LIBRARY PORTAL: DESIGN, DEVELOPMENT, HOSTING AND SECURITY

Satish Kumar,
Information Scientist, ARIES
Nainital, UK

ABSTRACT

Library Portal is one among the tools of information and knowledge management which acts as a platform to connect people, processes and information of library through a single point of access. There are many types of software which provided readymade solutions to build a portal with all the required features to manage knowledge and information for example Microsoft Office SharePoint Server, Plum tree and many more. When we think of establishing a portal for a library we need to be more economic but these share point and other solutions are really costly for library so we need to find an alternative. In the way of finding alternatives the features of web 2.0 and 3.0 provided best opportunity to build our own library portal without much investment. In this paper we have made an attempt to discuss the role of portal design, development, hosting and security for library by free web tools in managing the information and knowledge.

Keywords: *Website, Web Portal, Library Website, Website Security, Content Management, Portal Elements*

INTRODUCTION

"A portal is a framework for integrating information, applications and processes across the organizational boundaries."

One of the major challenges of library professionals in universities, colleges, institutions is to provide easily accessible information to the users in a seamless manner via the Internet and Intranet because of influence which the information and communication technologies on users. The concept of an online library portal is not new, although it has evolved overtime. The mid-1990s saw the advent of public portals like AltaVista, Excite, and Yahoo! These sites provided a key set of feature – most notably news, e-mail, weather, stock quotes and search combined with advertising. Presently we are in the Information age where there is a great demand for the management and integration of Multimedia content. It is a big challenge to the knowledge mangers to provide access to the information via internet/intranet to their users on time. The conventional method of publishing content needs technical expertise on web-authoring tools and HTML (Hyper Text Markup Language), CSS (Cascading Style Sheets), web server, web browser and client side and server side scripting languages like C Programs, Perl Scripts and Shell scripts using CGI (Common Gateway Interface), PHP (Hypertext pre-processor), ASP (Active Server Page), JSP (Java Server Page) and Java Scripts. Due to these barriers CMS has evolved for the effective and flexible management of Multimedia content[1].

The major task involved in this direction is to manage content. It forced the library professionals to adopt good web content management system software. Web content management software is a computer application used to create, edit, manage, search and publish

various kinds of digital media and electronic text on the web.

CMS is a computer program that allows non-IT / non technical persons to create, edit, manage, publish and maintain content through the central interface. CMS allows its end users to rate the content, search, upload/download and to provide feedback to the site owners. CMS is available on open source as well as commercial.

Library portal makes it easy for a library to connect with its resources and information over the Internet by:

- Connecting library with students, scholars, teachers and other people
- Providing library with a customized Web browser for educational application information and functionality
- Increasing library's visibility.

Library portal is growing in its importance as the preferred way of organizing and using information. Web portals are seen as positive potential frameworks for achieving order out of chaos. As portals become a primary means for transacting information and commerce, libraries of all types are becoming involved in thinking, planning and building various frameworks and services. Library portal reduce the barrier of users having to remember multiple log-ons. The portal gives the library a tool to channel users towards preferred resources. It increases the ability of the library to ensure that costly electronic journals and databases are used, by offering a simple way to browse the available resources. It supports searching by carrying users through from bibliographic searches to full text options.

ROLE OF A LIBRARY PORTAL

A library portal was viewed as a web site rather than an application that has a high value proposition to the organization/institute. The value of the portal comes from application and content that it can deliver. Library portal is playing an important role in increasing the value of an organization. Following are the points which will highlight the role of library portal[3].

- More effective library management with workflow and institute information integration
- Easy information sharing and document management
- Improved communication and collaboration
- Easy to search and navigate
- One entry point grants access to all information
- Improve the coordination with teachers, students and researchers
- Better management of intellectual assets
- Reduce distributing and sharing information costs
- Improved security of information
- Reduced time searching for information
- Better/faster decision making
- Decreased time-to-competency
- Improved teaching and learning process
- Reduced employee training costs
- Improved collaboration across departmental libraries
- Reduction in intranet/extranet administration costs
- Reduction in network and storage costs
- Support for team work
- Reduced development of new services through standardization of presentation layer
- Increased efficiency in content creation, management & delivery..etc

DEFINITION OF A PORTAL

- A **website/Portal** is a collection of Web pages, images, videos or other digital assets that is hosted on one or several Web server(s), usually accessible via the Internet, cell phone or a LAN.
- The definition of web page is a document, typically written in HTML, which is almost always accessible via HTTP, a protocol that transfers information from the Web server to display in the user's Web browser.

WHAT MAKES A WEBSITE?

The three primary things to make a website:

- domain name
- web host
- topic

Knowing what the website is about will help determine the domain name. Most domain names are 1 to 3 words that describe what the website is about. The domain name is the URL of the website, such as www.abc.com

Domain name

- A Domain Name is a unique 'web address name' that is used to direct people to your website across the World Wide Web.
- A domain name can be obtained through a company that specializes in registering web addresses which are called Domain Registrar.
- Domain names and IP addresses are linked in a computer that carries domain information called the domain name system (**DNS**).
- Domain names are unique (for a limited amount of time from the time it is registered)
- URL stands for **Uniform Resource Locator**, and is used to specify addresses on the World Wide Web. A URL is the fundamental network identification for any resource connected to the web (e.g., hypertext pages, images, and sound files).
- URLs have the following format:
    - ***protocol://domainname/other_information***. For example, the URL for DRDO's home page is: http://www.drdo.gov.in/
- Every website address also ends in what is called a top-level domain, which is a two or three letter reference that has something to do either with the country of origin, such as ***.in*** for the India, or with the type of web page it is, such as the very popular ***.com*** for commercial web pages.[6]

TERMS RELATED TO WEBSITE DESIGNING

- HTTP
- HTML
- CSS
- Client side scripting
- Server side scripting
- Database Management Systems(DBMS)
- Web browsers
- Web Servers
- Integrated Development Environment(IDE)

HYPERTEXT TRANSFER PROTOCOL (HTTP)

- Short for Hyper Text Transfer Protocol, HTTP is a set of standards that allow users of the World Wide Web to exchange information found on web pages.
- When anyone wants to access any web page enter http:// in front of the web

address, which tells the browser to communicate over HTTP. *For example*, the full URL for DRDO is http://www.drdo.com.

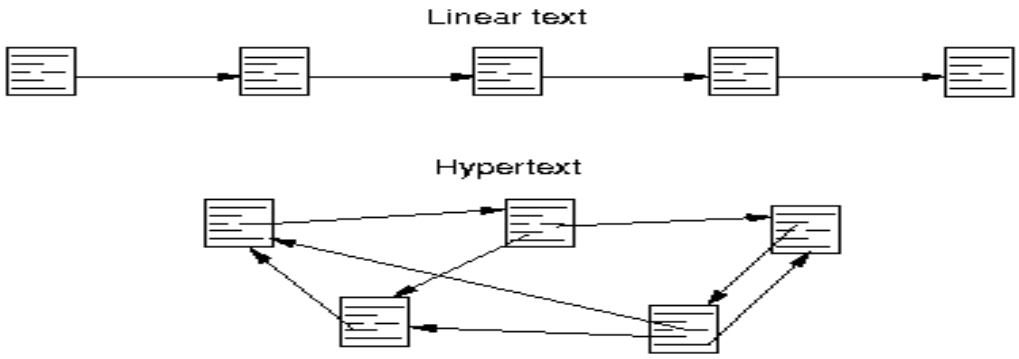
- Today's modern browsers no longer require HTTP in front of the URL since it is the default method of communication. However, it is still used in browsers because of the need to access other protocols such as FTP through the browser.

Hypertext -> hypertext is inherently nonlinear text

Transfer -> transferring on network

Protocol -> set of rules

HYPERTEXT



HYPERTEXT MARKUP LANGUAGE (HTML)

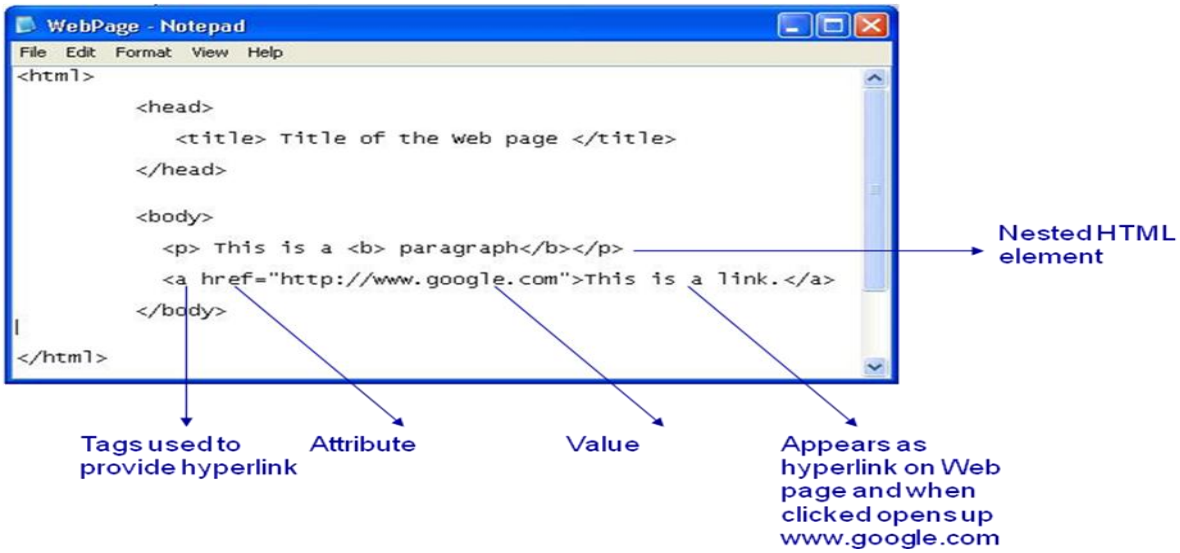
- HTML or HyperText Markup Language is the main markup language for creating web pages and other information that can be displayed in a web browser.
- HTML is written in the form of HTML elements consisting of tags enclosed in angle brackets (like <html>), within the web page content. The purpose of a web browser is to read HTML documents and compose them into visible or audible web pages. The browser does not display the HTML tags, but uses the tags to interpret the content of the page..
- HTML elements form the building blocks of all websites. HTML allows images and objects to be embedded and can be used to create interactive forms. It provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes and other items. It can embed scripts written in languages such as JavaScript which affect the behavior of HTML web pages.[5]

Markup-> markup is derived from the traditional publishing practice of "marking up" a manuscript
In Computer terminology-

HTML means a language which is used to marking up the hypertext.

SOME TAGS USED IN HTML ARE

- <html>,<xhtml>,<head>,<title>,<body>
- <a>
- <b>,<i>,<u>
- <img>
- <table><tr><td>
-
- ,<ol>,



CASCADING STYLE SHEET (CSS)

A style sheet language used for describing the look and formatting of a document written in a markup language. It can be used externally by

<link rel="stylesheet" type="text/css"
href="mystyle.css">
And internally by
<style type="text/css">
</style>

style and layout of multiple Web pages all at once. There are three ways of inserting a style sheet:

External style sheet

An external style sheet is ideal when the style is applied to many pages. With an external style sheet, you can change the look of an entire Web site by changing one file. Each page must link to the style sheet using the <link> tag. The <link> tag goes inside the head section:

<head>
<link rel="stylesheet" type="text/css"
href="mystyle.css">
</head>

Internal style sheet

An internal style sheet should be used when a single document has a unique style. You define internal styles in the head section of an HTML page, by using the <style> tag, like this:

```
<head>
    <style>
        hr          {color:sienna;}
        p           {margin-left:20px;}
        body        {background-image:url("images/back40.gif");}
    </style>
</head>
```

Inline style

To use inline styles you use the style attribute in the relevant tag. The style attribute can contain any CSS property.

*** This is a paragraph. ***

Cascading order

- Browser default
- External style sheet
- Internal style sheet (in the head section)
- Inline style (inside an HTML element)

CLIENT SIDE SCRIPTING

- Javascript
- Jquery
- VBscript

Javascript can be write in the tag

<script type="text/javascript">

</script>

SERVER SIDE SCRIPTING LANGUAGES

- JSP
- PHP
- Perl
- C#
- VB.net etc.

DATABASE S (DBMS)

- MySQL
- MSSQL
- PostgreSQL
- Oracle
- MS-Access
- IBM DB2

WEB BROWSERS

- Browser, short for web browser, is a software application used to enable computers users to locate and access web pages. The browser gets in contact with the web server and requests for information. The web server receives the information and displays it on the computer.
- Browsers translates the basic HTML (Hypertext Mark Up Language) code that allows us to see images, text videos and listen to audios on websites, along with hyperlinks that let us travel to different web pages. *Examples*- Internet Explorer, Mozilla Firefox, Google Chrome, Opera, Safari etc.

WEB SERVER & WEB HOST

- A web host is the computer that has the ability to display the website on the World Wide Web so that everyone can see it worldwide. The web host is a computer, just like you have at home, except that all it does is display websites on the World Wide Web. Web hosting is the service that provides you with the ability to display your website.[4]
- The term web server can refer to either the hardware (the computer) or the software (the computer application) that help to deliver web content that can be accessed through the Internet.

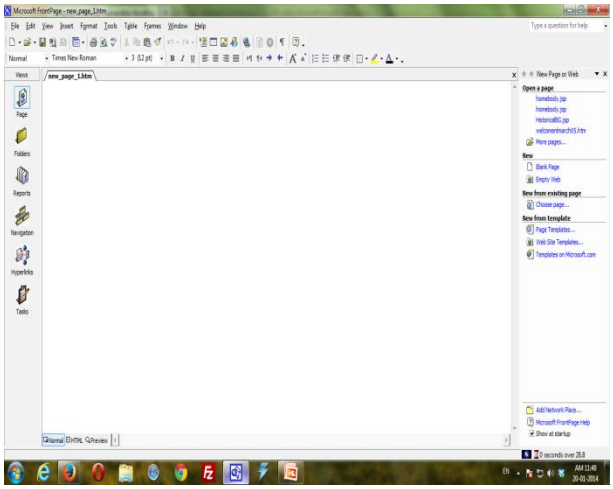
Web Servers

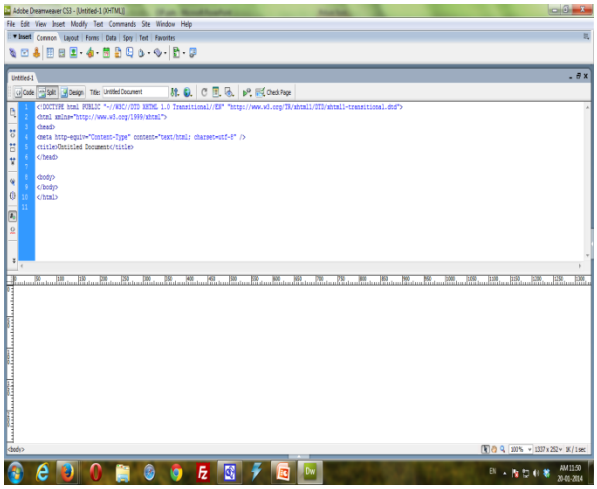
- IIS
- LAMP
- WAMP
- XAMPP
- Weblogic
- Apache Tomcat
- Glassfish
- websphere
- Oracle Application server etc.

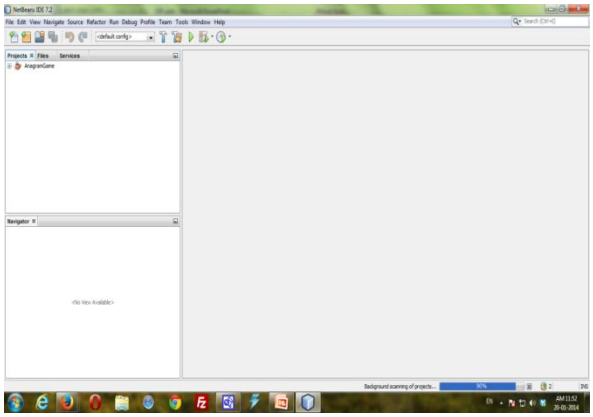
Integrated Development Environment (IDE)

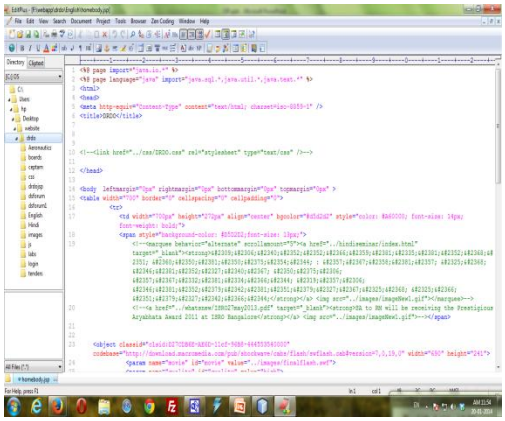
- MS- Frontpage
- Dreamweaver
- NetBeans

- Eclipse
- NusperePhpEd
- Notepad++
- EditPlus etc.









WEB SECURITY

Leading Design Guidelines for Web Applications

- Input validation
- Authentication
- Authorization
- Configuration Management
- Sensitive Data
- Session Management
- Cryptography
- Parameter Manipulation
- Exception Handling
- Auditing and Logging

Input Validation

- Attacks are performed by embedding malicious strings in query strings, form fields, cookies, and HTTP headers.
- Common examples would include command execution, cross-site scripting (XSS), SQL injection, and buffer overflow attacks.

Solution

- ✓ Do not trust input.
- ✓ Do consider centralized input validation.
- ✓ Do not rely on client-side validation.
- ✓ Be careful with canonicalization issues.
- ✓ Constrain, reject, and sanitize input.

- ✓ Validate for type, length, format, and range

Authentication

- This method involves guessing, cracking, or reusing valid credentials.
- Common examples are Identity spoofing, password cracking, elevation of privileges, and unauthorized access.

Solution

- ✓ Partition site by anonymous, identified, and authenticated area.
- ✓ Use strong passwords.
- ✓ Support password expiration periods and account disablement.
- ✓ Do not store credentials (use one-way hashes with salt).
- ✓ Encrypt communication channels to protect authentication tokens.
- ✓ Pass Forms authentication cookies only over HTTPS connections.

Authorization

- Improper or weak authorization leads to information disclosure and data tampering.
- Common examples are access to confidential or restricted data, tampering, and execution of unauthorized operations.

Solution

- ✓ Use least privileged accounts.
- ✓ Consider authorization granularity.
- ✓ Enforce separation of privileges.
- ✓ Restrict user access to system-level resources.

Configuration Management

- The consequences of a security breach to an administration interface can be severe, because the attacker frequently ends up running with administrator privileges and has direct access to the entire web application.
- Examples are unauthorized access to administration interfaces ability to update configuration data, and unauthorized access to user accounts and account profiles.

Solution

- ✓ Use least privileged process and service accounts.
- ✓ Do not store credentials in plaintext.
- ✓ Use strong authentication and authorization on administration interfaces.
- ✓ Secure the communication channel for remote administration.
- ✓ Avoid storing sensitive data in the Web space.

Sensitive Data

- The security of sensitive data while the data is stored in persistent storage and while it is passed across the network.
- Some examples are confidential information disclosure and data tampering.

Solution

- ✓ Avoid storing secrets.
- ✓ Encrypt sensitive data over the wire.
- ✓ Secure the communication channel.
- ✓ Provide strong access controls on sensitive data stores.
- ✓ Do not store sensitive data in persistent cookies.
- ✓ Do not pass sensitive data using the HTTP-GET protocol.

Session Management

- The security of sensitive data is an issue while the data is stored in persistent storage and while it is passed across the network.
- Some examples are capture of session identifiers resulting in session hijacking and identity spoofing.

Solution

- ✓ Limit the session lifetime.
- ✓ Secure the channel.
- ✓ Encrypt the contents of authentication cookies.

- ✓ Protect session state from unauthorized access.

Cryptography

- Web applications frequently use cryptography to secure data in persistent stores or as it is transmitted across networks.
- Some examples are access to confidential data or account credentials, or both.

Solution

- ✓ Use tried and tested platform features.
- ✓ Keep unencrypted data close to the algorithm.
- ✓ Use the right algorithm and key size.
- ✓ Avoid key management.
- ✓ Cycle your keys periodically.
- ✓ Store keys in a restricted location.

Parameter Manipulation

- In this the data sent between the client and Web application.
- This may be data sent using query strings, form fields, cookies, or in HTTP headers
- Examples are path traversal attacks, command execution, and bypass of access control mechanisms among others, leading to information disclosure, elevation of privileges, and denial of service.

Solution

- ✓ Encrypt sensitive cookie state.
- ✓ Do not trust fields that the client can manipulate (query strings, form fields, cookies, or HTTP headers).
- ✓ Validate all values sent from the client.

Exception Handling

- Without proper exception handling, information such as database schema details, operating system versions, stack traces, file names and path information, SQL query strings and other information of value to an attacker can be returned to the client
- Examples are denial of service and disclosure of sensitive system level details.

Solution

- ✓ Use structured exception handling.
- ✓ Do not reveal sensitive application implementation details.
- ✓ Do not log private data such as passwords.
- ✓ Consider a centralized exception management framework.

Auditing and Logging

- Log activity frequently provides early indications of a full-blown attack and the logs help address the repudiation threat where users deny their actions.
- Failure to spot the signs of intrusion, inability to prove a user's actions, and difficulties in problem diagnosis.

Solution

- ✓ Identify malicious behaviour.
- ✓ Know what good traffic looks like.
- ✓ Audit and log activity through all of the application tiers.
- ✓ Secure access to log files.
- ✓ Back up and regularly analyse log files[7].

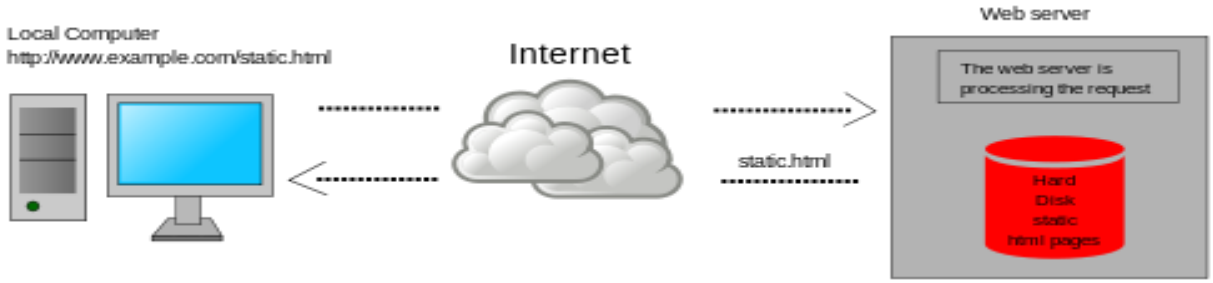
HYPERTEXT TRANSFER PROTOCOL OVER SECURE (HTTPS)

- Short for Hypertext Transfer Protocol over Secure, HTTPS is a secure method of accessing or sending information across a web page.
- All data sent over HTTPS is encrypted before it is sent, this prevents anyone from understanding that information if intercepted. Because data is encrypted over HTTPS, it is slower than HTTP, which is why HTTPS is only used when requiring login information or with

pages that contain sensitive information such as an online bank web page.

TYPES OF WEBSITES

Static website



- A static website is the simplest kind of website you can build. Static websites are written in HTML and CSS only, with no scripting. The only form of interactivity on a static website is hyperlinks. Audio or video might also be considered "static" content if it plays automatically or is generally non-interactive[2].
- A static web page (sometimes called a flat page/stationary page) is a web page that is delivered to the user exactly as stored, in contrast to dynamic web pages which are generated by a web application.
- If we want to share elements between pages (such as logos or menus), you'll have to duplicate the HTML on each page.

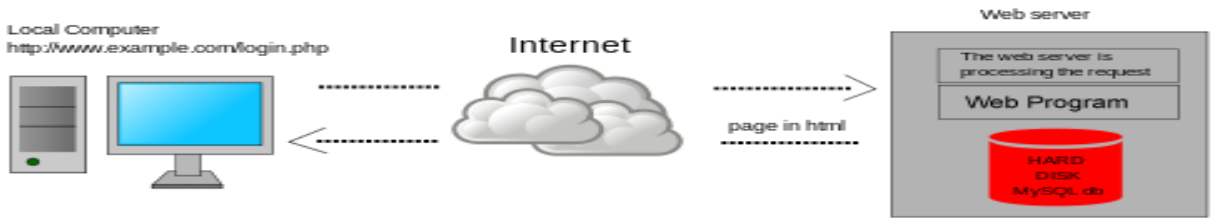
Advantages of static websites

- Quick to develop
- Cheap to develop
- Cheap to host

Disadvantages of static websites

- Requires web development expertise to update site
- Site not as useful for the user
- Content can get stagnant
- Non-Interactive

Dynamic Website



- A dynamic website is a website that not only uses HTML and CSS, but includes website scripting as well. There are two main reasons to use website scripting on website:

- *-we want an interactive web app that users can use, not just read*
- *-we want to be able to share HTML code between your pages.*

- It's not a full-blown web app like Facebook or Google, but it does have interactive elements like contact forms and search boxes.
- This website also shares the same HTML code for the header, menu and sidebar between all pages of the site.
- To create a dynamic site, learn web programming has to be used. Some common scripting languages and frameworks are JavaScript, PHP, Ruby on Rails and ASP.NET
- These websites may use databases and Content Mnagement Systems.

Advantages of dynamic websites

- Much more functional website
- Much easier to update
- New content brings people back to the site and helps in the search engines
- Can work as a system to allow staff or users to collaborate

Disadvantages of dynamic websites

- Slower / more expensive to develop
- Hosting costs a little more

CONCLUSION

For years together the role of library was to provide only information and resources in a controlled atmosphere. Limited access, rigid rules, time constrain were some of the limitations imposed by the library professionals. But now there is a shift from isolated information seeking to **interlink computing platform.**

With the growth and development of technology and information techniques and with the growing expectations from the present day users, it becomes necessary for the Librarians to transform themselves into Information Managers by developing tailor made information tools and train themselves to manage such environmental change. The end –users are at liberty to access, discuss, modify and post and retrieve information at anytime and anywhere. A portal offers one-stop shopping that takes patron from the initial need for information through its delivery without having to use several different tools. If well designed, a portal also provides effective navigation of complex, multiple collections. Thus with the modernized world, information and communication technology is changing and so the libraries. Therefore the Library professionals need to equip themselves with latest techniques and technologies in order to remain worthy.

All these portal development tools are really important to collaborate with people across the world and assist librarians to reveal library services and products to the user community. We never imagined that these technologies would impact on information services in the ways they have and we can begin to imagine the greater changes that tomorrow will bring in the world of the virtual reference library. This paper definitely helps librarians to understand and appreciate the tools and techniques for developing, hosting and securing library website.

REFERENCES

1. Tim O'Reilly. What is web 2.0 O'Reilly Network. (Available at http;//www.radar.oreillynet.com)
2. Koneru, Indira (2004). "Library portal; effective Information communication", Knowledge Organisation in Digital Environment in Libraries: Introspect and prospects, ILA Golden Jubilee Conference, 1-4 Dec 2004, pp.327-335.
3. Andress, Jason. (2011). "The basics of information security : understanding the

fundamentals of InfoSec in theory and practice, pp.16-190"

4. McInerney, Claire R. and Koenig, Michael E. D. (2011). "Knowledge Management (KM) Processes in Organizations: Theoretical Foundations and Practice", pp28-96.
5. "Overview of portals", Library Technology Reports, Nov-Dec 2002, pp.4-7.
6. Carden, Mark (2004)."Library portals and enterprise portals: why libraries need to be at the centre of enterprise portal projects", Information Services, No.24, pp.171-177.
7. Cox, Andrew (2003). "Choosing a library portal systems", The Journal of Information and Knowledge Management Systems, Vol.33 (1), pp.37-41.