

(/)

HTML: A Gentle Introduction

HOWTOs (/tag/howtos) by Eric Kasten on July 1, 1995

HyperText Markup Language (HTML) is a simple language for representing document format styles and links to other documents or media types, such as images or sound recordings. HTML can be used to create documents which contain styles such as underlined or bold-faced text. It can also be used to mix text, images, and sounds into a single document, the individual elements of which may be located on geographically distant systems around the world.

HTML is designed to create documents for the World Wide Web and it helps determine what is displayed when you are browsing documents with your favorite WWW browser. You can use it to create your home, or welcome page or to create a research document, article, or book. This document can then be viewed locally or made accessible for viewing by other Web surfers through the use of a WWW server, such as NCSA's or CERN's httpd daemon. This article introduces you to HTML basics so you may get started creating your own HTML documents.

Document Tags

You can use your favorite text editor to create an HTML document. The document will be composed of text, which will be displayed directly to the user, and markup tags, which are used to modify the appearance of the text or to incorporate images or sounds as part

of the document. Tags are also used when referencing other documents or different locations within a document. Document references are called *hypertext links* or simply "links".

A tag for indicating the start of a particular format is represented as a tag name enclosed in a pair of angle brackets. To indicate the termination of a format, the tag name is prefixed with a *I*. For instance, **<I>Italics**</I> would display the word "Italics" in italic format. Let's examine a simple HTML document.

<HEAD> <TITLE>Sample Document</TITLE> </HEAD> <BODY> <H1>A Sample HTML Document</H1> Here is some Bold text, and here is some <I>Italic text</I>. </BODY>

This makes use of a few basic tags. The text between the **<HEAD>** and **</HEAD>** tags is the document header. The header contains a **<TITLE>** tag which indicates the start of the document title tag and is terminated by a **</TITLE>**. The title usually isn't displayed as part of the document text by most browsers, but instead is displayed in a special location. For instance, Mosaic displays the title in the title box at the top of the browser window.

After the header is the body of the document, which is contained between the **<BODY>** and **</BODY>** tags. Inside the body the **<H1>** represents the start of a first level document heading. There are six levels of headings. Each increase in level results in a decrease in the prominence with which a heading is displayed. For instance, you might want to use an **H1** heading for displaying the document title in the document text, and then use **H2** for subheadings.

This is probably a good time to mention that tags are case-insensitive. Thus **<TITLE>** and **<title>** are the same tag; however, I will continue to capitalize document tags for clarity.

Physical Format Style Tags

Physical format styles are used to indicate the specific physical appearance with which to display text. The following is a list of physical format tags:

Displays text in bold face.

<l>*text*</l>

Displays text in italics.

<U>text</U>

Displays *text* underlined.

<TT>*text*</TT>

Displays *text* using a typewriter font.

The problem with physical formats is that there is no guarantee that a particular browser will display the text as expected. A user may modify the fonts that a browser uses, or the browser may not even have the specified font style available. For instance, if a text mode browser is used to display a document, it is unlikely that italic text can be displayed at all. To avoid the ambiguity associated with the display of physical formats, you may use logical format tags. In fact, it is usually recommended that you use logical format tags, in preference to physical format tags, wherever you can.

Logical Format Style Tags

I've already introduced you to one logical format. The H1 through H6 heading styles are logical formats for headings. If physical heading styles existed in HTML, you would have to specify a particular font and font size for each heading (for example, 30-point Courier). Following is a list of some of the more common logical format styles:

<CITE>text</CITE>

Used when *text* is the title of a creative work such as a book.

<CODE>text</CODE>

Used when *text* is a piece of computer code.

text

Display *text* with emphasis.

<SAMP>*text*</SAMP>

Used when *text* is a piece of sample output.

text

RM

Display text with strong emphasis.

Of particular note, it is better stylistic practice to use the **EM** logical style in place of the *italics* physical style, and to use the **STRONG** logical style in place of the **bold** physical style.

Page Breaks, Rules, and Preformated Text

Carriage returns and white space are usually ignored in HTML. This is done because different browsers have different display capabilities. Where one browser might be able

Find authenticated court documents without watermarks at docketalarm.com.

necessary to have a forced break, such as between paragraphs. Special tags were created so that an author could force a break in a document. Following is a list of tags which can be used to force a break in a document:

<P>

Force a paragraph break.

Force a line break.

<HR>

Force a line break with a horizontal rule.

As a general note, it is best to use break commands, as well as other HTML directives in ways that keep your document device independent. The browser your readers use might format documents with different line lengths. If you use a **BR** directive at each line break found in a paper document you were converting into HTML, a browser might insert additional line breaks as well. This could result in a choppy-looking document, probably with alternating long and short lines.

Notice that these break tags don't have an associated closing tag. For instance the **
** doesn't have an associated **</BR>**.

If you do want to display something exactly as it is typed, HTML provides a tag for preserving the format of preformated text. The **<PRE>***text***</PRE>** directive instructs a browser to display *text* exactly as it is typed in the document. Once again, keep in mind that characteristics of some browsers may make it difficult for them to display the text in the exact format you expect.

Lists and Definitions

HTML provides for bulleted and numbered lists. Following is an example of HTML code which will generate a bulleted or unnumbered list of items:

 Apples Oranges Pears

The **** tag indicates the start of an unnumbered list. Each item in the list is preceded by an **** tag to indicate the start of a new item in the list. Note that the **** tags down have closing tags, and that when displayed, each list item will be concreted by a carriage



return.

To change this list to a numbered list, just change the **** and **** tags into **** and **** tags. A numbered list will display each item in the list preceded by a number instead of a bullet.

Another type of list is a definition or description list. Following is a snippet of HTML code demonstrating how to code a definition list:

<DL> <DT>Apple <DD>A red colored fruit <DT>Orange <DD>An orange colored fruit </DL>

A **<DL>** tag begins a definition list, and a **</DL>** tag ends the list. The **<DT>** indicates a term to be defined, while a **<DD>** indicates a term definition. When the browser formats the list, each term and definition will be on a line by itself: the term is usually left justified with definition indented directly beneath it.

An Example Document

Let's look at an example document which contains many of the text markups which I just explained. The HTML source for the example document is listed below, and the formatted document, as displayed by Mosaic, as I have configured it, is shown in Figure 1, below.

Figure 1. (/files/linuxjournal.com/linuxjournal/articles/010/1081/1081f1.jpg)

DOCKET



Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time** alerts and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

