V Skills
Skills for a secure future

Certified HTML5 Developer Sample Material

V-Skills Certifications

A Government of India
&
Government of NCT Delhi Initiative

V-Skills

Skills for a secure future

1. HTML5

1.1 Introduction

Printed information around us whether in the form of newspaper, magazine, books or printed forms; are divided into small parts like a magazine into articles and each article has its own heading, sometimes a summary and then followed by the article which is further divided into paragraphs. Similarly a web site has many web pages composed of images, text, links to other pages and audio or video for an enriching experience.

Web is a collection of documents that all link together, and bear a strong similarity to the printed documents around us. Web pages are written in HTML (Hypertext Markup Language) or XHTML (eXtensible Hypertext Markup Language). Both are document-layout and hyperlink-specification language. They define how to display the contents of the document, including text, images, and other support media. The language also tells how to make hypertext links, which connect document with other documents.

In keeping with the principle of separation of concerns, the function of HTML is primarily to add structural and semantic information to the raw text of a document. Presentation and behavior are separate functions, which can be added as desired, ideally through links to external documents such as style sheets, graphics files, and scripts.

1.2 History

Physicist Berners-Lee specified HTML and wrote the browser, server software and the World Wide Web in the late 1990. World Wide Web's systems enabled hypertext linking, so documents automatically reference other documents, located anywhere thus, giving convenience and increasing productivity.

1.3 HTML Versions

There have been several versions of HTML and is overseen by an organization called the W3C (World Wide Web Consortium). The last major version of HTML was HTML 4.01 in December 1999. In January 2000, some stricter rules were added to HTML 4.01, called as XHTML (Extensible Hypertext Markup Language). HTML 5 is the latest revision of the HTML standard and currently remains under development.

1.4 HTML5 Enhancements

There have been several enhancements in HTML5, which are summarized as
HTML 5 New Doctype and Charset
HTML5 new tags – HTML5 has created tags to support newer structural elements as
✓ <section> - to define sections of pages
✓ <header> - defines the header of a page
✓ <footer> - defines the footer of a page
✓ <nav> - defines the navigation on a page
✓ <article> - defines the article or primary content on a page
✓ <aside> - defines extra content like a sidebar on a page
✓ <figure> - defines images that annotate an article

HTML 5 New Inline Elements - These new inline elements define some basic concepts and keep them semantically marked up as

- ✓ <mark> - to indicate content that is marked in some fashion
- ✓ <time> - to indicate content that is a time or date
- ✓ <meter> - to indicate content that is a fraction of a known range - such as disk usage
- ✓ <progress> - to indicate the progress of a task towards completion

HTML 5 New Dynamic Pages Support - HTML 5 was developed to help Web application developers, so there are a lot of new features to make it easy to create dynamic HTML pages as

- ✓ Context menus creation and use within web pages and applications
- ✓ async attribute - This is added to the script tag to tell the browser that the script should be loaded asynchronously so that it doesn't slow down the load and display of the rest of the page.
- ✓ <details> - provides details about an element like tooltips.
- ✓ <menu> - an old tag brought back and given new life allowing to create a menu system on your Web pages
- ✓ <command> - defines actions that should happen when a dynamic element is activated

HTML 5 New Form Types - HTML 5 adds a few more form input types as datetime, datetime-local, date, month, week, time, number, range, email and url

HTML 5 New Elements - There are a few exciting new elements in HTML 5 as

- ✓ <canvas> - an element to give a drawing space in JavaScript on Web pages.
- ✓ <video> - add video to Web pages with as a tag.
- ✓ <audio> - add sound to Web pages with as a tag.

1.5 Elements, Tags and Attributes

HTML is an embedded language, the language's directions or tags are inserted into the HTML document that browser loads for viewing. The web browser uses the information inside the HTML tags to decide how to display or otherwise treat the subsequent contents of a HTML document.

Similar to a word processor like MS-Word in which styles are applied to the text; in a HTML document markups or tags are applied to stylize text as bold or italicize. Specific tags are applied on the text for specific styling.

HTML documents are composed of a tree of HTML elements. Each element can have attributes specified. Elements can also have content, including other elements and text. HTML elements represent semantics, or meaning. For example, the title element represents the title of the document. In the HTML syntax, most elements are written with a start tag and an end tag, with the content in between like

<p>In the HTML syntax, most elements are written ...</p>

However, not all of these elements require the end tag, or even the start tag, to be present like the br element.

HTML element – It represents structure or semantics and generally consists of a start tag, content, and an end tag for example, following is a bold element

 This is in bold or more dark.

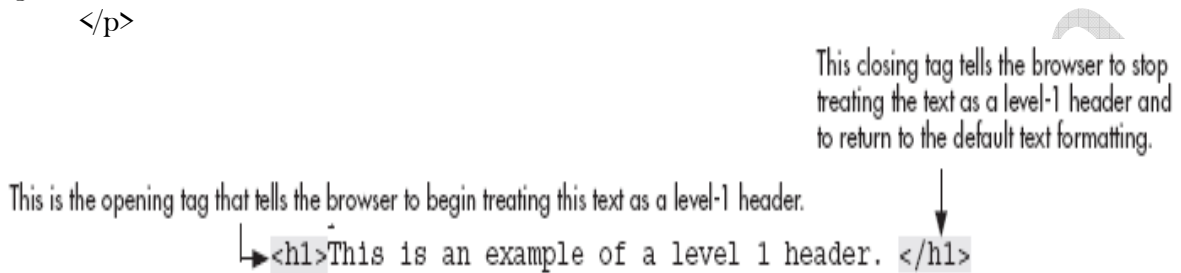**HTML tags –** They are used to mark the start and end of an HTML element.

Start tag has opening angle bracket (<) followed by element name, zero or more space separated attribute/value pairs, and a closing angle bracket (>).

A start tag with no attributes: <p>
A start tag with an attribute   <p class="info">

End tag has opening angle bracket followed by forward slash, the element name, and a closing angle bracket
        </p>

This closing tag tells the browser to stop treating the text as a level-1 header and to return to the default text formatting.

This is the opening tag that tells the browser to begin treating this text as a level-1 header.

        <h1>This is an example of a level 1 header. </h1>

Empty tag There are also some elements that are empty, meaning that they only consist of a single tag and do not have any content and look like opening tags

In XHTML. empty elements must have an end tag or the start tag must end with '/>' so it is used as

Difference - A tag consists of a left- and right-angle bracket and letters and numbers between those brackets, but an element is the opening and closing tags plus anything between the two tags.

HTML Attributes – It defines property for an element, consists of an name and value pair appearing within the element's start tag. Multiple name and value pairs are space separated.

All are illustrated in the figure

Opening **tag**

Element

<html xmlns="http://www.w3.org/1999/xhtml">

Attribute

</html>

Closing **tag**

1.6 Head and body tags

A web page is contained between the opening <html> and closing </html> tags.

<html> element

This element contains the whole HTML document. HTML document should have one opening <html> tag and document should end with a closing </html> tag. It has <head> and <body> tags in it.

<head> element

It is the head of the page, having information about the page but not the content of the page as it contain a title and a description of the page, or location for CSS and JavaScript files. The element has an opening <head> tag, the closing </head> tag, and everything in between.

<body> element

It is the body of the web page and has all the content seen on the web page. It consists of the opening <body> tag, closing </body> tag, and everything in between. Usually it appears after <head> element and the contents in it are also called as body content. It can have some paragraphs or complicated layouts having forms and tables for example

```
<body>
  <p>This is a paragraph tag.</p>
</body>
```

Nesting

An element containing other elements, must wholly contain that element. This is called as nesting elements correctly.

1.7 HTML Editor

HTML editors are used help a HTML5 developer efficiently and effectively develop web pages in HTML5 which are not only compliant to standards but also fulfill the user need. Some popular HTML editors are listed

✓ Amaya (www.w3.org/Amaya) Open-source software on Windows/Mac/Linux and has WYSIWYG visual editor and Spell-checking

✓ CoffeeCup (www.coffeecup.com) Paid software on Windows only and has WYSIWYG visual editor, FTP upload, Spell-checking, and Templates

✓ Dreamweaver (www.adobe.com/products/dreamweaver) Paid software on Windows/Mac and has WYSIWYG visual editor with browser preview, FTP upload, Spell-checking, Templates, Server-side scripting and Multi-user editing

1.8 Create a web page

The whole web page with html, head and body tag is made by joining earlier sections as

```
<html>

<head>
<title>Head tag example </title>
<meta name="Keywords" content=" Head tag, example" />
<meta name="description" content=" Head tag example" />
<base href="http://www.vskills.com" />
```

```
<link rel="stylesheet" type="text/css" href="style.css" />
<script type="text/javascript">
function show_alert()
{
alert("This is confidential page!");
}
</script></head>

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

</html>
```
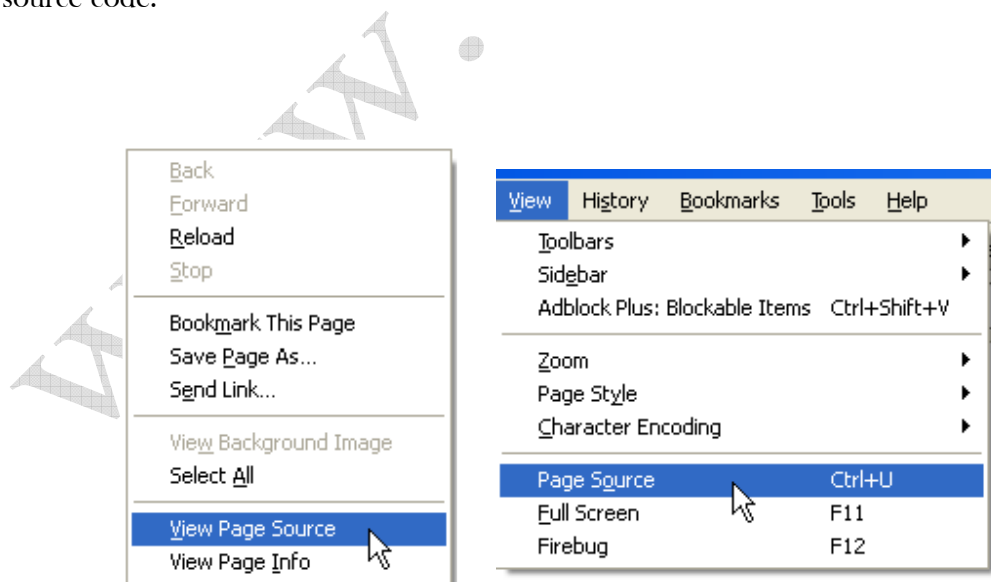
1.9 Viewing the Source

People learn about HTML by using the view source facility provided by almost all the browsers. Following steps are to be taken

✓ In the web browser, open the page whose code we like to see.
✓ Choose View > Page Source or Page > View Source, or View Source.

Viewing existing HTML code usually facilitate an inexperienced person to locate areas that cause problems. Furthermore it enables to learn various features on existing web pages that may encounter while surfing the web. There are two ways of viewing source HTML

Right clicking – Right click of mouse on the web page, and then clicking View source. In case of frames, source of frame is accessed and not that of the web page.

Menu Bar – In view menu which is usually present in all browsers, click the source option to view the source code.

Right Click   View Menu

HTML code can easily be copied to be used for other pages or they can be taken from existing web pages, and then modifying that code to suite needs.

1.10 White Space and Flow

Spaces, tabs and new lines are all called as white spaces. White spaces given in the source HTML is taken as a single space and would also appear as one space by the browser. This phenomenon is called white space collapsing. A new line or consecutive empty lines, all these are treated as one space.

<p>The
lines of
this paragraph
are present on
multiple lines
and also space between
words are

treated as a single space. Also

the large spaces between
some of the
words will also not appear
in the

browser.

This is how white space just collapses as a single space.</p>

Browser displays text in the full width of the screen and wraps the text onto new lines when no space is present. White space collapsing helps in adding spaces in web page for commenting which does not show in a browser as shown

The lines of this paragraph are present on multiple lines and also space between words are treated as a single space. Also the large spaces between some of the words will also not appear in the browser. This is how white space just collapses as a single space.

1.11 HTML Comments

It is part of HTML which is ignored by web browser and is used usually for commenting HTML tag blocks, to indicate sections of a document or any specific instructions about it. It helps to understand the HTML code more easily by adding comments to them. But if comments are used around HTML code then, the code block will not be shown hence, commenting is to be done with care and precaution.

Single Line Comment

It is indicated by special starting tag "<!—"and closing tag "-->" placed at the beginning and end of line to be treated as a comment. Comments do not nest and there should be no space in the start-of-comment string like "< !—" will not be treated as comment but will be shown on the screen. An example of valid comment is

<!--   This is single line comment. -->

Multi line Comments

Commenting multiple lines is done by placing the multiple or different lines between the starting tag "<!—" and closing tag "-->"to be treated as a comment like

```
<!--
This is multi line comment
and can comment many lines
present in this block.
-->
```

Conditional Comments

They are supported by Internet Explorer 5 onwards on Windows, and give instructions specific to Internet Explorer for browser specific instructions. They can also distinguish between different versions 5.0, 5.5 and 6.0 for example

```
<!--[if IE 6]>
Special instructions for IE 6 here
<![endif]-->
```

They are similar to a comment by using "<!-- -->"so that other browsers treat them as comment but internet explorer recognize it and executes the instructions.

Comment tag

Some browsers support <comment> tag for commenting.

```
<comment>This is a comment.</comment>
```

1.12 HTML Meta Tags

Meta tags specify metadata about the document. Metadata is information about a document and not the content of document and usually include items like author, keywords, date and time of publishing the content, etc. Meta element is used and has name/value pairs describing metadata. The <meta> tag being a empty element does not have a closing tag, but have information in attributes.

Metadata given by meta tag is important part of webpage as it helps search engines to index the webpage as per different search criteria so as to find best match for a user search. Search engines usually scan keywords on the webpage for its relevance. <meta> tags are placed between <head> and </head> tags and its attributes are discussed

Attribute	Detail
Name	Name of property like keywords, description, author etc.
content	Value of the property'.
scheme	How to interpret value of property (given in content attribute).
http-equiv	http response message headers details like to refresh the page or to set a cookie. Values include content-type, expires, refresh and set-cookie.

Attributes with examples are listed
Keywords – It is used by search engines for indexing a web page and is given by

```
<head>
<meta name="keywords" content="Vskills, HTML, Certification" />
</head>
```

Description – It is also used by search engines for indexing the web page as it provides some information about the web page and it is given by

```
<head>
<meta name="description" content="Certification in HTML." />
</head>
```

Revision date – It is to indicate last update time of the document and is given by

```
<head>
<meta name="revised" content="Vskills, 6/12/2011" />
</head>
```

Refreshing – It specifies duration after which web page will refresh itself and for a 15 seconds refresh is given by

```
<head>
<meta http-equiv="refresh" content="15" />
</head>
```

Page Redirection – It is for redirecting to another web page using meta tag and usually used with refresh but if no refresh value is given then redirection takes place immediately.

```
<head>
<meta http-equiv="refresh"
        content="15; url=http://www.vskills.in" />
</head>
```

Cookies - They store data on client computers to be used by web server for tracking a web page visitor.

```
<head>
<meta http-equiv="cookie" content="userid=new; course=HTML;
        expires=Wednesday, 18-Jun-12 23:59:59 GMT;" /> </head>
```

If no expiration date and time is given, cookie becomes a session cookie and is deleted when the user closes the browser.

Author – It is given as

```
<head>
<meta name="author" content="Vskills" />
</head>
```

1.13 HTML Attributes

As discussed earlier, attributes specify properties of an element by name and value pairs thus, attributes are made up of a name and a value.

The name is the property of element like <font> element has an attribute face for telling the typeface to use. The value is assigned in double quotation marks separated by equals sign to the

name and the value can be arial or courier for face name of font element. Different attributes are separated by space as in example

<font face="courierl" color="#ffffff">

Some HTML tags have attributes specific to them only but all have some common attributes.

1.14 XHTML First Line

XHTML language is derived from Extensible Markup Language (XML) which creates markup languages, and can have an optional XML declaration at start.

<?xml version="1.0" encoding="UTF-8" ?>

If included then this declaration should be at the start of document and nothing above it. Encoding attribute tells about the encoding used in web document to represent characters for display. Encoding is mentioned as different hardware and software uses different method to represent character (or encoding).

1.15 DTD (Document Type Declaration)

XHTML is stricter in rule implementation as compared to HTML but being in transition from HTML, XHTML gives three levels of compliance to a web page which are indicated by Doctype or Document Type Declaration and which are

Transitional XHTML 1.0 to allow use of outdated markups of HTML 4.1

Strict XHTML 1.0 is used with no outdated markup. It follows XHTML stricter syntax.

Frameset XHTML 1.0 to allow frames in web pages.

Usually transitional XHTML 1.0 is used and the doctype declaration is before <html> tag, and after optional XML Declaration. Different doctype declarations are

For transitional -
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

For strict -
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">

For frameset -
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">

1.16 HTML5 new Doctype and Charset

HTML5 defines an HTML syntax that is compatible with HTML4 and XHTML1 documents published on the Web. The HTML syntax of HTML5 requires a DOCTYPE to be specified to ensure that the browser renders the page in standards mode. The DOCTYPE declaration is <!DOCTYPE html> and is case-insensitive in the HTML syntax.

For the HTML syntax of HTML5, setting the character encoding is done by using a meta element with a charset attribute that specifies the encoding within the first 1024 bytes of the document. E.g. <meta charset="UTF-8"> could be used to specify the UTF-8 encoding. This replaces the need for <meta http-equiv="Content-Type" content="text/html; charset=UTF-8"> although that syntax is still allowed.

Following example of a HTML5 web page illustrates the usage of new doctype and charset

```
<!doctype html>
<html>
  <head>
   <meta charset="UTF-8">
   <title>Example document</title>
  </head>
  <body>
   <p>Example paragraph</p>
  </body>
</html>
```

1.17 Special Characters

Some characters other than alphanumeric characters have special meaning and some have no keyboard equivalent to enter like the angle brackets can be thought as markup by the browser. Hence, for them special characters called character entity or escape character are used. Character entities can be typed as either a numbered entity or a named entity. All character entities begin with an ampersand (&) and end with a semicolon (;). Example of them is

Character	Numbered Entity	Named Entity
"	"	"
&	&	&
©	©	©
®	®	®
<	<	<
>	>	>

1.18 Capitalization

Earlier version of HTML was case-insensitive so <html>, <HTML> or <HTml> were treated equally but HTML4, HTML 5 and XHTML are case-sensitive and requires all tags to be lowercase so only <html> is valid. Hence it is recommended to use all-lowercase tags.

1.19 Quotations

HTML and XHTML requires all values to be in quotation marks, as shown below

1.20 Nesting

It refers to the process of containing one HTML tag inside another like <body> and <head> tags are nested in <html> tag as shown

All tags should close as they have been opened within the tag it was opened.

1.21 Spacing and Breaks

Spacing is needed between tags and attributes and more rules need to be followed in applying spacing which is illustrated by the image

1.21 HTML5 Global attributes

Attributes which apply to all elements are called global attributes like accesskey, class, dir, id, lang, style, tabindex and title. There are also several new global attributes in HTML5 as

- ✓ The contenteditable attribute for an editable area.
- ✓ The contextmenu attribute for a context menu.
- ✓ The data-* collection of author-defined attributes.
- ✓ The draggable and dropzone attributes.
- ✓ The hidden attribute for an element which is not yet, relevant.
- ✓ The role and aria-* collection attributes.
- ✓ The spellcheck attribute.

HTML5 also makes all event handler attributes from HTML4, which take the form onevent-name, global attributes and adds several new event handler attributes for new events it defines. E.g. the play event which is used by the API for the media elements (video and audio).

Self Assessment Questions

Q.1 XHTML standard came before HTML 4.0
A. True
B. False
C. Can not say
D. None

Q.2 XHTML standard came before HTML 5.0
A. True
B. False
C. Can not say
D. None

Q.3 All elements should have a closing tag
A. True
B. False
C. Can not say
D. None

Answers : 1-B, 2-A,3-B