



Faculty of Science and Technology
Department of Computer Science

WEB TECHNOLOGIES (CSC 3222)

Lecture Note 1

Week 1

Supta Richard Philip
Richard@aiub.edu

Contents

Introduction to Web Technology	3
Introduction to Internet.....	3
History of Internet	3
World Wide Web	4
Client/Server model.....	5
Three-tier Client Server Architecture	5
Peer to Peer (P2P).....	6
Internet and World Wide Web (www)	6
Uniform Resource Identifier (URI)	7
Uniform Resource Locator (URL)	7
Browsers.....	8
Hypertext Transport (HTTP) Protocol.....	8
Process of http	9
HTTP Request and Response.....	10
Request Line	11
Status Line	11
HTML	12
A Simple HTML Document.....	12
HTML Tag	13
Exercise	14
HTML FORM	14
Exercise	15
XML.....	16
XHTML	17
DOM	18
DHTML	19
References	19

Introduction to Web Technology

In this lecture, we will learn about Web Technology and few terminologies: HTTP, HTML, XML and XHTML. We will learn about how web technology can help in the Business World as well.

Introduction to Internet

A global computer network providing a variety of information and communication facilities, consisting of interconnected networks using standardized communication protocols. "the guide is also available on the Internet" The Internet is the global system of interconnected computer networks that use the Internet protocol suite (TCP/IP) to link devices worldwide. It is a network of networks that consists of private, public, academic, business, and government networks of local to global scope, linked by a broad array of electronic, wireless, and optical networking technologies. The Internet carries a vast range of information resources and services.

History of Internet

This marvelous tool has quite a history that holds its roots in the cold war scenario. A need was realized to connect the top universities of the United States so that they can share all the research data without having too much of a time lag. This attempt was a result of Advanced Research Projects Agency (ARPA) which was formed at the end of 1950s just after the Russians had climbed the space era with the launch of Sputnik. After the ARPA got success in 1969, it didn't take the experts long to understand that how much potential can this interconnection tool have. In 1971 Ray Tomlinson made a system to send electronic mail. This was a big step in the making as these opened gateways for remote computer accessing i.e. telnet.

During all this time, rigorous paper work was being done in all the elite research institutions. From giving every computer an address to setting out the rules, everything was getting penned down. 1973 saw the preparations for the vital TCP/IP and Ethernet services. At the end of 1970s, Usenet groups had surfaced up. By the time the 80s had started, IBM came up with its PC based on Intel 8088 processor which was widely used by students and universities for it solved the purpose of easy computing. By 1982, the Defense Agencies made the TCP/IP compulsory and the term —internetl was coined. The domain name services arrived in the year 1984 which is also the time around which various internets based marked their debut. A worm, or a rust the computers, attacked in 1988 and disabled over 10% of the computer systems all over the world. While most of the researchers regarded it as an opportunity to enhance computing as it was still in its juvenile phase, quite a number of computer companies became interested in dissecting the cores of the malware which resulted to the formation Computer Emergency Rescue Team (CERT). Soon after the world got over with the computer worm, World Wide Web came into existence. Discovered by Tim Berners-Lee, World Wide Web was seen as a service to connect documents in websites using hyperlinks.

World Wide Web

The World Wide Web (abbreviated WWW or the Web) is an information space where documents and other web resources are identified by Uniform Resource Locators (URLs), interlinked by hypertext links, and can be accessed via the Internet. English scientist Tim Berners-Lee invented the World Wide Web in 1989. He wrote the first web browser computer program in 1990 while employed at CERN in Switzerland. The Web browser was released outside CERN in 1991, first to other research institutions starting in January 1991 and to the general public on the Internet in August 1991.

The World Wide Web has been central to the development of the Information Age and is the primary tool billions of people use to interact on the Internet. Web pages are primarily text documents formatted and annotated with Hypertext Markup Language (HTML). In addition to formatted text, web pages may contain images, video, audio, and software components that are rendered in the user's web browser as coherent pages of multimedia content.

Embedded hyperlinks permit users to navigate between web pages. Multiple web pages with a common theme, a common domain name, or both, make up a website. Website content can largely be provided by the publisher, or interactively where users contribute content or the content depends upon the users or their actions. Websites may be mostly informative, primarily for entertainment, or largely for commercial, governmental, or non-governmental organizational purposes.

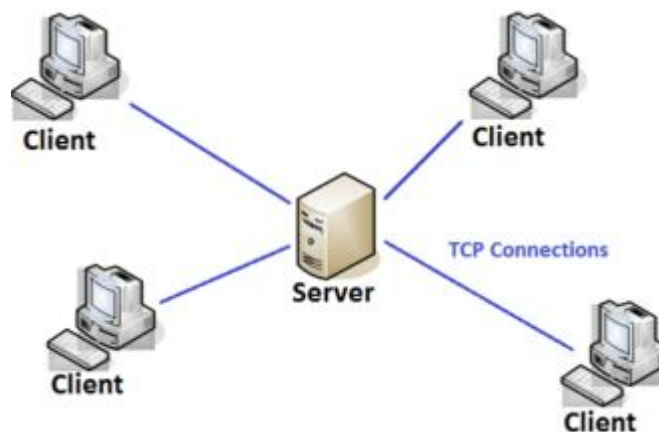


Figure 1: Client/Server model

WWW is another example of client/server computing. Each time a link is followed, the client is requesting a document (or graphic or sound file) from a server (also called a Web server) that's part of the World Wide Web that "serves" up the document. The server uses a protocol called HTTP or Hyper Text Transport Protocol. The standard for creating hypertext documents for the WWW is Hyper Text Markup Language or HTML. HTML essentially codes plain text documents so they can be viewed on the Web.

Client/Server model

Client Server Architecture is a computing model in which the server hosts, delivers and manages most of the resources and services to be consumed by the client. This type of architecture has one or more client computers connected to a central server over a network or internet connection. This system shares computing resources.

Client/server architecture is also known as a networking computing model or client/server network because all the requests and services are delivered over a network.

Three-tier Client Server Architecture

- The traditional client/server architecture involves two levels, a client level and a server level. Another common design of client/server systems uses three tiers:
- A client that interacts with the user
- An application server that contains the business logic of the application
- A resource manager that stores data.

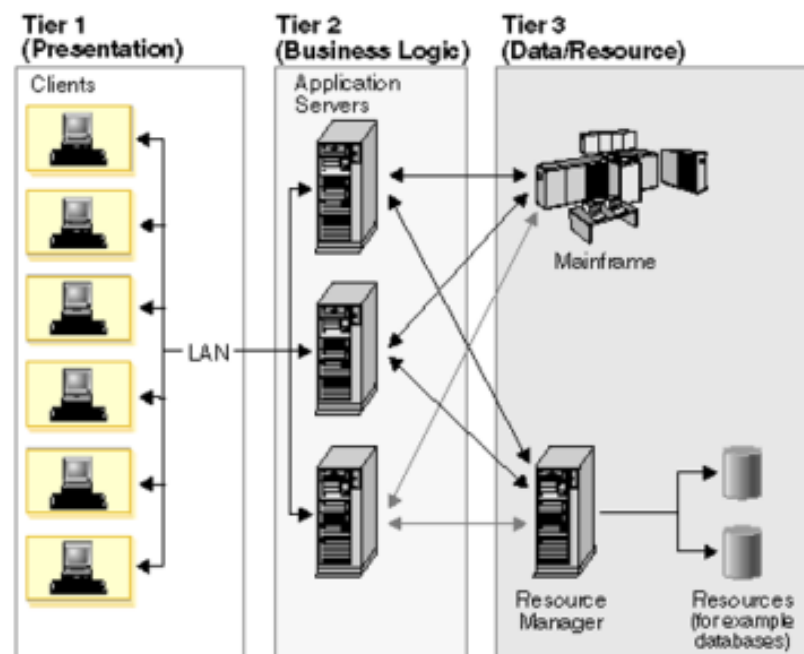


Figure 2 Three-tier Client Server Architecture

Peer to Peer (P2P)

- Distributed application architecture that partitions tasks or workloads between peers.
- Peers are equally privileged, equipotent participants in the application.
- They are said to form a peer-to-peer network of nodes.

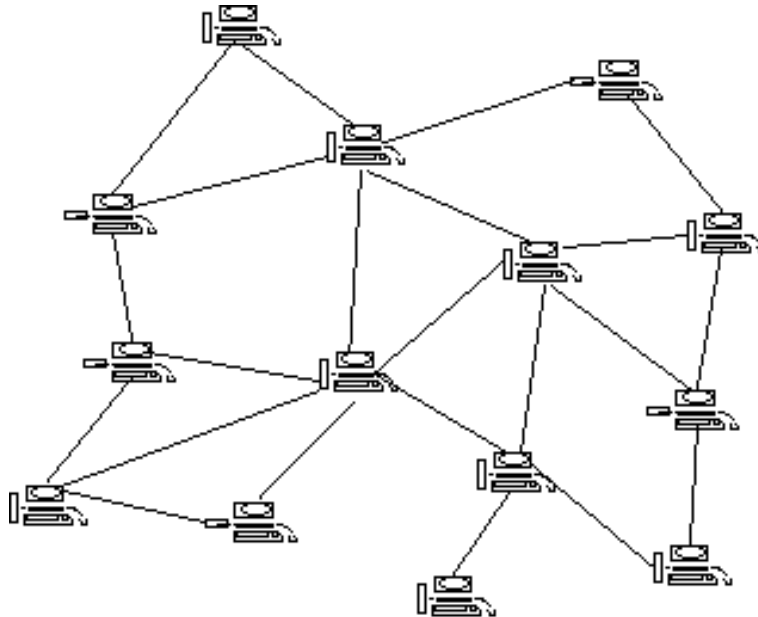


Figure 3 Peer to Peer (P2P)

Internet and World Wide Web (www)

- **Internet is used for Global system of interconnected computer networks.**
 - Use the **standard Internet Protocol Suite (TCP/IP)**.
 - Serve billions of users worldwide.
 - Network of networks that consists of millions of private, public, academic, business, and government networks.
 - Carries a vast range of **information resources** and services, such as the inter-linked **hypertext** documents of the **World Wide Web** and the infrastructure to support **electronic mail**.
- **World Wide Web** commonly known as the **Web (Abbreviated as WWW)**.
- A system of interlinked hypertext documents accessed via the Internet. With a web browser, anyone can view web pages that may contain text, images, videos, and other multimedia and navigate between them via hyperlinks.
- Internet is global system of interconnected computer networks. In short, the Web is an application running on the Internet.

- WWW is one of the services that run on the Internet. It is a collection of interconnected documents and other resources, linked by hyperlinks and URLs. An application running on the Internet.
- Viewing a web page on the **World Wide Web** normally begins
 - typing the URL of the page into a web browser,
 - Following a hyperlink to that page or resource.
- The web browser then initiates a series of communication messages, behind the scenes, in order to fetch and display it.
- First, the server-name portion of the URL is resolved into an **IP address** using the global, distributed Internet database known as the **Domain Name System (DNS)**.
- The browser then requests the resource by sending an HTTP request to the Web server at that particular address.
- While receiving these files from the web server, browsers may progressively render the page onto the screen as specified by its HTML, Cascading Style Sheets (CSS), or other page composition languages.
- Uniform Resource Identifier (URI)- a universal system for referencing resources on the Internet, such as hypertext documents and images.
- HyperText Transfer Protocol (HTTP)- how the browser and server authenticate each other.

Uniform Resource Identifier (URI)

A Uniform Resource Identifier (URI) is a string of characters that unambiguously identifies a particular resource. To guarantee uniformity, all URIs follow a predefined set of syntax rules. The most common form of URI is the Uniform Resource Locator (URL), frequently referred to informally as a web address.

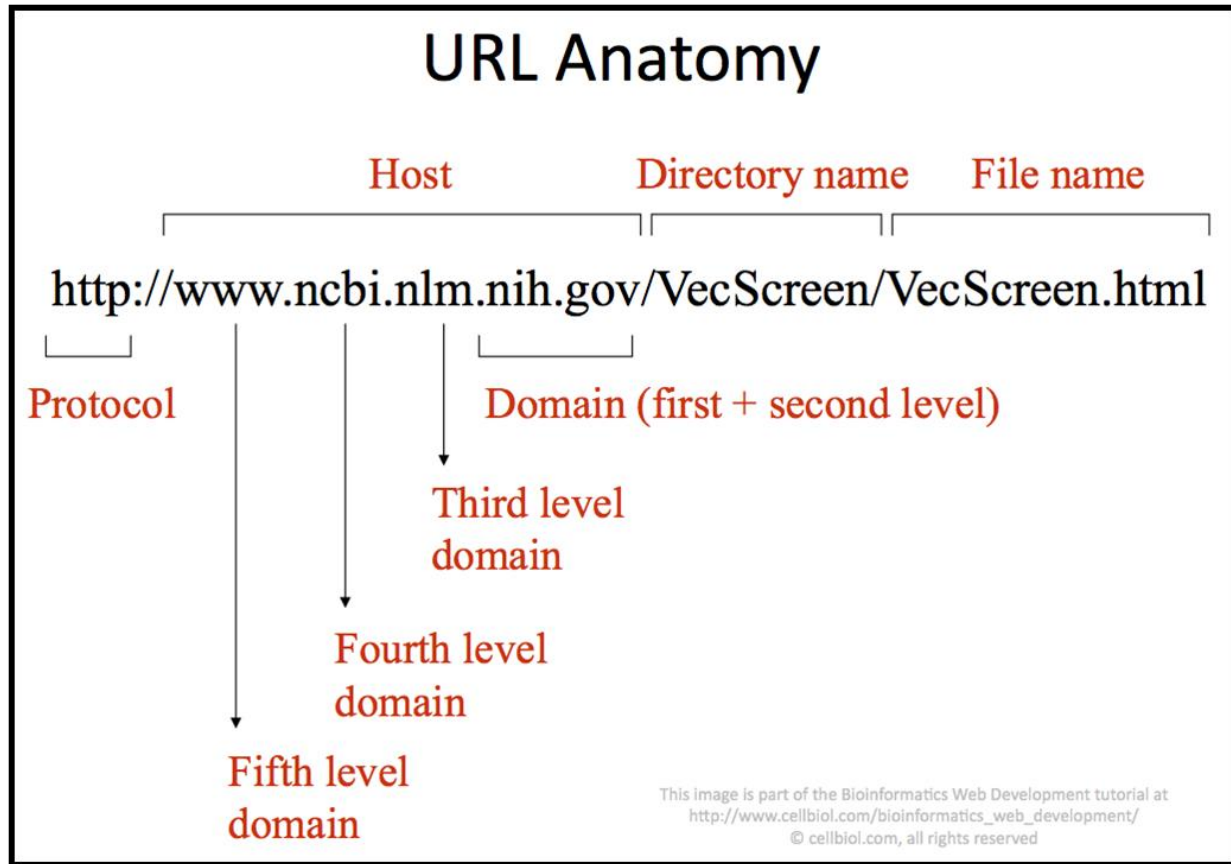
Uniform Resource Locator (URL)

A Uniform Resource Locator or URL is the address of a document found on the WWW. Browser interprets the information in the URL in order to connect to the proper Internet server and to retrieve your desired document. Each time a click on a hyperlink in a WWW document instructs browser to find the URL that's embedded within the hyperlink. The elements in a URL: **Protocol://server's address/filename**

Hypertext protocol: <http://www.aiub.edu>

File Transfer Protocol: <ftp://ftp.dartmouth.edu>

Telnet Protocol: <telnet://pac.carl.org>

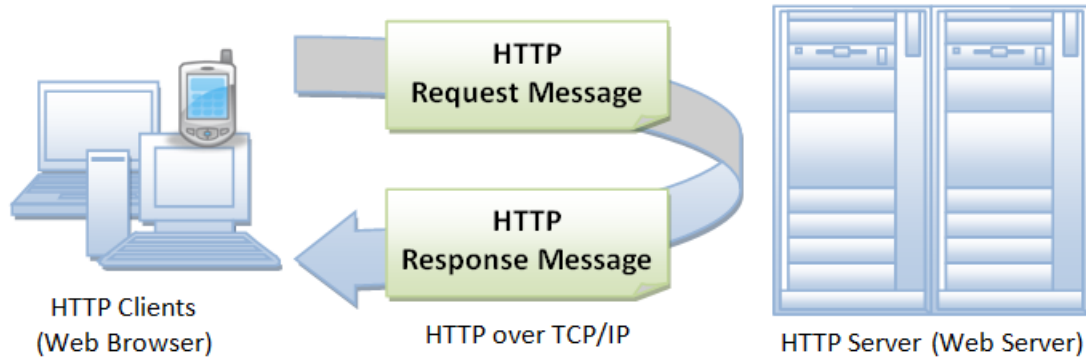


Browsers

WWW Clients or "Browser": The program you use to access the WWW is known as a browser because it "browses" the WWW and requests these hypertext documents. Browsers can be graphical, allows seeing and hearing the graphics and audio; text-only browsers (i.e., those with no sound or graphics capability) are also available. All of these programs understand http and other Internet protocols such as FTP, gopher, mail, and news, making the WWW a kind of "one stop shopping" for Internet users.

Hypertext Transport (HTTP) Protocol

HTTP (Hypertext Transfer Protocol) is perhaps the most popular application protocol used in the Internet (or The WEB).



- HTTP is an asymmetric request-response client-server protocol as illustrated. An HTTP client sends a request message to an HTTP server. The server, in turn, returns a response message. In other words, HTTP is a pull protocol, the client pulls information from the server (instead of server pushes information down to the client).
- HTTP is a stateless protocol. In other words, the current request does not know what has been done in the previous requests.
- HTTP permits negotiating of data type and representation, so as to allow systems to be built independently of the data being transferred.
- Quoting from the RFC2616: "The Hypertext Transfer Protocol (HTTP) is an application-level protocol for distributed, collaborative, hypermedia information systems. It is a generic, stateless, protocol which can be used for many tasks beyond its use for hypertext, such as name servers and distributed object management systems, through extension of its request methods, error codes and headers."

Process of http

Whenever you issue a URL from your browser to get a web resource using HTTP, e.g. `http://www.nowhere123.com/index.html`, the browser turns the URL into a *request message* and sends it to the HTTP server. The HTTP server interprets the request message, and returns you an appropriate response message, which is either the resource you requested or an error message. This process is illustrated below:

A URL (Uniform Resource Locator) is used to uniquely identify a resource over the web. URL has the following syntax:

Protocol://hostname:port/path-and-file-name

There are 4 parts in a URL:

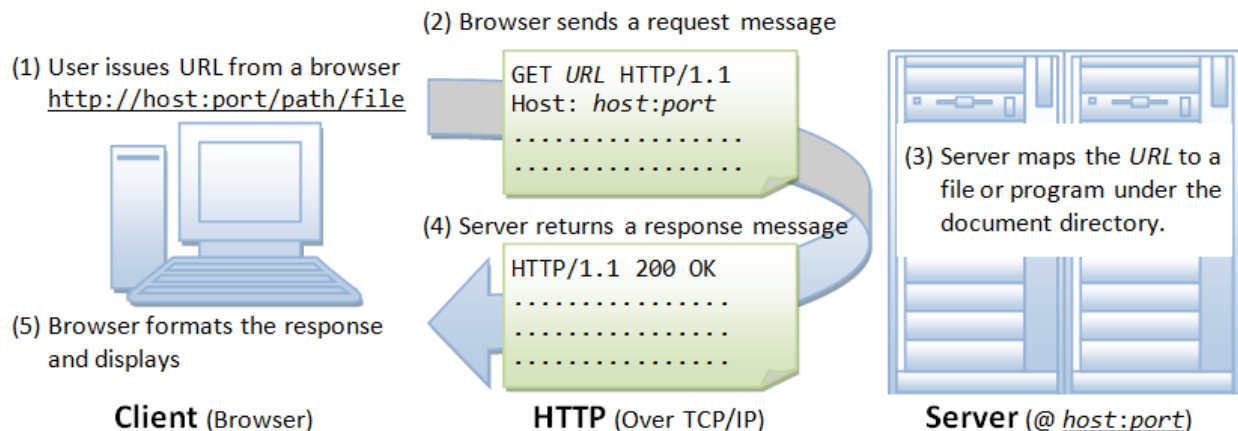
Protocol: The application-level protocol used by the client and server, e.g., HTTP, FTP, and telnet.

Hostname: The DNS domain name (e.g., `www.nowhere123.com`) or IP address (e.g., `192.128.1.2`) of the server.

Port: The TCP port number that the server is listening for incoming requests from the clients.

Path-and-file-name: The name and location of the requested resource, under the server document base directory.

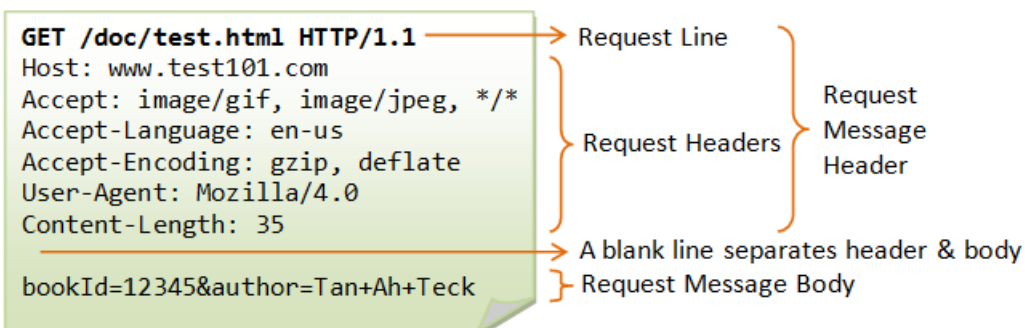
For example, in the URL `http://www.nowhere123.com/docs/index.html`, the communication protocol is HTTP; the hostname is `www.nowhere123.com`. The port number was not specified in the URL, and takes on the default number, which is TCP port 80 for HTTP. The path and file name for the resource to be located is `/docs/index.html`.



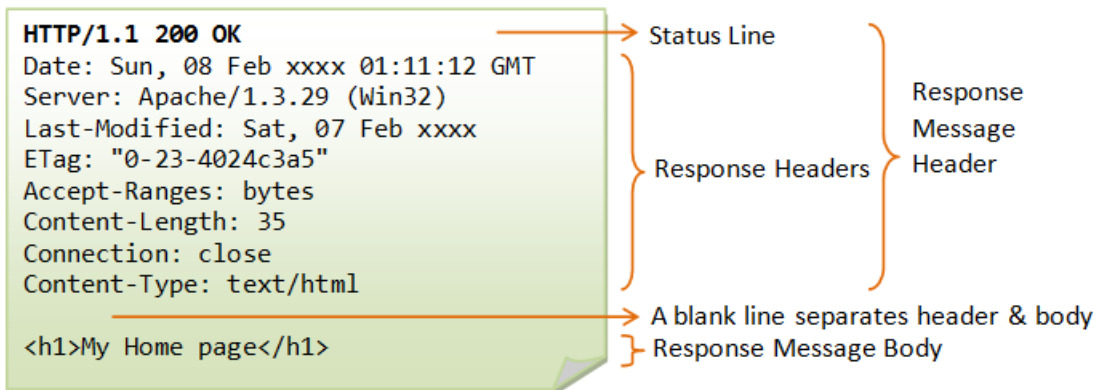
HTTP Request and Response

As mentioned, whenever you enter a URL in the address box of the browser, the browser translates the URL into a request message according to the specified protocol; and sends the request message to the server.

For example, the browser translated the URL `http://www.test101.com/doc/test.html` into the following request message:



The following shows a sample response message:



Request Line

The first line of the header is called the *request line*, followed by optional request headers.

Examples of request line are:

GET: A client can use the GET request to get a web resource from the server.

```
GET /test.html HTTP/1.1
```

POST: Used to post data up to the web server. HTML form uses POST request.

```
POST /index.html HTTP/1.1
```

PUT: Ask the server to store the data.

```
PUT /index.html HTTP/1.1
```

DELETE: Ask the server to delete the data.

```
DELETE /index.html HTTP/1.1
```

Status Line

The first line is called the *status line*, followed by optional response header(s).

Examples of status line are:

```
HTTP/1.1 200 OK
```

```
HTTP/1.0 404 Not Found
```

```
HTTP/1.1 403 Forbidden
```

HTML

- HTML stands for Hyper Text Markup Language
- HTML describes the structure of a Web page
- HTML consists of a series of elements
- HTML elements tell the browser how to display the content
- HTML elements are represented by tags
- HTML tags label pieces of content such as "heading", "paragraph", "table", and so on
- Browsers do not display the HTML tags, but use them to render the content of the page

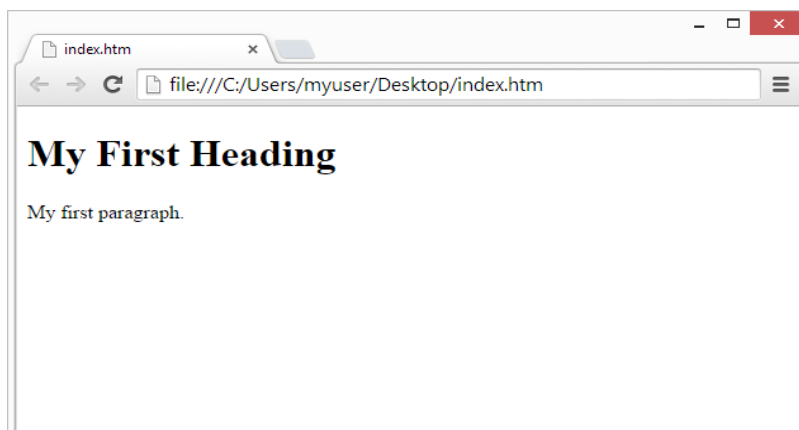
A Simple HTML Document

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>

<h1>My First Heading</h1>
<p>My first paragraph.</p>

</body>
</html>
```

- The `<!DOCTYPE html>` declaration defines this document to be HTML5
- The `<html>` element is the root element of an HTML page
- The `<head>` element contains meta information about the document
- The `<title>` element specifies a title for the document
- The `<body>` element contains the visible page content
- The `<h1>` element defines a large heading
- The `<p>` element defines a paragraph



HTML tags are element names surrounded by angle brackets:

<tagname>content goes here...</tagname>

HTML Tag

HTML Headings

<h1>Heading 1</h1>

-----upto-----

<h6>Heading 6</h6>

HTML Paragraphs

<p>This is a paragraph.</p>

HTML Links

Visit our HTML tutorial

HTML Images

HTML Images

HTML Tables

<table style="width:100%">

<tr>

<th>Firstname</th>

<th>Lastname</th>

<th>Age</th>

</tr>

<tr>

<td>Jill</td>

<td>Smith</td>

<td>50</td>

</tr>

</table>

HTML Lists

<ul style="list-style-type:disc;">

Coffee

Tea

Block element div in HTML:

<div>Hello World</div>

Exercise

1. Design a personal CV/your web page using html elements.

HTML FORM

- An HTML form is used to collect user input. The user input can then be sent to a server for processing.
- The HTML `<form>` element defines a form that is used to collect user input.
- The `<input>` element is the most important form element.
- The `<input>` element is displayed in several ways, depending on the **type** attribute.
 - `<input type="button">`
 - `<input type="checkbox">`
 - `<input type="date">`
 - `<input type="email">`
 - `<input type="file">`
 - `<input type="hidden">`
 - `<input type="image">`
 - `<input type="password">`
 - `<input type="radio">`
 - `<input type="range">`
 - `<input type="reset">`
 - `<input type="submit">`
- **The Action Attribute:** The **action** attribute defines the action to be performed when the form is submitted. If the **action** attribute is omitted, the action is set to the current page.
- **Method Attribute:** The **method** attribute specifies the HTTP method (**GET or POST**) to be used when submitting the form data.

```
<form action="/action_page.php" method="get">
<label for="fname">First name:</label>
<input type="text" id="fname" name="fname"><br>
<label for="lname">Last name:</label>
<input type="text" id="lname" name="lname"><br>

<label for="email">Email:</label>
<input type="text" id="email" name="email"><br>

<input type="radio" id="male" name="gender" value="male">
<label for="male">Male</label><br>
<input type="radio" id="female" name="gender" value="female">
<label for="female">Female</label><br>

<input type="submit" value="Submit">

<input type="reset">

</form>
```

HTML form

First name:

Last name:

E-mail:

☐ Male
☐ Female

Exercise

1. Design a login form, change password form and registration form as follows:

REGISTRATION

Name :

Email : i

User Name :

Password :

Confirm Password :

Gender

☐ Male ☐ Female ☐ Other

Date of Birth

/ / (dd/mm/yyyy)

LOGIN

User Name :

Password :

☐ Remember Me

[Forgot Password?](#)

CHANGE PASSWORD

Current Password :

New Password :

Retype New Password :

XML

- XML stands for eXtensible Markup Language.
- XML was designed to store and transport data.
- XML was designed to be both human- and machine-readable.

```
<?xml version="1.0" encoding="UTF-8"?>
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```

- Save the file **name.xml**
- XML was designed to carry data, not to display data
- XML tags are not predefined. You must define your own tags
- XML is designed to be self-descriptive.
- XML is a W3C Recommendation
- Maybe it is a little hard to understand, but XML does not DO anything. XML was created to structure, store, and transport information.

- XML is the most common tool for data transmissions between all sorts of applications.
- XML Documents Form a Tree Structure
- XML documents must contain a root element. This element is "the parent" of all other elements.
- The elements in an XML document form a document tree.
- The tree starts at the root and branches to the lowest level of the tree.
- All elements can have sub elements (child elements)

XML elements must follow these naming rules:

- Names can contain letters, numbers, and other characters
- Names cannot start with a number or punctuation character
- Names cannot start with the letters xml (or XML, or Xml, etc)
- Names cannot contain spaces
- Any name can be used, no words are reserved.
- All XML Elements Must Have a Closing Tag
- XML Tags are Case Sensitive
- XML Elements Must be Properly Nested
- XML Documents Must Have a Root Element
- XML Attribute Values Must be Quoted

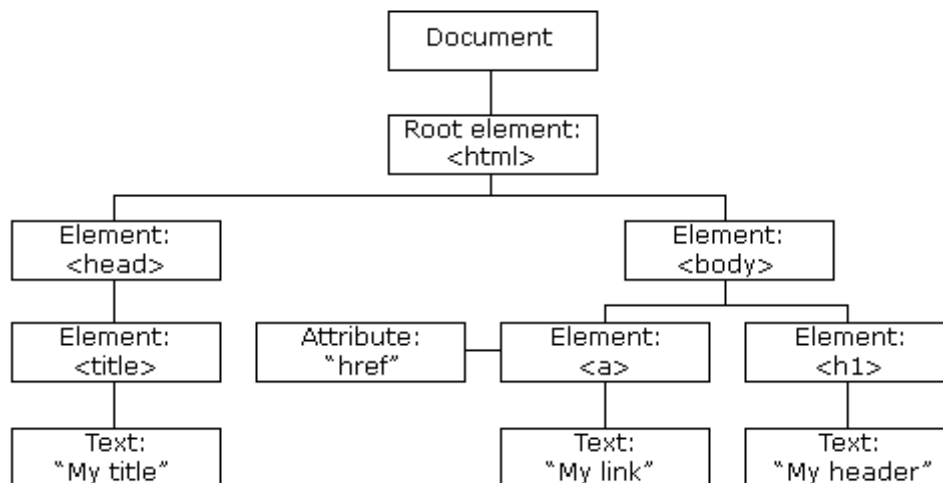
XHTML

- XHTML is a stricter, more XML-based version of HTML.
- XHTML stands for EXtensible HyperText Markup Language
- XHTML is a stricter, more XML-based version of HTML
- XHTML is HTML defined as an XML application
- All XHTML elements must be closed.
- XHTML elements must be properly nested
- `<i>bold and italic</i>` is wrong
- XHTML documents must be well-formed
- XML is a markup language where everything must be marked up correctly, which results in "well-formed" documents. Today's market consists of different browser technologies, some browsers run on computers, and some browsers run on mobile phones or other small devices. The last-mentioned do not have the resources or power to interpret a "bad" markup language.
- Combining the strengths of HTML and XML, W3C recommended a markup language i.e XHTML.

- XHTML elements must be properly nested
- XHTML elements must always be closed
- XHTML elements must be in lowercase
- XHTML documents must have one root element

DOM

- DOM a standard for accessing documents like XML and HTML
- "The W3C Document Object Model (DOM) is a platform and language-neutral interface that allows programs and scripts to dynamically access and update the content, structure, and style of a document."
- The DOM is separated into 3 different parts / levels:
- Core DOM - standard model for any structured document
- XML DOM - standard model for XML documents
- HTML DOM - standard model for HTML documents
- The DOM defines the objects and properties of all document elements, and the methods (interface) to access them.



When you want to access HTML elements with JavaScript, you have to find the elements first.

There are a couple of ways to do this:

- Finding HTML elements by id
- Finding HTML elements by tag name
- Finding HTML elements by class name
- Finding HTML elements by CSS selectors
- Finding HTML elements by HTML object collections.

DHTML

- DHTML stands for Dynamic HTML, it is totally different from HTML.
- The DHTML is based on the properties of the HTML, JavaScript, CSS, and DOM (Document Object Model which is used to access individual elements of a document) which helps in making dynamic content.
- The DHTML make use of Dynamic object model to make changes in settings and also in properties and methods.
- **DHTML Example(Changing HTML Content)**
 - The HTML document above contains an <h1> element with id="id01"
 - We use the HTML DOM to get the element with id="id01"
 - A JavaScript changes the content (innerHTML) of that element to "New Heading"

```
<!DOCTYPE html>
<html>
<body>
<h1 id="id01">Old Heading</h1>
<script>
var element = document.getElementById("id01");
element.innerHTML = "New Heading";
</script>
</body>
</html>
```

- The following example changes the style of a <p> element:

```
<html>
<body>
<p id="p2">Hello World!</p>
<script>
document.getElementById("p2").style.color = "blue";
</script>
<p>The paragraph above was changed by a script.</p>
</body>
</html>
```

References

- [1] https://www.ntu.edu.sg/home/ehchua/programming/webprogramming/HTTP_Basics.html
- [2] <https://www.w3schools.com/html/default.asp>
- [3] https://w3schools.sinsixx.com/dhtml/dhtml_javascript.asp.htm