NEERAJ®

WEB PROGRAMMING

By:
Gajendra Nayal, M.C.A

Reference Book Cum *Question Bank*

New Edition

NP

NEERAJ PUBLICATIONS

(Publishers of Educational Books)
(An ISO 9001 : 2008 Certified Company)
1507, Ist Floor, NAI SARAK, DELHI - 110 006
Ph. 011-23260329, 69414105 Fax 011-23285501
E-mail: info@neerajbooks.com
Website: www.neerajbooks.com

Price
₹ 160/-

Notes:

1. For the best & upto-date study & results, please prefer the recommended textbooks/study material only.

2. This book is just a Guide Book/Reference Book published by NEERAJ PUBLICATIONS based on the suggested syllabus by a particular Board /University.

3. The information and data etc. given in this Book are from the best of the data arranged by the Author, but for the complete and upto-date information and data etc. see the Govt. of India Publications/textbooks recommended by the Board/University.

4. Publisher is not responsible for any omission or error though every care has been taken while preparing, printing, composing and proof reading of the Book. As all the Composing, Printing, Publishing and Proof Reading etc. are done by Human only and chances of Human Error could not be denied. If any reader is not satisfied, then he is requested not to buy this book.

5. In case of any dispute whatsoever the maximum anybody can claim against NEERAJ PUBLICATIONS is just for the price of the Book.

6. If anyone finds any mistake or error in this Book, he is requested to inform the Publisher, so that the same could be rectified and he would be provided the rectified Book free of cost.

7. The number of questions in NEERAJ study materials are indicative of general scope and design of the question paper.

8. Subject to Delhi Jurisdiction only.

How to get Books by V.P.P.

CONTENTS

WEB PROGRAMMING

■■

WEB PROGRAMMING

Web 2.0 and XHTML

1

INTRODUCTION

In this chapter, we will discuss about various web development strategies. We will also discuss what is web? How to create a web page?

After that we will discuss about the web, we define the basic terminology of web like search, blogging social media, etc.

We will also introduce the XHTML. XHTML document can utilize the applications like scripts and applets that rely upon either the HTML DOM or XMLDOM. We will also discuss WML and syntactic difference between HTML and XHTML.

In this section, we will discuss about the standard XHTML document structure that help to create the web page, and we will also discuss about the XHTML elements like Images, Lists and Tables.

CHAPTER AT A GLANCE

WHAT IS WEB 2.0?

The term 'Web 2.0' refers to the idea of the "New Internet", or the second wave of the World Wide Web. Web 2.0 is not a specific application or technology, but explains two paradigm shifts within Information Technology, 'user-generated content' and 'thin client computing'.

User-generated content refers to social networking sites such as Facebook, Myspace and YouTube, Blogs, and any web application that enables users to create, elaborate, personal web pages without any prior technical programming knowledge. User-generated content of Web 2.0 is changing the way we use the Internet. Users have transformed the World Wide Web into a pool of knowledge and news that is created and reported on by 'citizen journalists'. Web 2.0 is radically changing journalism, creating new opportunities on the Internet and enhancing globalization at a pace faster than critics can comprehend. A major point of interest with Web 2.0 is the equalization on a mass scale between user, client and big corporations.

Thin Client Computing refers to data and applications that are housed on a web server, providing the user with universal access to information from any computer. Although, it is not a new concept for the World Wide Web, thin client computing has the potential to revolutionize the Internet into one giant application server for all users.

Web 2.0 also refers to the application that creates information to seek out the user and provide him with specific and pointed information. Algorithms are employed to direct this information based on the user's profile and browsing history.

The client-side/web browser technologies used in Web 2.0 development are Asynchronous JavaScript and XML (Ajax), Adobe Flash and the Adobe Flex framework, and JavaScript/Ajax frameworks such as YUI Library, Dojo Toolkit, MooTools, and jQuery. Ajax programming uses JavaScript to upload and download new data from the web server without undergoing a full page reload.

To allow users to continue to interact with the page, communications such as data requests going to the server are separated from data coming back to the page (asynchronously). Otherwise, user would have to routinely wait for the data to come back before they can do anything else on that page, just as a user has to wait for a page to complete the reload. This also increases overall performance of the site, as the sending of requests can complete quicker independent of blocking and queuing required to send data back to the client.

The data fetched by an Ajax request is typically formatted in XML or JSON (JavaScript Object Notation) format, two widely used structured data formats. Since both of these formats are natively understood by JavaScript, a programmer can easily use them to transmit structured data in their web application. When this data is received via Ajax, the JavaScript program then uses the Document Object Model (DOM) to dynamically update the web page based on the new data, allowing for a rapid and interactive user experience. In short, using these techniques, Web designers can make their pages function like desktop applications. For example, Google Docs uses this technique to create a Web based word processor.

Web 2.0 can be described in 3 parts which are as follows:

- **Rich Internet Application (RIA):** It defines the experience brought from desktop to browser whether it is from a graphical point of view or usability point of view. Some buzzwords related to RIA are Ajax and Flash.
- **Service-Oriented Architecture (SOA):** It is a key piece in Web 2.0 which defines how Web 2.0 applications expose their functionality so that other applications can leverage and integrate the functionality providing a set of much richer applications (Examples are: Feeds, RSS, Web Services, Mash-ups).
- **Social Web:** It defines how Web 2.0 tends to interact much more with the end user and make the end-user an integral part.

Web 2.0 websites include the following features and techniques:

Search: Finding information through keyword search.

Links: Connects information together into a meaningful information ecosystem using the model of the Web, and provides low-barrier social tools.

Authoring: The ability to create and update content leads to the collaborative work of many rather than just a few web authors. In wikis, users may extend, undo and redo each other's work. In blogs, posts and comments of individuals build up over time.

Tags: Categorization of content by users adding "tags" short, usually one-word descriptions to facilitate searching, without dependence on pre-made categories. Collections of tags created by many users within a single system may be referred to as "folksonomies" (i.e., folk taxonomies).

Extensions: Software that makes the Web an application platform as well as a document server. These include software like Adobe Reader, Adobe Flash player, Microsoft Silver light, ActiveX, Oracle Java, QuickTime, Windows Media, etc.

Signals: The use of syndication technology such as RSS to notify users of content changes.

INTRODUCTION TO WEB 2.0 TERMS

As such, Web 2.0 draws together the capabilities of client and server side software, content syndication and the use of network protocols. Standards-oriented web browsers may use plug-ins and software extensions to handle the content and the user interactions. Web 2.0 sites provide users with information storage, creation, and dissemination capabilities that were not possible in the environment now known as "Web 1.0".

SEARCH

Millions of searches are conducted each day on popular search engines by people all around the world. What are they looking for? A number of major search engines provide a way to glimpse into the web's query stream to discover the most popular search keywords or topics. These are:

- **AOL Hot Searches:** Top current queries, or see those in the last hour, last day and within particular categories.
- **Ask IQ:** See top searches at Ask.
- **Dogpile SearchSpy:** Choose to see either a filtered or non-filtered sample of top, real-time search terms from this popular meta search service. Sister site MetaCrawler offers a similar MetaCrawler MetaSpy service.
- **Google Trends:** Allows you to tap into Google's database of searches, to determine what's popular. View the volume of queries over time, by city, regions, languages and so on. Compare multiple terms, as well. See our review: Google Trends: Peer Into Google's Database Of Searches.
- **Google Zeitgeist:** What people are searching for at Google and its associated specialty services in a variety of categories. There are versions for various countries, as well.
- **Lycos 50:** Long-standing service showing top searches at Lycos each week.
- **MSN Search Insider:** Top 200 queries on MSN Search (annoyingly in random order), top "movers" in TV, sports and music, and a "duels" feature pitting top queries in a race against each other.
- **Yahoo Buzz Index:** Shows you what's hot and what's not in terms of search topics at Yahoo.

CONTENT NETWORKS

A content delivery network or content distribution network (CDN) is a system of computers containing copies of data placed at various nodes of a network. When properly designed and implemented, a CDN can improve access to the data it caches by increasing access bandwidth and redundancy and reducing access latency. Data content types often cached in CDNs include web objects, downloadable objects (media files, software, documents), applications, live streaming media, and database queries.

The Internet was designed according to the end-to-end principle. This principle keeps the core network relatively simple and moves the intelligence as much as possible to the network end-points: the hosts and clients. As a result, the core network is specialized, simplified, and optimized to only forward data packets.

Content Delivery Networks augment the end-to-end transport network by distributing on it a variety of intelligent applications employing techniques designed to optimize content delivery. The resulting tightly integrated overlay uses web caching, server-load balancing, request routing, and content services. These techniques are briefly described below:

Web caches store popular content on servers that have the greatest demand for the content requested. These shared network appliances reduce bandwidth requirements, reduce server load, and improve the client response times for content stored in the cache.

Server-load balancing uses one or more techniques including service based (global load balancing) or hardware based, i.e., layer 4-7 switches, also known as a web switch, content switch, or multilayer switch to share traffic among a number of servers or web caches. Here the switch is assigned a single virtual IP address. Traffic arriving at the switch is then directed to one of the real web servers attached to the switch. This has the advantage of balancing load, increasing total capacity, improving scalability, and providing increased reliability by redistributing the load of a failed web server and providing server health checks.

A content cluster or service node can be formed using a layer 4-7 switch to balance load across a number of servers or a number of web caches within the network.

Request routing directs client requests to the content source best able to serve the request. This may involve directing a client request to the service node that is closest to the client, or to the one with the most capacity. A variety of algorithms are used to route the request. These include Global Server Load Balancing, DNS-based request routing, dynamic metafile generation, HTML rewriting, and any casting. Proximity-choosing the closest service node-is estimated using a variety of techniques including reactive probing, proactive probing, and connection monitoring.

CDNs use a variety of methods of content delivery but not limited to manual asset copying, active web caches, and global hardware load balancers.

BLOGGING

Educational blogging takes advantage of the desire to express oneself and to receive feedback, but within the confines of the technology and the educational environment it is implemented in. And when done as part of a teacher or parent-initiated program, educational blogging starts with the assumption that the teacher or parent will be actively watching the content and the dialogue. The ability to contribute, through posting content and comments to the web, in an academic discipline accomplishes something of significance: it gives youth a vision of their ability to add to the accumulated knowledge and understanding of the world.

SOCIAL NETWORKING

Social networking sites, they include the ability to post written material to the web like blogs do, revolve much more around the ability of an individual to build a web "presence" and to create visible links with others in the network. What is written on a social networking site is often in computer slang, intended to be "different." These sites often allow the ability to place music, photo, and video content on the site, allowing the individual to showcase their likes and dislikes. A form of text messaging is often included as well, with a history of the messages appearing on the site. Again, all of the features of social networking sites are intended to create social interactions with others.

SOCIAL MEDIA

The term Social Media refers to the use of web-based and mobile technologies to turn communication into an interactive dialogue. Andreas Kaplan and Michael Haenlein define social media as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content. Social media are media for social interaction, as a superset beyond social communication. Enabled by ubiquitously accessible and scalable communication techniques, social media substantially change the way of communication between organizations, communities, as well as individuals. Social media take on many different forms, including magazines, Internet forums,

weblogs, social blogs, micro blogging, wikis, podcasts, photographs or pictures, video, rating and social bookmarking.

People obtain information, education, news and other data from electronic media and print media. Social media are distinct from industrial or traditional media, such as newspapers, television, and film. They are relatively inexpensive and accessible to enable anyone (even private individuals) to publish or access information, compared to industrial media, which generally require significant resources to publish information.

One characteristic shared by both social media and industrial media is the capability to reach small or large audiences; for example, either a blog post or a television show may reach no people or millions of people. Some of the properties that help describe the differences between social media and industrial media are:

1. **Reach:** Both industrial and social media technologies provide scale and are capable of reaching a global audience. Industrial media, however, typically use a centralized framework for organization, production, and dissemination, whereas social media are by their very nature more decentralized, less hierarchical, and distinguished by multiple points of production and utility.

2. **Accessibility:** The means of production for industrial media are typically government and/or privately owned; social media tools are generally available to the public at little or no cost.

3. **Usability:** Industrial media production typically requires specialized skills and training. Conversely, most social media production does not require specialized skills and training, or requires only modest reinterpretation of existing skills; in theory, anyone with access can operate the means of social media production.

4. **Immediacy:** The time lag between communications produced by industrial media can be long (days, weeks, or even months) compared to social media (which can be capable of virtually instantaneous responses; only the participants determine any delay in response). However, as industrial media begin adopting aspects of production normally associated with social media tools, this feature may not prove distinctive over time.

5. **Permanence:** Industrial media, once created, cannot be altered (once a magazine article is printed and distributed changes cannot be made to that same article) whereas social media can be altered almost instantaneously by comments or editing.

Community media constitute an interesting hybrid of industrial and social media. Though community-owned, some community radios, TV and newspapers are run by professionals and some by amateurs. They use both social and industrial media frameworks.

RICH INTERNET APPLICATIONS (RIAs)

A Rich Internet Application (RIA) is a Web application that has many characteristics of desktop application software, typically delivered either by way of a site-specific browser, via a browser plug-in, independent sandboxes, extensive use of JavaScript, or virtual machines. Users generally need to install a software framework using the computer's operating system before launching the application, which typically downloads, updates, verifies and executes the RIA. This is the main differentiator from JavaScript-based alternatives like Ajax that use built-in browser functionality to implement comparable interfaces. As can be seen on the List of Rich Internet application frameworks which includes even server-side frameworks, while some consider such interfaces to be RIAs, some consider them competitors to RIAs; and others, including Gartner, treat them as similar but separate technologies.

RIA present indexing challenges to Web search engines, but Adobe Flash content is now at least partially index able. Security can improve over that application software (for example, through use of sandboxes and automatic updates), but the extensions themselves remain subject to vulnerabilities and access is often much greater than that of native Web applications. For security purposes, most RIAs run their client portions within a special isolated area of the client desktop called a sandbox. The sandbox limits visibility and access to the file-system and to the operating system on the client to the application server on the other side of the connection. This approach allows the client system to handle local activities, calculations, reformatting and so forth, thereby lowering the amount and frequency of client-server traffic, especially versus client-server implementations built around so called thin clients.

Rich Internet Applications provide the end user with an interface that is faster and more responsive than traditional applications. Providing an unparallel