

## WEB AUTHORING - A 'MUST' SKILL FOR THE INFORMATION PROFESSIONALS AT THE DAWN OF THE NEW MILLENNIUM

**M G Sreekumar**  
Librarian  
Indian Institute of  
Management Kozhikode  
REC P. O., Calicut  
Kerala - 673 601

**M Bavakutty**  
Professor & Head  
Department of Library &  
Information Science  
University of Calicut  
Calicut University P.O.  
Kerala - 673 635

*The explosive growth of the Internet and the World Wide Web (WWW) in recent years has impact on almost every field of human activity while its effect on librarianship is more intensive owing to the information component attached to the profession. For the first time in history, the WWW with its suite of standards and respective technologies made significant strides in the field of information transfer and knowledge management by liberating the basic static nature of the ordinary text, making it dynamic and powerful. The Web has opened up a new school of thought, the hypertext world, with its unquestioned ability in distributed computing on a global scale. The purpose of this paper is to give information professionals a sense of what is unique about the design of Web Pages on the WWW and an indication in general terms of what types of information and materials can be disseminated through this medium. It describes the very need for a Site, its unlimited scope, opportunities and implications and suggests several guidelines for Web site development that will enhance the potential for improved visibility of the library/information centre, its information resources, systems and the array of services, to be seen and utilised by outside world. Also discusses the method of creating an HTML document supplemented with advanced applications of DHTML, XML, VRML, scripting languages, style sheets etc., how the page could be published in the Net and further publicising the page.*

### RATIONALE

Information technology enabled value-added information resources and services are becoming the de-facto standard the world over. Already the

World Wide Web boasts over 70 million users and is growing at a phenomenal rate. Information professionals cannot ignore this population no matter what the transactions or resources are. The Web page of a library facilitates a user to have online access to its resources and services and thereby providing the library services round the clock. Being a part of this community indicates that besides being interested in our patrons we have professional commitment to organisation, retrieval and dissemination of information. Experience of Web Sites possessing libraries and information centres the world over shows that they have received wider popularity, increased business and revenue and also enhanced the visibility.

### INTRODUCTION

The explosive growth of the Internet and the World Wide Web in recent years has its impact on the information profession too. It has registered a sea change in the information seeking approach as well as the mode of dissemination of information. The most appreciated fact about the Internet is that it allows millions of people all over the world to communicate and share. As librarians and information professionals, it is our prime responsibility to acquire, organise, preserve, retrieve and disseminate pertinent information to our clientele. It is in this context that we have to look at Internet with a friendly attitude. This global forum, an emerging medium of communication and a proven and concrete technology in sharing and exchanging information, has a lot to offer to the information professionals. Studies reveal that the world over the world of information is going digital. Digital information has the unique quality of condensing the time element in the creation,

storage, retrieval and dissemination process. This is characterised by the fact that the moment information is transferred into electronic form, it becomes mobile and dynamic rather than static. It can then be processed, transmitted, stored, retrieved and disseminated at enormous speed and convenience. With the advent of the Internet and communication technologies this information can be transmitted across space. Further, there are three clear benefits to going digital. Firstly it helps librarians to preserve rare and fragile documents without denying access to those who wish to access them. A second benefit is convenience as once information is converted to digital form, users can retrieve it in seconds rather than in minutes. The third advantage of electronic copies is that they occupy millimeters of space on a magnetic disk rather than meters on a shelf. Another latest important development is the host of digital information services which attract more users than ever before. Internet thus plays a commendable role here.

Internet and the WWW provide a variety of opportunities to the information professionals in outreaching their clientele. As active professionals, we should be able to publicise and market our resources and services in the Net. We should learn the Web publishing technology, the art of creating and launching Web Sites. The technology is becoming more and more friendly as time passes by. What is important is our initiative, enthusiasm, vision and perseverance. A genuine question arises here as to what is its use when the library does not even have a Net connection. Well, this is a tricky situation. Still, we would suggest we should go on. If not today, tomorrow this shall have to be a reality. Also if the library has a PC, or the librarian has access to a PC, this could be tried and be readied to load and launch the site straight away. The site could well be utilised within the library/institution as a means of presenting the activities. Once the authorities find that there is something worth showing to the outside world, they shall give green signal at least for a dial-up TCP/IP connection. A number of our activities could be projected on the site. It could act as a bulletin of the various events the library organises, the host of services it offers to its patrons, an interactive catalogue - say the books collection, a periodical list of additions, an array of pointers to other useful sites, etc. Its scope is only limited to our own

imagination. In this paper, we intend to summarise the basics of the Internet and its popular services, the WWW and as to how a website is created and launched.

## **INTERNET**

The secret to understanding the Internet is that it is populated by two types of computer programs, namely, servers and clients. Servers are programs that provide resources and clients are programs that we use to access these resources. E-Mail, LISTSERV/Mail lists, USENET/Newsgroups, FTP, Telnet, Gopher, Archie, WWW (World Wide Web) etc., are among the prominent services of the Internet. Each type of service in the Internet has its own client. For example, to access the WWW, we need to use a Web client such as "Netscape" or "Internet Explorer". As already mentioned, Internet is a collection of networks covering the world. These networks contain many different types of computers and there should be a tool to hold the whole thing together and that is "TCP/IP", which stands for Transmission Control Protocol/Internet Protocol. A protocol is a set of rules describing, in technical terms, how something should be done. TCP/IP is the common name for a collection of more than 100 protocols used to connect computers and networks.

## **THE WORLD WIDE WEB (WWW)**

The WWW is almost synonymous with the Net. It is a large system of servers which offers all kinds of information on the Net. The information can be in the form of regular text, as well as pictures, sounds, video clips, and other types of data. Invented in early 1990s by physicists under the leadership of Tim Berners-Lee, at the European Particle Physics Laboratory (CERN), Switzerland, the Web has rapidly become the Graphical User Interface (GUI) of the Net. In fact, the Web is just one Internet application, a way of using the vast interconnected network and view information from around the world. The main use of the Web is for information retrieval, whereby multi-media documents are copied over the Net for local viewing. It uses a protocol called HTTP (Hypertext Transfer Protocol). The Web is in fact a multi-media hypertext - any sort of digital data can be distributed inside a Web document, and each document contains links to other documents, as

shown by highlighted or underlined text. Simply click on the word or the highlighted area (where the palm appears) and we can travel to the document in question. Web documents are written using a markup language called Hypertext Markup Language (HTML). HTML is independent of platforms. So it does not matter if you are using a Macintosh, a PC or a Sun workstation to access them. The Massachusetts Institute of Technology's (MIT) World Wide Web Consortium (W3C) is the international standardising agency for HTML and HTML 4.0 is the latest standard. HTML has tags for providing references to other Web pages, which can be on the same server or any other server on the Net irrespective of the geographical location. This facilitates hypertext links across the documents on the Net. Web pages can contain references of images (GIF and JPEG formats), audio files (AU and WAV formats) and video files (MPEG format) which thus adds the multi-media dimension to the information provided on the Web. The clients or browsers access the Web pages on the servers, render and format them according to the HTML tags to display on the client's system. When the user selects a hypertext link (indicated by an underscore) on a Web page, client can follow the link and fetch the referred document irrespective of the location of the document on the Net. Every Web document has a unique address, called the Uniform Resource Locator (URL). It is a simple way of describing almost any information resource, using a standard format for locating information on the Net as:

```
<protocol>://<host.domain>:<port>/<protocol-specific>
```

It is an Internet address which is unique to that resource. It consists of a protocol, a host name, a port (optional), a directory and a file name. That is, URL consists of the computer on which the document is stored and a file name. For example, the URL of the Indian Institute of Management at Kozhikode, is: "http://www.iimk.org". It is to be noted that the file name is missing in the example, as it refers to the home page.

The key to the Web is a browser program, which is used to retrieve and display Web documents. The browser (HTML viewer) is an Internet compatible program that runs on our local

computer, whether it is a Macintosh, IBM PC, UNIX Workstation, and does three things for the Web documents, it

- a) uses the Internet to retrieve documents from other computers, called servers;
- b) displays these documents on our screen, using the formatting specified in the documents; and
- c) makes the displayed documents active, so that pointing and clicking on a cross-referenced item in a document will take us to the reference. There are two popular Web browsers, viz., the Navigator of Netscape and the Internet Explorer, from Microsoft. However, Netscape has become the de facto standard for browsers.

### WEB SITE DESIGNING

Creating a Web Page for a library or information centre is exciting and at the same time challenging. A well formulated strategy is pre-requisite for a successful site. Quite often, it is felt that setting up a site is easier than to properly maintain it. To build a coherent Web (or intranet) site, one has to have an overall organisational plan for the site. Web engineering is an emerging discipline and the Web design process is something challenging and at the same time rewarding. It is a real blend of art and science. If one know some bit of HTML, lots more on the Web page can be done. Though it sounds very complex and difficult to learn, nothing could be farther from the truth. In fact, one can probably pick up basic HTML in a few hours. It may be difficult to digest the codes in the beginning, but with little bit of interest, enthusiasm and determination, it can be a reality.

A cool library page should have the following features:

- (i) The page should highlight the name of the library and its parent organisation in an appealing manner
- (ii) A brief listing of the major information resources with hyperlinks to each of them

- (iii) A brief listing of the major information services with hyperlinks wherever necessary
- (iv) Highlights on some important events in the library, special collections, etc.
- (v) Hyperlinks to general information about library, library guide, library publications, recent additions list, library bulletins, etc.
- (iv) Shortcuts to online resources, preferably in a pop-up menu/combo box format, which may include:
  - Electronic journals, popular magazines, newspapers etc.
  - Institutions of interest and other important libraries and information centres
  - Internet guides, conference, symposia, workshops etc.

It would be ideal to have an interactive catalogue of the books and other major collections such as journals, videos, microfilms/microfische etc., which the users shall be able to consult straightaway. This requires the involvement of advanced applications and hence this should be attempted with due preparations. Though pictures shall enhance the appeal of pages, precaution should be taken to make them as small as possible, and also the pixels kept moderately low, say 300-400 dpi.

Given below are the web sites of two of the world's most famous libraries, viz., the British Library, UK (Fig.1., "www.bl.uk/index.html") and the U.S. Library of Congress (Fig.2."www.loc.gov"). One should notice the various features of these pages, the highlights, hyperlinks etc.

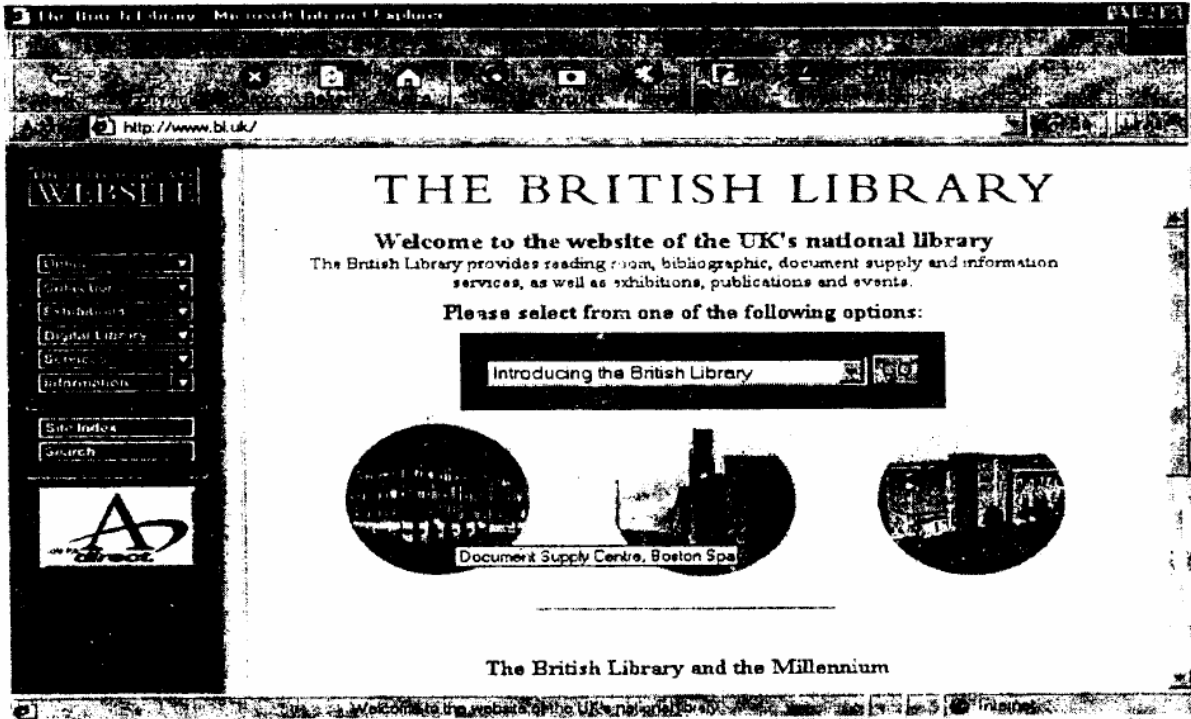


Fig: 1 : The British Library Web Site

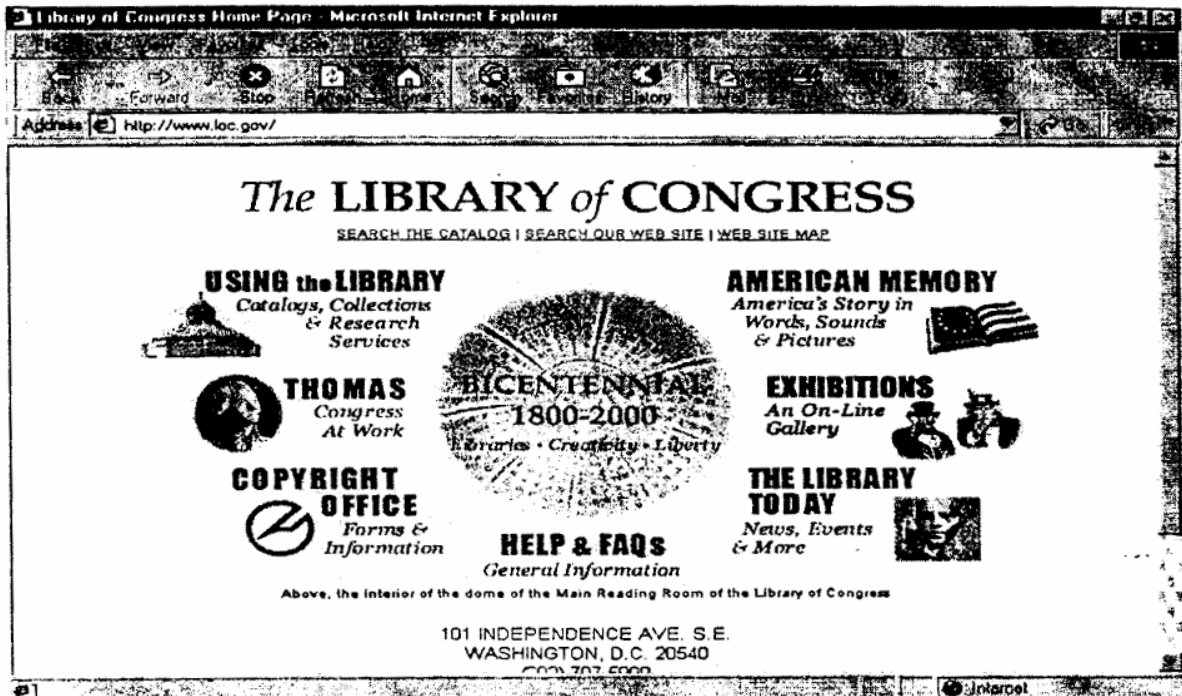


Fig. 2 : The Home Page of the Library of Congress

### HYPER TEXT MARKUP LANGUAGE (HTML)

HTML, as mentioned earlier, stands for HyperText Markup Language, is not a programming language. It is the lingua franca for publishing hypertext on the World Wide Web. It is a non-proprietary format based upon SGML (Standard General Markup Language) and can be created and processed by a wide range of tools, from simple plain text editors - you type it in from scratch - to sophisticated WYSIWYG authoring tools. It is a way of telling your Web browser how to display the contents of a page. All you need is to create a basic HTML document using a text editor like Microsoft Windows Notepad, DOS edit, Windows Write, Macintosh Teach Text/Simple Text, UNIX vi or emacs, or VAX/VMS edit. Word processors like

WordPerfect or Microsoft Word may also be used, as long as you save the file as a text file or in ASCII mode. Simply put, an HTML document is a plain text file with different formatting codes. These codes are enclosed in angle brackets ("**<**", "**>**") and are referred to as tags. When a browser reads the page and sees a tag in front of some text, it knows that it has to display it in a particular way. But the tag itself is not displayed. For instance, if we give "**<B>** Make this text bold **</B>**", the browser shall format the contents of the tags to appear in boldface. The tag that says **<B>** is a message to the browser - it means 'whatever text follows this tag must appear in bold face'. The tag "**</B>**" at the end is the closing tag that tells the browser that it no longer has to follow the formatting that the previous tag has set. The closing tag always

has a forward slash before it. Learning HTML is merely knowing what each tag does.

An important thing to note about HTML markup codes is that they are not case sensitive, <body> is the same as <bODY> or <boDY> or <BodY>. Most HTML authors use uppercase consistently for HTML markup code because it makes the markup stand out visually from the actual text of the HTML document, easing the chore of proof reading.

### **Types of tags**

There are three main types of tages in HTML grouped as paired tags, stand alone tags and tags with attributes. The paired tags require both an opening and closing tag. An example would be <B> Make this text bold </B>. Stand alone tags are always on their own as they do not have closing tags. An example to this could be <HR>, which shall insert a horizontal line, where the tag appears. It does not require a closing tag as it is not specifying any text attributes. Tags with attributes form yet another group. One can also add some properties or attributes to a paired tag or a stand alone tag to make them do more. For example, if we give <HR SIZE ="5" WIDTH=50%>, the browser shall make the line 5 pixeis high and shall give it a width of 50% of the width of the page.

### **HTML Body Structure**

The anatomy of an HTML page basically requires a few tags that are mandatory for every HTML page that you create. The document structure is as follows:

```
<HTML>
<HEAD>
<TITLE> Library Web page </TITLE>
</HEAD>
<BODY>
Welcome!
[Put all your stuff here]
</BODY>
</HTML>
```

Save it as a file called index.htm on the hard disk. If you are using Notepad, set the "Save as type" field in the Save dialog box to "All files", then save it as index.htm. Go to the folder in which you have

saved your file and then double click index.htm. This should launch your browser and load up the file. You will see only the "Welcome!" as the output. We can now build on from this basic structure to a full-fledged Web Page.

Every HTML document comprises two parts which are header and body. The header contains the author's name and date, the name of the HTML editor program (used for creating the page) and the text for the title bar. The body contains the content of the Web page, with HTML tags. <HTML> and its closing tag instruct the Web browser to begin and stop interpreting HTML tags. <HEAD> and (BODY> tags indicate the start of the header and body sections. <TITLE> indicates text for the browser title bar. Within the body, you can type headings and text or insert tables, graphics and hyperlinks. There are plenty of HTML primers and tutorials on the Internet that list the HTML commands. Check out the tutorials at the Yahoo! and NCSA (National Centre for Supercomputing Applications). The Internet site (www.internet.com) is another useful location to pick up HTML commands. There are in addition plenty of computer books and journals that deal with HTML comprehensively and are readily available in the market. The images we find in the sites we visit could be downloaded using the right mouse button, by giving suitable file names. Also the cut and paste method could be used in getting some of the useful applications found in some interesting sites.

### **Authoring tools (HTML Editors)**

There exist a number of tools to help us build pages. "Netscape Composer", a component of the Netscape Communicator suite, is a WYSIWYG (What You See Is What You Get) HTML editor. The nicest thing about Composer is its smooth integration with the publication wizard and browser (Netscape Navigator). Click the preview button and the Netscape browser appears with your page in it. Click another button and the publishing wizard pops up (for uploading your pages to a remote Web site). Videos and background sounds can be added on the page simply by inserting links to the appropriate files. A new HTML feature that has been added to the Composer is the use of Style sheets. It is a real joy to work with Composer and one may explore all other possibilities all by

## THE INFORMATION PROFESSIONAL AT THE DAWN OF THE NEW MILLENNIUM

oneself. The "Front Page Express" is another WYSIWYG full blown HTML editor, which comes with Microsoft Internet Explorer 4.x. It is the light version of Front Page 98. The Web editor enables one to see page layout and formatting exactly as it would appear when viewed through a browser - through the entire process of Web page creation. Though it will be beyond the scope of this

paper to enlist all the HTML tags and its attributes, it is presumed that a basic listing shall definitely provide a great deal of understanding of the various HTML tags used. Developers should consult other books and HTML tutorial guides to learn more about Web page designing. Given below are some of the useful tags widely used in creating pages.

Tag	Function
Structure Tags	
<!--...-->	Creates a comment (not visible on the page, but visible for anyone viewing the source code)
<HTML>...</HTML>	Indicates the start and end of the entire HTML document
<HEAD>...</HEAD>	Denotes the header information in the HTML document
<BODY>...</BODY>	Encloses the body (text and tags) of the entire HTML document
<META name="Keywords" content="your own keywords">	Notifies Web search engines about the contents of the site so that people searching for the page will find it
Text Function Tags	
<TITLE>...</TITLE>	Denotes the title of the document, which will appear in the list of Web search results
<H1>...</H1> to <H6>...</H6>	Encloses headings 1 to 6, with heading 1 being the largest and 6 the smallest
<P>...</P>	Indicates that the enclosed text is a basic paragraph
<OL>...</OL>	Encloses ordered (numbered) list
<UL>...</UL>	Encloses an unordered (bulleted) list
<LI>...</LI>	Denotes a list item for either type of list
Formatting Tags	
<B>...</B>	Encloses bold text, you can also use <STRONG>...</STRONG>
<I>...</I>	Marks italic text, you can also use <EM>...</EM> (for emphasis)
<U>...</U>	Denotes underlined text
<TT>...</TT>	Indicates monospaced text, good for user instructions, explaining computer messages
<CITE>...</CITE>	Signifies book, film, or other title citations with italics
<CODE>...</CODE>	Identifies source code using a monospaced font
<HR>...</HR>	Inserts an embossed horizontal rule
 ...</BR>	Preserves a line break
<BLOCKQUOTE>...</BLOCKQUOTE>	Encloses long quotes or citations in a different font or indented margins
<ADDRESS>...</ADDRESS>	Denotes an address block with information about the Web page, its author and the last update



---

Link and Image Tags	
<code>&lt;A HREF="URL"&gt;Click Here&lt;/A&gt;</code>	Anchor tag defining a Web address or file link for the "Click Here" hypertext
<code>&lt;IMG SRC="filename.gif"&gt;</code> <code>&lt;ALT="Text describing image")</code>	Inserts the image whose file name appears in the tag Displays text instead of an image for browsers not displaying images
<code>(A HREF="URL"&gt;&lt;IMG SRC=</code>	Links the URL or the file name "filename.gif"> referenced to a Web page image

---

Advanced HTML Tags	
<code>&lt;TABLE&gt;...&lt;/TABLE&gt;</code>	Defines the beginning of a Table and all its contents
<code>&lt;TR&gt;...&lt;/TR&gt;</code>	Denotes table rows
<code>&lt;TD&gt;...&lt;/TD&gt;</code>	Marks table cells within rows
<code>&lt;TH&gt;...&lt;/TH&gt;</code>	Encloses cell headings (usually columns) for the table
<code>&lt;FRAME&gt;</code>	Defines a Frame (please note End Tag is Illegal)
<code>&lt;FRAMESET&gt;...&lt;/FRAMESET&gt;</code>	Defines the layout of Frames within a Window
<code>&lt;FORM&gt;...&lt;/FORM&gt;</code>	Creates a form that holds controls for user input

---

### Advanced HTML Features

Some of the advanced HTML features are:

**Forms** is probably the most important feature of HTML for collecting user input, in addition to including hyperlinks and images in a page. One can get input from users and feed it to computer programs for almost any purpose, such as sign-up registration, taking orders, or updating a database. Pull-Down Menus and Scrollboxes with multiple choices also could be provided in the Page, as desired.

**Frames** is a new feature of HTML, that attracted immediate attention of everyone, and Web pages using it popped up all over the Internet. Frames are much like the split-screen video tricks used by television networks to retain viewers between shows. Frames can contain ordinary HTML markup, can be scrollable and can even hold clickable images or image maps. While you interact with one frame, the contents of the other frames still remain on the screen.

**CGI.** Once you know how to collect information with HTML forms, we can spend a bit of time learning about processing it. As mentioned above, the basic concept is taking the information entered on the form and passing it to the Web server and CGI is the standard way of doing so. For CGI-based forms, the input will be directly passed on to the Web server with the help of the CGI program commands. The CGI program usually will be located in the cgi-bin directory, cgi-bin being the short form for Common Gateway Interface binary program, a computer program. One can write CGI programs in almost any programming language, including UNIX shell scripts, DOS batch files, Visual BASIC, Apple Script, the C language, or others. PERL, short form for Larry Wall's Practical Extraction and Report Language, is the most widely used language for CGI programming. It is freely available for most systems on which Web servers run, including PCs and Macs. Whatever language is used, CGI script must accept as input the information the user has entered into the form, then process it in some way. The Fig. 3 shows the interaction between the Client, the remote Server, and the CGI application.



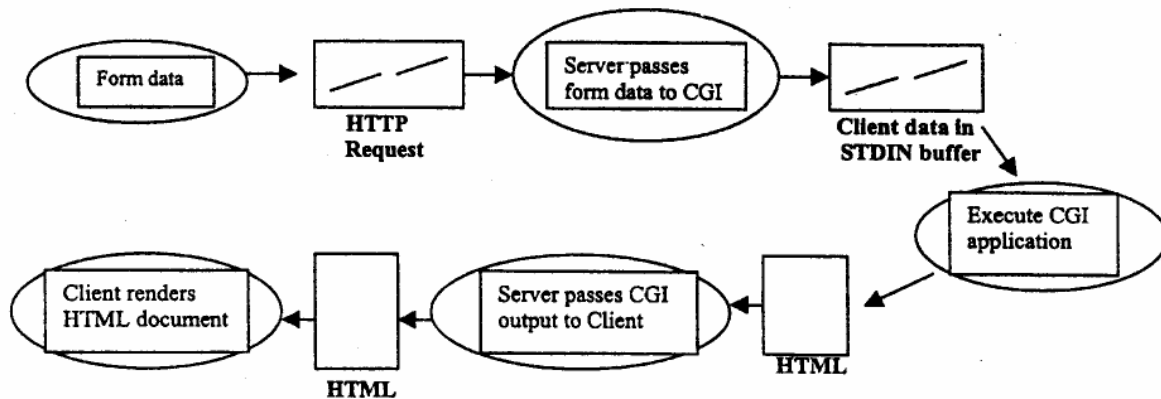


Fig. 3

**Stylesheets**, a recent development in HTML document presentation, allows the user to specify the formatting instructions as an entity separate from the text markup, and to specify these instructions using a language designed for formatting details. Stylesheets are independent of HTML. This approach turns the act of designing documents into a two-part process. In the first part, the author marks up the document itself, for example using HTML, to denote the main structural components, and to distinguish these components (headers, body, footers etc.) one from the other. In the second part, the developer designs a collection of formatting instructions that will specify the desired formatting for the different structural components. This collection of instructions is called the document stylesheet, as it contains the formatting or styling information. Cascading Stylesheets (CSS) is the currently deployed Web stylesheet language.

New HTML features such as Microsoft's **Marquees** tag allow one to place scrolling text on

a Web page, similar to a banner, or the Netscape's Blink feature. When the user opens the page, the text starts to scroll from left to right or right to left across the screen.

To make the Web page more attractive, interactive and dynamic, one can use **Java, VRML, ActiveX Control, Shockwave** etc. Java is a programming language developed by Sun Micro Systems, USA. Web page developers create mini-programs called applets, that can be placed right on Web pages. Java applets can run on any operating system - Windows, Mac, or UNIX, and hence the developer does not have to create a separate application for each operating system. This allows developers to place active content on Web pages without having to worry about which operating system the visitor is using. VRML, Virtual Reality Modeling Language, is similar to Java. With the VRML programming language, developers can create three-dimensional virtual worlds, which can be entered and explored. Shockwave is a player created by Macromedia, whose Director program is the

leading tool for putting together multimedia presentations. With Director, developers combine still pictures, animations, and sounds, and include point-and-click interaction. Macromedia Flash is a program that puts together small and fast Shockwave multimedia. Many Web sites contain presentations created with Flash. Shockwave lets one to view Director and Flash presentations over the Web. A Web site designed for use with Shockwave can present video and sound without having to request each animation and sound file. It also allows one to interact by clicking buttons. All these special applications need the respective enabling browsers, such as Java enabled, ActiveX, VRML browser etc. However, one does

not need a copy of Director or Flash, unless a personal presentation is required to be designed. In addition, there are Helper Applications and Plug-Ins available for help, which allow to incorporate multimedia (audio and video) into the HTML documents. Helper applications take the data which Web browsers cannot interpret and deal with it, displaying unsupported images or playing sound or video files, calling up a new window to display the data. Plug-Ins perform much of the same role as helper applications, but are more tightly integrated into newer Web browsers (Netscape and Explorer). Rather than popping up a new window for the multimedia data, it is displayed right in the browser window.

Some of the popular file formats, which are generally used in Web environment, are listed below:

File extension	Format
.au	Sound
.avi	movie format
.dtd	text (SGML file description)
.gif	pictures/graphics
.html/.htm	HTML files
.jpg	pictures/graphics
.mid	music
.mpg	movie/sound
.pdf	Acrobat Document
.png	pictures/graphics
.rtf	text
.sig	signature
.wav	sound

### **HTML AND BEYOND**

Dynamic HTML (DHTML) is newer form of HTML and not a new version of HTML. It uses a model similar to that of Object Oriented Programming (OOP) to build on HTML tags; yet it permits dynamic styles, content, and positioning as well as data binding to a browser. This means that the site becomes interactive for the user. A visitor who reaches a DHTML Web page can have richer, faster browsing through client-side processing. To put it briefly, it is the marriage of several previously separate Web technologies, expanded to

accommodate a new way of thinking about Web page design that is style sheets, a scripting language, and a document object model to tie it all together. Style Sheets, Sheets, or more specifically Cascading Style Sheets, as specified earlier, will allow one to describe precisely how the pages should look and exactly where and how things should appear.

A scripting language, such as a JavaScript or VBScript, will give the Web pages the power to follow and react to every mouse movement anywhere in the document, and in any way it is

## THE INFORMATION PROFESSIONAL AT THE DAWN OF THE NEW MILLENNIUM

wanted. A Document Object Model (DOM) will make every paragraph, every image, even every single character in the document a distinct, individual entity with the ability to be addressed, manipulated, and even completely changed whenever and however wanted.

XHTML 1.0 is the first major change to HTML since HTML 4.0 was released in 1997. It brings the rigor of XML to Web pages and is the keystone in World Wide Web Consortium's (W3C) work to create standards that provide richer Web pages on an ever increasing range of browser platforms including cell phones, televisions, cars, wallet sized wireless communications, kiosks and desktops. XHTML is modular making it easy to combine with markup tags for things like vector graphics, multimedia, mathematics, electronic commerce and more. Content providers will find it easier to produce content for a wide range of platforms, with better assurances as to how the content is rendered. The modular design reflects the realization that a one-size-fits-all approach will no longer work in a world where browsers vary enormously in their capabilities. A browser in a cellphone can't offer the same experience as a top of the range multimedia desktop machine. The cellphone doesn't even have the memory to load the page designed for the desktop browser. XHTML 1.0 is the first step and the HTML working group is busy on the next. XHTML 1.0 reformulates HTML as an XML application. This makes it easier to process and easier to maintain. XHTML 1.0 borrows the tags from W3C's earlier work on HTML 4, and can be interpreted by existing browsers, by following a few simple guidelines. This allows one to start using XHTML. One can roll over the old HTML documents into XHTML using W3C's Open Source HTML Tidy utility. This tool also cleans up markup errors, removes clutter and prettifies the markup making it easier to maintain. HTML 4.01 is a revision of the HTML 4.0 Recommendation first released on 18th December 1997.

Here is the set of modules recommended in the latest draft:

- Applet - applet, param
- Block phrasal - address, blockquote, pre, h1-h6
- Block presentational - center, hr

- Block structural - div, p
- In line phrasal - abbr, acronym, cite, code, dfn, em, kbd, q, samp, strong, var
- In-line presentational - b, basefont, big, font, i, s, small, strike, sub, sup, tt, u
- Inline structural - bdo, br, del, ins, span
- Linking - a, base, link
- Lists - dir, dl, dt, dd, ol, ul, li, menu
- Simple forms - form, input, select, option, textarea
- Extended forms- button, fieldset, label, legend, optgroup, option, select, textarea
- Simple tables - table, td, th, tr
- Extended tables - caption, col, colgroup, tbody, tfoot, thead
- Images - img
- Image maps - area, map
- Objects - object, param
- Frames - frameset, frame, iframe, noframes
- Events - Onclick, ondblclick, onmousedown, onmouseup, onmouseover, onmousemove, onmouseout, onkeypress, onkeydown, onkeyup
- Metadata - meta, title
- Scripts - noscript, script
- Styles - style element and attribute
- Structure - html, head, body

XML stands for Extensible Markup Language and is an extremely simple dialect of SGML (Standard Generalized Markup Language). XML has been developed to fill the gap between the power and complexity of SGML at one end of the spectrum, and the inadequacy and simplicity of HTML at the other end of the spectrum. The specifications of XML still are not yet finished. XML is an attempt to find a common ground between SGML and HTML. XML is not intended to be a replacement for HTML. XML is meant to be a supplement and an alternative to HTML, when it is needed XML and HTML can happily coexist, each being used for what it is most suited to; HTML for quick applications for Web pages and XML for applications that need more intelligent documents and more processing ability.

### GUIDELINES FOR AUTHORIZING

Here are some rough guidelines for HTML authors. If you use these, you are more likely to end up with pages that are easy to maintain, look acceptable to users regardless of the browser they

are using, and can be accessed by the many Web users with disabilities. Meanwhile W3C have produced some more formal guidelines for authors. Some of the detailed accessibility guidelines are:

1. A question of style sheets. For most people the look of a document - the color, the font and the margins are as important as the textual content of the document itself. But make no mistake! HTML is not designed to be used to control these aspects of document layout. What you should do is to use HTML to mark up headings, paragraphs, lists, hypertext links, and other structural parts of the document, and then add a style sheet to specify layout separately, just as one might do in a conventional Desk Top Publishing Package. That way, not only is there a better chance of all browsers displaying the document properly, but also, if one wants to change such things as the font or color, it is really simple to do so. See the touch of style.
2. FONT tag considered harmful! Many filters from word-processing packages and also some HTML authoring tools, generate HTML code which is completely contrary to the design goals of the language. What they do is to look at a document almost purely from the point of view of layout, and then mimic that layout in HTML by doing tricks with FONT, BR and &nbsp; (non-breaking spaces). HTML documents are supposed to be structured around items such as paragraphs, heading and lists. Yet some of these documents barely have a paragraph tag in sight. The problem comes when the content of pages needs to be updated or given a new layout or re-cast in XML (which is now to be the new mark-up language). With proper use of HTML, such operations are not difficult, but with a muddle of non-structural tags it is quite a different matter and maintenance tasks become impractical. To correct pages suffering from injudicious use of FONT, try the HTML Tidy program, which will do its best to put things right and generate better and more manageable HTML.
3. Make the pages readable by those with disabilities. The Web is a tremendously useful tool for the visually impaired or blind user, but

bear in mind that these users rely on speech synthesizers or Braille readers to render the text. Sloopy mark-up, or mark-up which does not have the layout defined in a separate style sheet is hard for such softwares to deal with. Wherever possible, use a style sheet for the presentational aspects of the pages, using HTML purely for structural mark-up.

Also, remember to include descriptions with each image and try to avoid server-side image maps. For tables, you should include a summary of the table's structure and remember to associate table data with relevant headers. This will give non-visual browsers a chance to help orientate people as they move from one cell to the next. For forms, remember to include labels for form fields.

It is suggested here that the accessibility guidelines for a more detailed account of how to make the Web pages really accessible be consulted.

#### **IIMK LIBRARY WEB PAGE**

The Web page for the library and information centre (LIC) of the Indian Institute of Management Kozhikode was prepared as a sub system providing a hyperlink from the IIMK Home Page (<http://www.iimk.org>). The important desired features of an attractive, that is, a cool library site have already been discussed earlier. A thorough systems study and analysis was made taking into consideration the relevant features and significant information that has to go into the LIC page. All the important functional areas, activities, components, resources and services were identified. More importantly, the user's access points (user approaches) were also identified. Fig.4 is the LIC page when viewed through a Web browser such as Netscape or Internet Explorer.

As already discussed earlier, Web engineering is an increasingly important emerging discipline and developers should take extreme care and precautions while embarking on site development. Some of the major system design considerations include the time frame, general look and feel of the site, guidelines for content, structure, budget, site maintenance and technology issues. Cool sites are always under better control and it also gives a professional look to the outside world.

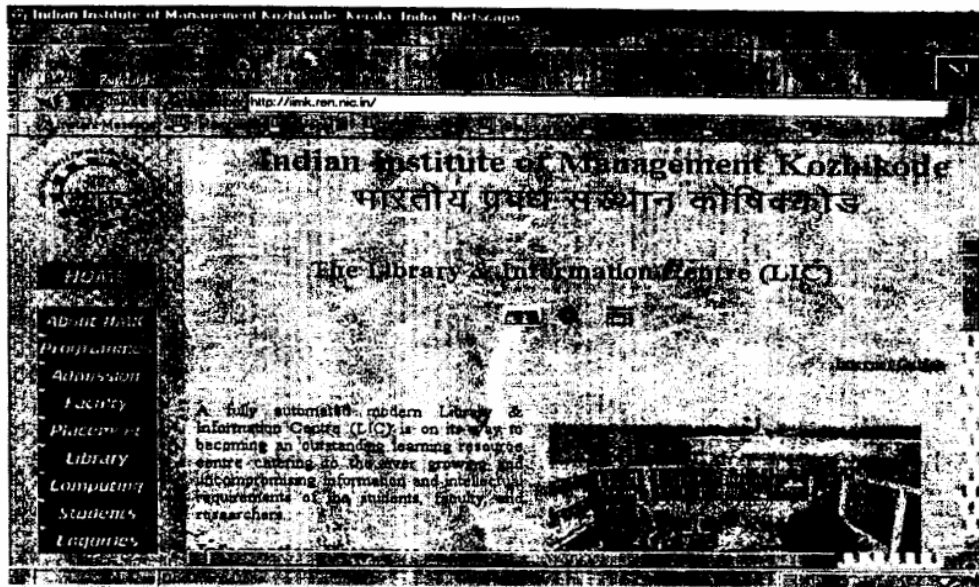


Fig. 4 : The IIMK LIC Web page

### Web Server Selection

In the above case since the LIC's page is part of the IIMK home page, its launching is automatically taken care of at the time of the launching the home page itself. Otherwise, after designing all the HTML files, which include the main "index.html" also, the next step is to load them on to a computer from where the same can be accessed by a common user. The following three options are considered for launching a site.

- a) The ideal situation would be to have a computer fully configured to work as Web Server located inside the Institute's premises for better management and maintainability and to have your own Domain Name registered using which the users can access your Web site. However, taking into consideration the cost and the time required for Domain Name registration and setting up our own Web Server, the option is not a viable one.
- b) There are many Web Servers available where one can launch the Web site on payment basis. These servers provide a fixed amount of hard disk space for storing required HTML and other related files with file transfer permission to the user for uploading the files to the server using FTP, as and when needed, on predetermined payment terms. This also requires that your own Domain Name must already be registered so that the site can be accessed using the registered Domain Name. However, this also will require some time and efforts on your part to make it operational.
- c) There are many Web Servers available where one can launch the Web site free of cost. The URLs of some such Web servers include [www.geocities.com](http://www.geocities.com); [www.rediff.com](http://www.rediff.com); [www.angelfire.com](http://www.angelfire.com); [www.tripod.com](http://www.tripod.com); [www.firefly.com](http://www.firefly.com); [www.zoom.com](http://www.zoom.com), and so on. In order to put your library page on any server, you must first register yourselves on that server.

### **UPLOADING USING CuteFTP**

CuteFTP helps to transfer files across the Internet. Whether publishing a Web page, downloading the latest digital music and software or transferring high-volume files between computers. CuteFTP provides the tools one needs and its user-friendly interface will help transfer files faster.

### **DESIGN ASSISTANCE**

**CoolText.com** is a totally free online service which provides real-time generation of graphics customized exactly the way you want them. Please choose what kind of image you would like to create from one of the options listed at [www.cooltext.com](http://www.cooltext.com) on the left. Then simply fill out the form and click on the render button to have your custom graphic generated on the fly. The service will always be available for use free of charge.

**V Splash** is an online web site creation tool that helps develop a web interface faster, without any technical knowledge. This is a priced service and the total cost of developing and hosting the site will be around Rs.5000/-. The uniqueness of [vspalsh.com](http://vspalsh.com), as per their claim, is that it offers the average computer user to design a web site the way he wants it.

### **DOMAIN NAME REGISTRATION**

You have to apply to InterNic/Network Solutions Inc. for the registration of your own domain name so that your site could be accessed with a URL of your choice. The DNS entry will be allotted to you by the agency after due formalities.

It would be advisable to have a payment based Web server say, 'hostserver.com' (<http://www.hostserver.com>), which shall guarantee your site's presence round the clock and also provide you with a 50 MB of hard disk space and file transfer permission using FTP. Once you transfer all the HTML files along with the main file "index.html", the site will be available on the URL of your choice, and it shall be pointing to the IP address of hostserver.com computer.

### **PUBLICISING THE SITE**

When it is time to announce the site to the world, submit the URL to Internet search engines such as Altavista, Lycos, Excite, Infoseek, Askeeves, Khoj, etc and also to Indexing services such as Yahoo etc. One can also announce the site to others through banner advertisements put on Web pages. There are special newsgroup services that announce new Web sites. A mailing list of our preference (say LIS-FORUM of NCSI) may also be used for announcing the site. There are Internet-based companies ([www.webcom.com](http://www.webcom.com), [www.doog.com](http://www.doog.com), [www.register-it.com](http://www.register-it.com), etc.) that offer publicity services specifically for Web sites. Also, <META> tag notifies Web search engines about the contents of the site so that people searching for the page will find it.

Add Me! is another popular website promotion and submission engine. It is a free service that lets one submit to the website to over 30 popular search engines and directories on the web.

### **Acknowledgement**

The first author expresses his profound thanks to Prof. A.H. Kalro, Director, Indian Institute of Management, Kozhikode for the kind encouragement and support.

### **REFERENCES**

1. The Economist. Ed., Going digital: How new technology is changing our lives. *The Economist*. 1996.
2. GRAHAM (Ian S). HTML stylesheet sourcebook. 1997. John Wiley & Sons Inc.
3. KRAYNAK (Joe) and HABRAKEN (Joe). Internet 6 inl. PHI. 1997.
4. TITTEL (Ed.) et al, Ed. CGI Bible. 1997. Comdex Computer Publishing.
5. PRASHER (R G), Ed., Library and Information Science. V.2: Information Science, IT and its application. 1997. Concept Pub. Co.

## THE INFORMATION PROFESSIONAL AT THE DAWN OF THE NEW MILLENNIUM

6. CRUMLISH (Christian). The Internet. 1998. BPB.
7. MUDRY (Jon Robert). The DHTML companion, Prentice HallInc., 1998.
8. ZIMMERMAN (Scott), BROWN (Christopher L T). Web site construction kit for Windows 95. Techmedia. 1998.
9. POWERS (Shelley) et al., Dynamic Web publishing. 2nd ed. Techmedia. 1998.
10. SREEKUMAR (M G). 'A client/server system for interactive access to a national database over the Internet'. Proceedings of the Conference on Advances in Computing (CADCOM98), Calicut, (6-8 April 1998), pp. 289-294.
11. SREEKUMAR (M G) and PATHAK (Ashok), WWW and HTML - The papyrus and script of the cyber age. National Convention on Business and Management Information, Ahmedabad, 4-6 March 1999.
12. See <http://developer.netscape.com/library/documentation/htmlguid/index.htm>
13. See <http://www.internet.com>
14. See <http://www.microsoft.com/workshop/author/newhtml/>
15. See <http://www.w3.org>
16. See <http://www.w3.org/pub/WWW/Style/>
17. See <http://www.w3.org.usm.edu/~ila/>
18. See <http://www.servtech.com/~mvail/women.html>
19. See <http://www.cuteftp.com/products/cuteftp/index.html>
20. See <http://www.networksolutions.com/>
21. See <http://www.internic.net/>
22. See <http://www.addme.com>