

United Nations Educational Scientific and Cultural Organization  
and  
International Atomic Energy Agency

THE ABDUS SALAM INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

**ON HTML AND XML BASED WEB DESIGN  
AND IMPLEMENTATION TECHNIQUES**

T. Bezboruah<sup>1</sup>

*Department of Electronics Science, Gauhati University, Guwahati-781014, Assam, India  
and*

*The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy*

and

M. Kalita

*Department of Electronics Science, Gauhati University, Guwahati-781014, Assam, India.*

**Abstract**

Web implementation is truly a multidisciplinary field with influences from programming, choosing of scripting languages, graphic design, user interface design, and database design. The challenge of a Web designer/implementer is his ability to create an attractive and informative Web. To work with the universal framework and link diagrams from the design process as well as the Web specifications and domain information, it is essential to create Hypertext Markup Language (HTML) or other software and multimedia to accomplish the Web's objective. In this article we will discuss Web design standards and the techniques involved in Web implementation based on HTML and Extensible Markup Language (XML). We will also discuss the advantages and disadvantages of HTML over its successor XML in designing and implementing a Web. We have developed two Web pages, one utilizing the features of HTML and the other based on the features of XML to carry out the present investigation.

MIRAMARE – TRIESTE

May 2006

---

<sup>1</sup> Junior Associate of ICTP. [zbt\\_gu@yahoo.co.in](mailto:zbt_gu@yahoo.co.in)

## 1. INTRODUCTION

In principle the Web is still based on the system of hypertext that Tim Berners-Lee and others developed at the European Laboratory for Particle Physics in Geneva, Switzerland in the late 1980s [1, 2, 3]. But the Web today is more diverse technologically and more diffused within society and culture. The range of technologies that a Web developer can choose from is now more varied than ever. Besides an array of techniques and tools to work with HTML/XML, developers can also use many technologies to add new kinds of multimedia and interactive content to on-line services. New kinds of software to view Web content are being developed, and the competition for the provider of Internet software has risen to the highest priority in the personal computer industry. The initial outlook of the Web in 1989 was a text-based browser deployed on an internal network. But today the Web is a global medium that encompasses many software and communication systems from corner to corner of many networks.

The Java language [4], specifically designed for network communications, rose prominently in 1995, as a new way of communication over the Web. Following Java's success, a new company, Lucent Technologies, spun off from American Telephone and Telegraph (AT&T) and presented its Inferno system [5] to the world to address the need for a network operating system as well as a language for network distributed content. Both Java and Inferno represent major innovative ways of thinking about online communications and technically supporting it. The transmission of sound, video, and other multimedia effects were also possible before 1995-1996, but during this period systems emerged providing high-quality solutions to the problems of distributing multimedia on global networks. The RealAudio [6] emerged as an outstanding solution for providing audio on demand over the Internet. Multimedia developers also gained new possibilities for networked communication with Macromedia's Shockwave product [7]. Shockwave provides a set of plugins for its existing Macromedia authorware and enables developers to create multimedia content.

The method of distributing RealAudio or Shockwave content to users who freely obtain available viewer software, remains a usual model for distributing new media on the Web. Virtual Reality Modeling Language (VRML) [8] is another major technology that emerged during this period as a component of Web developer's choices for manifestation. VRML opens up the possibility for three-dimensional worlds, a cyberspace more in line with the visions of science fiction writers of the past. Combined with Java, the 3-D worlds of VRML can have the behavior, shapes and structures that can respond to a user's presence and input. Microsoft's [9] strategic focus on the Internet, as a key part of personal computing, certainly shifted the whole playing ground of Web technology development. Netscape Communications' Web software continued to dominate, but Microsoft's Internet Explorer and other browsers [10] emerged as stern alternatives for browsing the Web.

## **2. THE STATE OF ART IN WEB IMPLEMENTATION**

The explosion of interest in the World Wide Web (WWW) and its native languages, HTML and XML has fostered enormous competitions among Internet software vendors. The big players, Netscape [11] and Microsoft [9] remain in strong competition to create the most popular software for the Internet. For the user, this competition has given rise to a very wide range of choices for software. But for the developer, the competing commercial forces, combined with the popular interest in the WWW, have resulted in a more fragmented communications environment. The nature of this fragmentation has many folds as mentioned below:

(i) There are now more variations in HTML and XML elements. Different brands of Web browsers recognize different HTML and XML elements, many of which have not been recognized as standards by the World Wide Web Consortium (W3C) [12]. The slow process for the ratification of HTML and XML standards has given browser manufacturers strapping power to set those standards.

(ii) There are new kinds of associated software that work with Web browsers. Some Web delivered information depends on plugins and add-on software to be used with the browser for displaying information. In many cases, these plugins are available only for certain platforms and can be used only with certain brands of browsers. Languages such as Java assure to break the platform independence, reducing the model for transmitting Web information.

(iii) There are new kinds of media involved in Web information and communication, new kinds of ways for sound, video, animation, and other affects to be implemented on the Web. Many of these media types involve special formats and plugins for interpretation. Others, such as Java based media types, assure to provide a more platform independent and seamless way for users to recognize media.

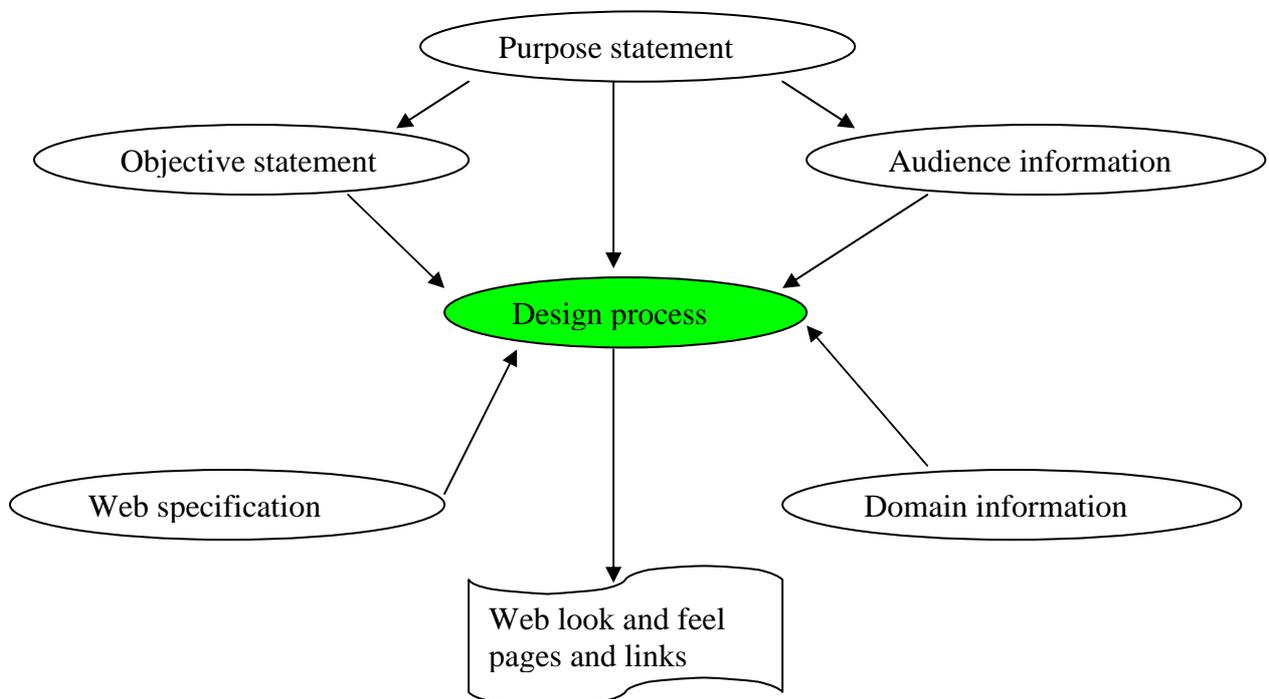
(iv) Automated techniques for HTML and XML implementation have emerged. No longer do many professional Web implementers depend on hand-crafted HTML and XML pages. Instead, the trend is toward automatic generation of HTML and XML pages based on databases and/or templates. Despite all the advances and changes in Web-implementation technologies, the requirement for trained and talented implementers who take a process-oriented view toward implementation has changed. The concept of implementation as one part of an overall methodology to meet the Web user's requirements hasn't been changed. Web implementation performed within the framework using a repeatable, process-oriented approach is one way to cope with constantly changing technology.

## **3. OUR PRESENT WEB DESIGN TECHNIQUES**

A Web's design is essentially its look and feel. We have taken into account all the Web elements, e.g. audience information, purpose and objective statements, domain information and Web specifications and have combined them to produce an arrangement for implementing the Web. We then utilize this design and the Web specifications to create an

attractive and informative Web. We make our own choices about how to best achieve the effects called for by the Web planning process, the objective statements, and audience information. We have drawn on a catalog of techniques for packaging, linking, and cueing information using design methodologies. Throughout this process, we are very sensitive to users' experiences of the Web's information space, consistency, and cues. There are practical issues which are involved in design, such as considerations for inline images and graphics, how much to put on a single page, and which text or images should be made a link in contrast to which should not.

The design process, however, is just one process in the interlocking Web development processes. A successful Web requires that all processes and all elements work together. Figure 1 illustrates how the Web-design process takes information from all elements of Web weaving and combines them to produce a look-and-feel design which can be used through the implementation process to create an effective Web. By separating the design from the implementation process, information about the Web's structure and operation can be cast in a hypertext as well as language independent form. While the design process is influenced by knowledge of what is possible in the target design language, its product can be implemented in any language that can capture the features used in the design. In this way, the design process can be adopted with successors or alternatives to the widely used HTML and XML. In our present design we adopted both HTML and XML.



**Figure 1: An overall Web design process**

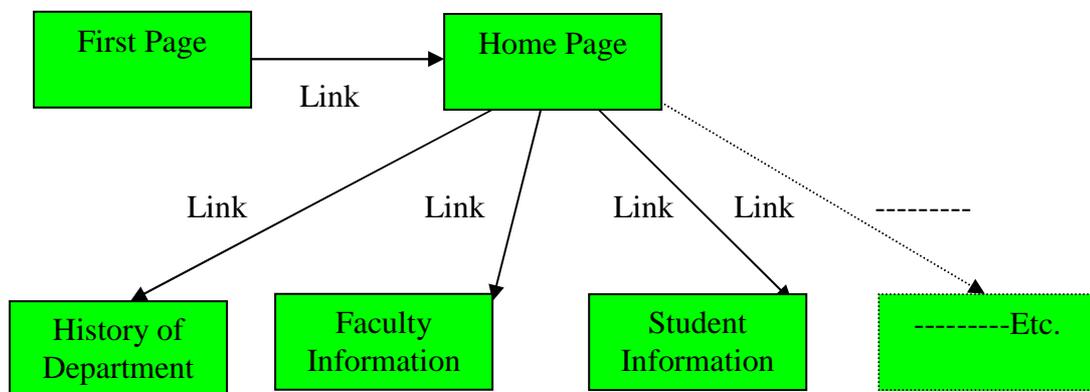
### 3.1 The HTML

HTML is a language for describing how pages of text, graphics and other information are organized, formatted and linked together. The Web mainly uses HTML documents for the sharing of data across distributed, heterogeneous hardware and software environments. Business can use a browser interface to access data from sources around the world. HTML provides an excellent means of sharing content with a variety of Web based clients but it has several disadvantages [13] such as:

- a. It gives much emphasis on presentation rather than content. HTML focuses mainly on computer-to-human interface, but it has limited value in applications like transferring of information between databases (computer-to-computer).
- b. HTML has a fixed set of markup tags. It lacks support for creating new, application specific tags.
- c. HTML does not help applications to validate data, as it is entered or imported.

To overcome HTML's limitations W3C created XML, which is more extensible than HTML. This new language lets information publishers to invent their own tags for their benefit.

The flow diagram for information processing on the Web in our present design technique, in its simplest form, is shown in figure 2. Two sample HTML files created for our present design are also shown in figures 3.1 and 3.2



**Figure 2: Flow diagram for receiving information on the Web in its simplest form**

**File 1:**

```

<HTML>
<HEAD>
  <TITLE>Right frame</TITLE>
  <FRAMESET BORDER=0 ROWS="30%, *">
    <FRAME SRC="lftpage.html" scrolling="no" name="topframe">
    <FRAMESET BORDER=0 ROWS="22%,*,15%">
      <FRAME SRC="menu.html" scrolling="no" name="menuframe">
      <FRAME SRC="leftmn.html" scrolling="yes" name="bottomframe">
      <FRAME SRC="design.xml" scrolling="no" name="designframe">
    </FRAMESET>
  </FRAMESET>
</HEAD>
</HTML>

```

**Figure 3.1: Sample program for demonstration of links and frameset using HTML**

**File 2:**

```

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<HTML>
<HEAD>
  <TITLE>Left bottom menu</TITLE>
</HEAD>
<BODY bgcolor="royalblue" vlink="green" alink="navyblue">
  <table row="1" border="3" cellspacing="1" cellpadding="3"><th bgcolor="plum">
  <font size="2" color="black">
  <a target="rightbframe" href="post.xml">Post Graduate Studies</a><br/>
  <a target="right frame" href="qualif.xml">Essential Qualification</a><br/>
  <a target="rightbframe" href="student.html">Present Student</a><br/>
  <a target="rightbframe" href="event.xml">Events</a><br/>
  <a target="rightbframe" href="staff.html">Teaching Staff</a><br/>
  <a target="rightbframe" href="research.xml">Research
  Group</a><br/><a target="rightbframe" href="rstudent.xml">Research Student</a><br/>
  <a target="rightbframe" href="job.html">Job Opportunity</a><br/>
  <a target="rightbframe" href="alumni.html">Alumni</a><br/><a target="rightbframe"
  href="journal.html">Books and Journals</a><br/><a target="rightbframe"
  href="history.html">History of the Department</a><br/>
  <a target="rightbframe" href="admin.html">Technical & Office Staff</a><br/>
  <a target="rightbframe" href="course.html"> Courses Offerred</a>

```

```
</table>  
</BODY>  
</HTML>
```

**Figure 3.2: A sample program for creating the menu**

### 3.2 The XML

The W3C began work on XML in 1996 and adopted version 1.0 of the technology in 1998 [14, 15]. There are two main reasons behind this development:

a. Business-to-business (B2B) e-commerce required companies on different platforms to communicate seamlessly with one another.

b. Also many companies wanted to make legacy data in back-end systems, such as database, available to new Web-based applications.

The XML is simply a set of rules that is used to control the creation of a markup language that can identify document content so that information can be formatted consistently throughout the document. This set of rules identifies how one can define tags that separate a document into parts and subparts. XML markup is a meta-language which means that the markup can be used to describe other languages. This allows one to create his own XML tags to provide information about the information for which the XML based document structure is created. XML is a very easy to use, easy to read by a computer, easy to debug, and easy to use for creating markup languages suitable for any industry that uses structured data such as spreadsheets, databases, financial information and technical drawings. It is best suited for documents that have large amounts of similarly organized information such as catalogs, address books, mathematical functions and even account records. XML allows a document to be read by any type of software.

The XML tags provide metadata, which is the information about data that a document contains. However recipients must have access to the tags meaning to access a document correctly. XML has used Document Type Definition (DTD), a mechanism for constraining the use of markup tags, to accomplish this [16]. A DTD which a sender transmits with an XML based document, tells a browser or other DTD readers what the tags are called, what they mean, how they are used, where they may occur and how they fit together [17]. As the DTDs cannot adequately model complex data, the W3C subsequently developed an XML schema specification which adds data types, relationships and constraints [18]. Applications can use XML parser to validate imported data for conformation to a DTD or schema. Two sample XML files for our present design are given in figures 4.1 and 4.2

**File 1:**

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/css" href="rstudentcs.css"?>
<!--
  Document   : rstudent.xml
  Created on : June 4, 2005, 3:57 PM
  Author    : Resarch Students
  Description:
    Purpose of the document follows.
-->

<!DOCTYPE root [
<!ELEMENT root (fst+)>
<!ELEMENT fst (#PCDATA)>
]>
<root>
<fst>Ms. M. Kalita</fst>
<fst>Mrs. K. RoySinha</fst>
<fst>Mr. G. Deka</fst>
<fst>Mr. M. Bora</fst>
<fst>Mr. S. Haque</fst>
</root>
```

**Figure 4.1: A sample XML program for displaying information**

**File 2:**

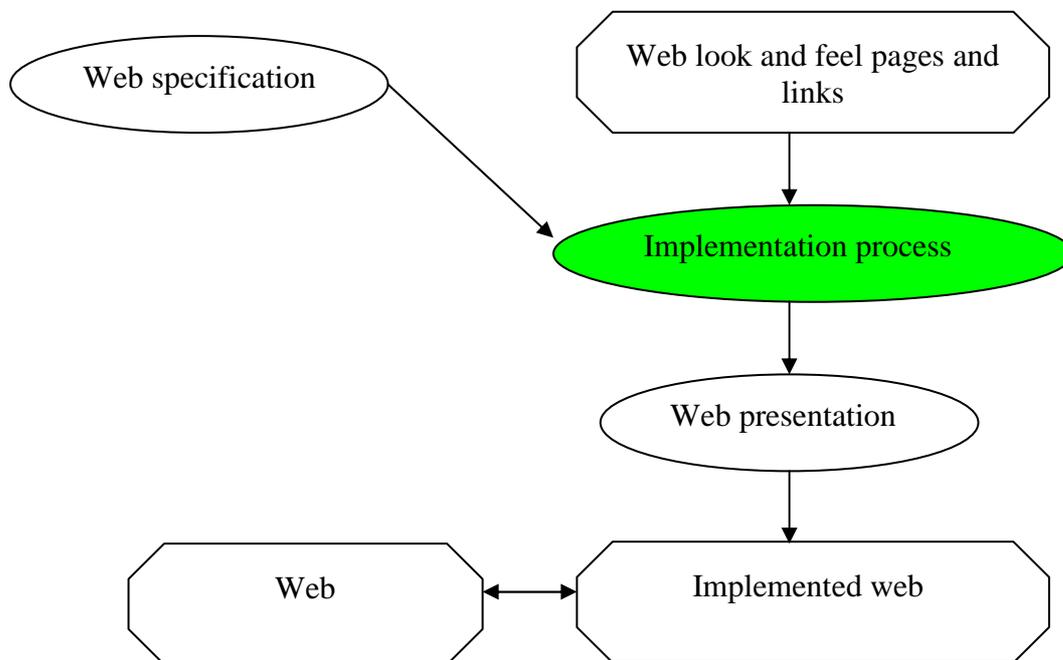
```
/*
  Document   : rstudents
  Created on : June 11, 2005, 1:01 PM
  Author    : Resarch Students
  Description:
    Purpose of the stylesheet follows.
*/

root {
  background-position: "100% 100%";
  background-color: "tan";
}
fst {
  display: "block";
  color: "blue";
  font-size: "16pt";
  font-weight: "bold";
  text-align: "left";
  text-indent: 1cm;
  padding: 10pt;
  background-position: "100% 100%";
  background-color: "tan";
}
```

**Figure 4.2: A sample CSS files for formatting data**

#### 4. OUR IMPLEMENTATION OVERVIEW

Figure 5 shows an overview of the implementation process. The objective of our implementation is to combine the output from the design and planning processes, as well as any updates or maintenance specified as a result of the analysis process, to create an effective Web. The key feature of the entire Web development process is that Web design, planning, and other processes are separated from Web implementation. This allows developers to make decisions about design that are independent of language-specific or implementation issues [19]. The job of the implementer is to bridge this gap between these abstract processes and the specific needs of implementation. The separation also helps designers and planners to focus on the needs of audience without disturbing the changing methods and practices for writing HTML or XML. The implemented Web is the object that users often consider to be the whole work of the Web for creating an effective Web.



**Figure.5: Web Implementation**

After the completion of Web design, the next step is to implement the Web within the limitations of its technical makeup used to define its specifications. The initial implementation is a prototype which is not released publicly, but available for analysis as used by a set of representative users.

The process essentially involves creating HTML/XML, Common Gateway Interface (CGI) programs (Java scripts and/or applets can also be used). The implementation process resembles software development because it involves using a specific syntax for creating hypertext structures in HTML/XML or writing programming language code statements in computer files.

The key features for our implementation practices are:

- (i) In the very beginning we create a stable, extendible directory and file structure to manage the Web's files and/or software components (including CGI /Java programs).
- (ii) The HTML/XML tools, such as editors and authoring environments, are used as required. It is wise to note at this point that some of these editors are "translators" from authoring programs meant for paper based publications and are not always customized for hypertext development.
- (iii) One should make sure of the Web's implementation in various browsers to ensure that the HTML/XML can be interpreted properly. Templates for supporting the consistent look and feel defined in the Web's design can also be used.

## **5. DISCUSSION AND CONCLUSION**

Developing information for the Web requires the focus on achieving the user needs. To accomplish this, the methodology involves six elements e.g. audience information, purpose statement, objectives list, domain information, Web and Web presentation. It also requires six continuously ongoing processes e.g. planning, analysis, design, implementation, promotion, innovation and information. A Web designer/implementer should have excellent knowledge of HTML and other languages, skills in using the computer systems, on which the Web is being deployed, and excellent file management and organizational abilities. He should also have a good writing skills, a talent for layout and design, and a sense of how the intended audience uses and thinks about the information presented in the Web.

The HTML based Web Design and implementation has several disadvantages, such as, the designer needs to emphasis more on presentation. Because it focuses mainly on a computer-to-human interface it has limited value in applications like transferring of information between databases (computer-computer). HTML has fixed a set of markup tags and lacks support for creating new, application specific tags. To overcome this, XML is one of the best choices for designer/implementer, which is more versatile than HTML. This new language lets information publishers invent their own tags and their benefits. One of the disadvantages of XML is that XLink does not support some of the older versions of Web browsers.

## **ACKNOWLEDGMENTS**

The present work has been carried out within the framework of the Associate Scheme of the Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy. The authors are grateful to ICTP for providing the necessary facilities and literature to prepare the manuscript. Dr. Tulshi Bezboruah is also thankful to Y.A. Ahmed, Centre for Energy Research and Training, Ahmadu Bello University, Nigeria for his valuable suggestions in designing the manuscript. The authors are also thankful to the Department of Science and Technology, Government of India for providing excellent computational, Networking and

Internet facilities to the Department of Electronics Science, Gauhati University, INDIA under FIST Program.

## REFERENCES

1. Berners-Lee, T., Information Management: A Proposal, CERN, 1989
2. Rosenfeld, L. and Morville, P.: Information Architecture for the World Wide Web/ Paperback / O'Reilly & Associates, 1998
3. Cailliau, R., Gilles, J., Cailliau, R.: How the Web was Born: The Story of the World Wide Web, Oxford University Press, ISBN 0192862073, 2000
4. <http://www.java.sun.com/>
5. <http://www.lucent.com/inferno/>
6. <http://www.realaudio.com/>
7. <http://www.macromedia.com/>
8. <http://www.vrml.org/>
9. <http://www.microsoft.com/>
10. <http://www.browserwatch.com/>
11. <http://www.netscape.com/>
12. <http://www.w3C.org/>
13. Len Seligman, Arnon Rosenthal, "XML's Impact on Databases and Data sharing", IEEE Computers 59-67, June 2001.
14. Allen Brown, Mathew Fuchs, Jonathan Robie, Philips Wadler, "MSL: a model for W3C XML Schema", Computer Networks 39 (2002) 507-521.
15. Maria-Angeles Grado-Caffaro and Martin Grado-Caffaro, "The challenges that XML faces", IEEE Computers, 15-18, October 2001.
16. Wenfei Fan, Gabriel M. Kuper, Jerome Simeon, "A unified constraint model for XML", Computer Networks, 39 (2002) 489-505.
17. Charles E. Campbell, Anrew Eisenberg and Jim Melton, "XML Schema", Sigmod Record, 32(2) 96-101, June 2003.
18. Peter Buneman, Susan Davidson, Wenfei Fan, Carmem Hara, Wang-Chiew Tan, "Keys for XML", Computer Networks 39 (2002) 473-487.
19. Tate, B., Crossing borders: Continuations, Web development, and Java programming, J2Life, LLC (2006)