can apply to IVSL by submitting an on-line application, which is sent to the registrar, who verifies the applicant as a valid member of the institution and approves or rejects the applicant. Approved users are given a user name and password and can then access the IVSL
News	Contains advance News presentation and Editing system	Contains simple News presentation and Editing system
User Registration System	It uses LDAP Directory Server for storing the users information	Unpublished
Access to Publishers	Uses EZProxy Server	Uses EZProxy Server

Abbreviations

IVL	Iraqi Virtual Library
IVSL	Iraqi Virtual Science Library
LDAP	Light Weight Directory Access Protocol
JNDI	Java Naming And Directory Interface
JDBC	Java Data Base Connectivity
URL	Uniform Resource Locator
J2EE	Java 2 Enterprise Edition
RBAC	Role Based Access Control
HTTP	Hyper Text Transfer Protocol
SSL	Secure Socket Layer
MD5	Message Digest
DIT	Directory Information Tree
JSF	Java Server Faces
JSP	Java Server Pages
MVC	Model View Controller
CVS	Concurrent Version System

10. References

[1] URL: https://www.ivsl.org.

[2] Jennifer Rowley, "The Electronic Library", Facet Publishing, London, 2002.

[3] DJ Patil, James Simon and Susan Cumberledge, "Digital Science Libraries Practical Approaches to Supporting Science in Developing Countries", UN Chronicle Online Edition, United Nations, 2006. Available: http://www.un.org/Pubs/chronicle/2006/issue3/0306p60.htm.

[4] Jalal B Raouf, "Design of Iraqi Virtual Science Library", 2007. Available: http://e-science-library.dev.java.net.

[5] Jayson Falkner, Kevin Jones"Servlets and JavaServer Pages™: The J2EE™ Technology Web Tier", Addison Wesley, United State of America, September 19, 2003.

[6] Ira H. Fuchs, "Remote Authentication and Authorization for JSTOR", Chief Scientist, JSTOR. Vice President for Computing and Information Technology, Princeton University, September 8, 2006. Available: URL: http://www.jstor.org/about/remote.html.

[7] URL: http://www.usefulutilities.com.

[8] James McGovern,Rahim Adatia, Yakov Fain "Java™ 2 Enterprise Edition 1.4 Bible", Wiley Publishing, Inc., Indianapolis, Indiana, United States of America, 2003.

[9] Paul DuBois, "MySQL Cookbook", O'Reilly, United States of America, October 2002.

[10] William Crawford, Jim Farley, "Java Enterprise in a Nutshell, 3rd Edition", O'Reilly, United States of America, November 2005.

[11] David Geary ,CAY HORSTMANN "core JavaServer™ Faces", Addison Wesley, United States of America, June 15, 2004.

[12] Jalal B Raouf, "IVSL Platform", 2007. Available: http://e-science-library.dev.java.net.

[13] URL: https://www.dev.java.net/scdocs/ddCVS.html.

[14] Dragutin Petkovic, Rainer Todtenhoefer and Gary Thompson), "Teaching Practical Software Engineering and Global Software Engineering: Case Study and Recommendations", the Computer Science Departments of San Francisco State University (SFSU) USA in conjunction with the University of Applied Sciences, Fulda University, Germany, 2005. Available: http://librarymanagementgreen.dev.java.net.

تصميم وتنفيذ المكتبة العراقية الافتراضية

بهاء ابراهيم كاظم*   محمد نجم عبد الله**   جلال باسم رؤوف***
جين نويل كولن****   باري ليفاين*****

* مركز التطوير والتعليم المستمر/ جامعة بغداد
البريد الالكتروني : drbahaa@gmail.com
** قسم هندسة الحاسبات وتكنولوجيا المعلومات/ الجامعة التكنولوجية
البريد الالكتروني : mustafamna@yahoo.com
*** كلية الادارة والاقتصاد / جامعة بغداد
البريد الالكتروني: jalal_bassim@hotmail.com
**** كلية علوم الحاسبات / جامعة نامور/ بلجيكا
البريد الالكتروني : jnc@info.fundp.ac.be
***** قسم علوم الحاسبات/ جامعة سان فرانسيسكو/ الولايات المتحدة الامريكية
البريد الالكتروني: Levine@sfsu.edu

الخلاصة

في الدول النامية لايتمكن اغلب الطلاب والباحثين من تحمل التكاليف الباهضة للتسجيل في دور النشر العالمية مثل (JSTOR, ELSEVIER) من اجل الحصول على المصادر العلمية الموجوده فيها. لهذا تسعى حكومات وجامعات هذه الدول من اجل الحصول على تسجيل جماعي لطلابها وباحثيها في دور النشر العالمية ليتمكنوا من الحصول على المصادر العلمية مجانا او باسعار مخفضة. لتحقيق هذا المسعى فان ذلك يتطلب بناء نظام يكون موقعة بين دار النشر و المستخدم (الطالب او الباحث) ويعمل هذا النظام كبوابة و موجة للمعلومات, ويجب ان يقوم هذا النظام بتسجيل المستخدمين كما يجب ان يحتوي على نظام امني موثوق لكي يضمن ان تكون خدمات النظام متاحة فقط للطلاب والباحثين التابعين لتلك الجامعة. يطلق على هذا النظام اسم المكتبة الافتراضية. هذا العمل يركز على تصميم المكتبة العراقية الافتراضية.