
Hypertext Markup Language

INFO/CSE 100, Spring 2006
Fluency in Information Technology

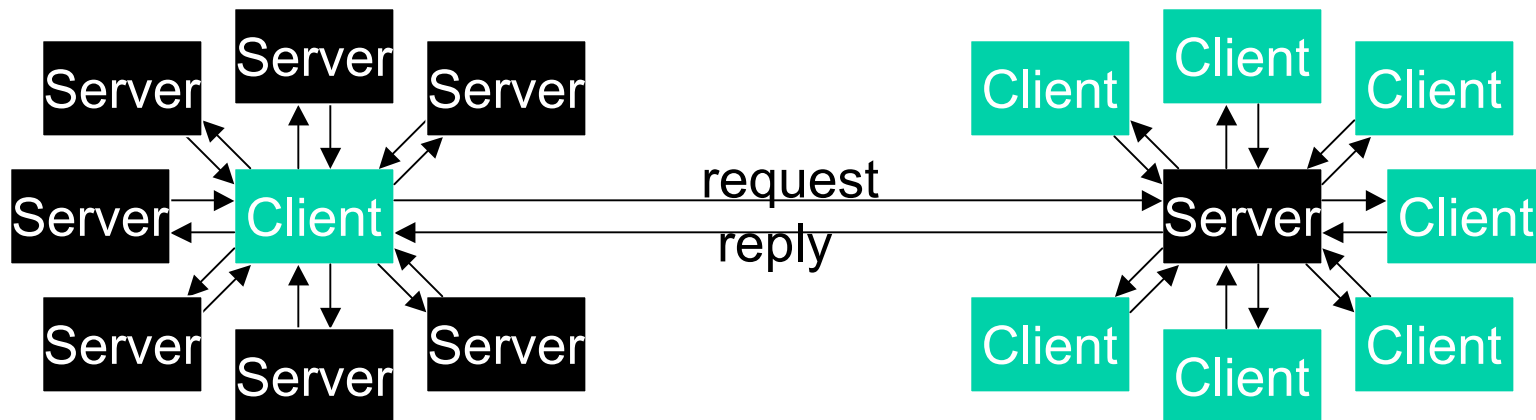
<http://www.cs.washington.edu/100>

Readings and References

- Reading
 - » *Fluency with Information Technology*
 - Chapter 4, Marking up with HTML
- References
 - » World Wide Web Consortium
 - "The World Wide Web Consortium (W3C) develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential."
 - <http://www.w3.org/>
 - <http://w3c.org/MarkUp/Activity>
 - <http://w3c.org/MarkUp/Guide/>

Client/Server Interaction

- For Web pages, the client requests a page the server returns it: there's no permanent connection, just a short conversation
 - » Details of the conversation are specified by HTTP



Simple HTTP Request

method request protocol

GET /pub/WWW/TheProject.html HTTP/1.1

Host: www.w3.org

host

A diagram illustrating the components of an HTTP request. The request line is "GET /pub/WWW/TheProject.html HTTP/1.1". Red arrows point from labels to parts of the request: "method" points to "GET", "request" points to the entire request line, "protocol" points to "HTTP/1.1", and "host" points to "www.w3.org" in the "Host: www.w3.org" header line.

A Typical Browser Request

GET /pub/WWW/TheProject.html HTTP/1.1

**Accept: image/gif, image/x-xbitmap, image/jpeg,
image/pjpeg, application/vnd.ms-powerpoint,
application/vnd.ms-excel, application/msword,
application/x-shockwave-flash, */***

Accept-Language: en-us

Accept-Encoding: gzip, deflate

**User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows
NT 5.0; APC)**

Host: www.w3.org

Connection: Keep-Alive

Server Response

HTTP/1.1 200 OK

Date: Monday, 23 May 2005 22:38:34 GMT

Server: Apache/1.3.27 (Unix) (Red-Hat/Linux)

Last-Modified: Wed, 08 Jan 2003 23:11:55 GMT

Etag: "3f80f-1b6-3e1cv03b"

Accept-Ranges: bytes

Content-Length: 438

Connection: close

Content-Type: text/html

<html>

<head><title>A Sample Page</title></head>

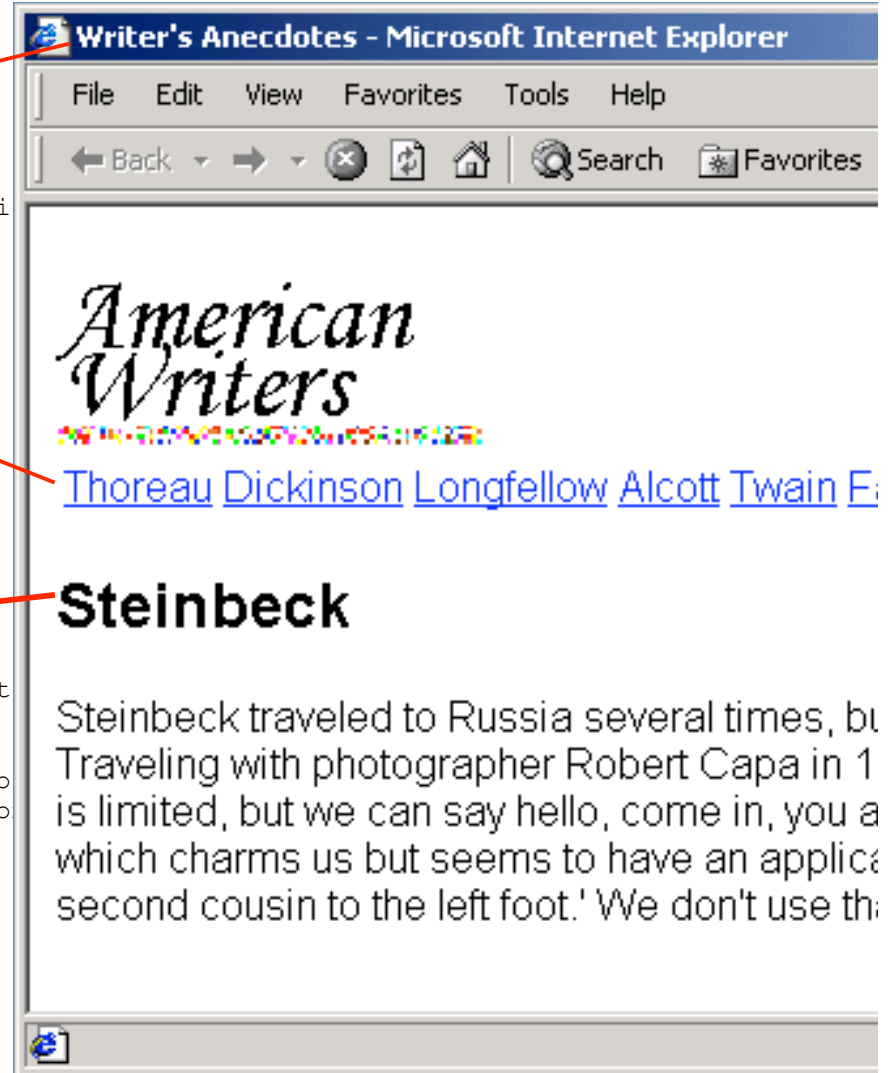
<body>

...



Web Pages in HTML

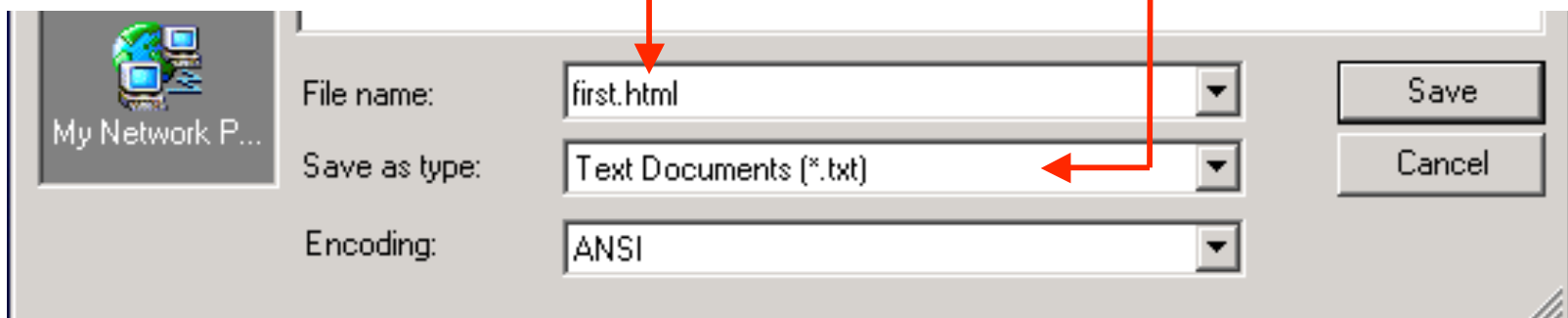
```
<html>
<head><title>Writer's Anecdotes</title></head>
<body bgcolor="white" text="black"><font face="Helvetica"
  
  <table>
    <tr>
      <td><a href="hdt.html">Thoreau</a></td>
      <td><a href="ed.html">Dickinson</a></td>
      <td><a href="hwl.html">Longfellow</a></td>
      <td><a href="lma.html">Alcott</a></td>
      <td><a href="sc.html">Twain</a></td>
      <td><a href="wf.html">Faulkner</a></td>
      <td><a href="rf.html">Frost</a></td>
      <td><a href="eh.html">Hemingway</a></td>
      <td><a href="js.html">Steinbeck</a></td>
    </tr>
  </table>
  <h2>Steinbeck</h2>
  <p>Steinbeck traveled to Russia several times, but
  never mastered the language. Traveling with
  photographer Robert Capa in 1947 he wrote, "...I
  admit our Russian is limited, but we can say hello
  come in, you are beautiful, oh no you don't, and o
  which charms us but seems to have an applica
  rarely needed, 'The thumb is second cousin to the
  left foot.' We don't use that one much."</p>
</body>
</html>
```



HTML Must Be Text

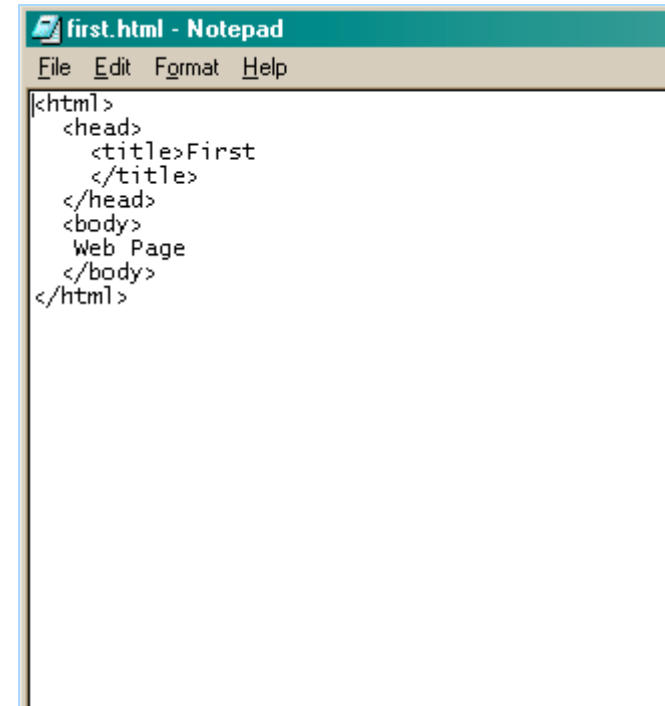
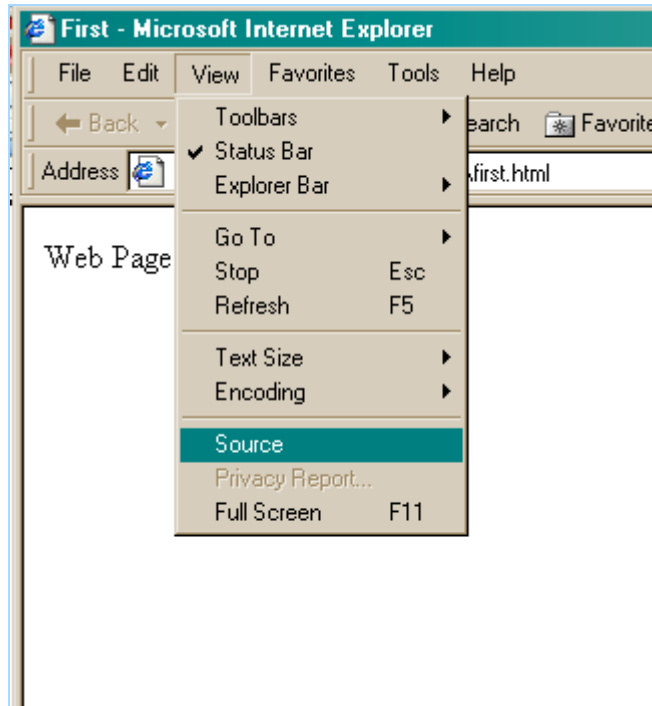
- Word processors insert formatting tags, confusing browsers
 - Create source in NotePad, jEdit, etc.
 - Save in **Text** or **txt** format
 - Save with file extension **.html**

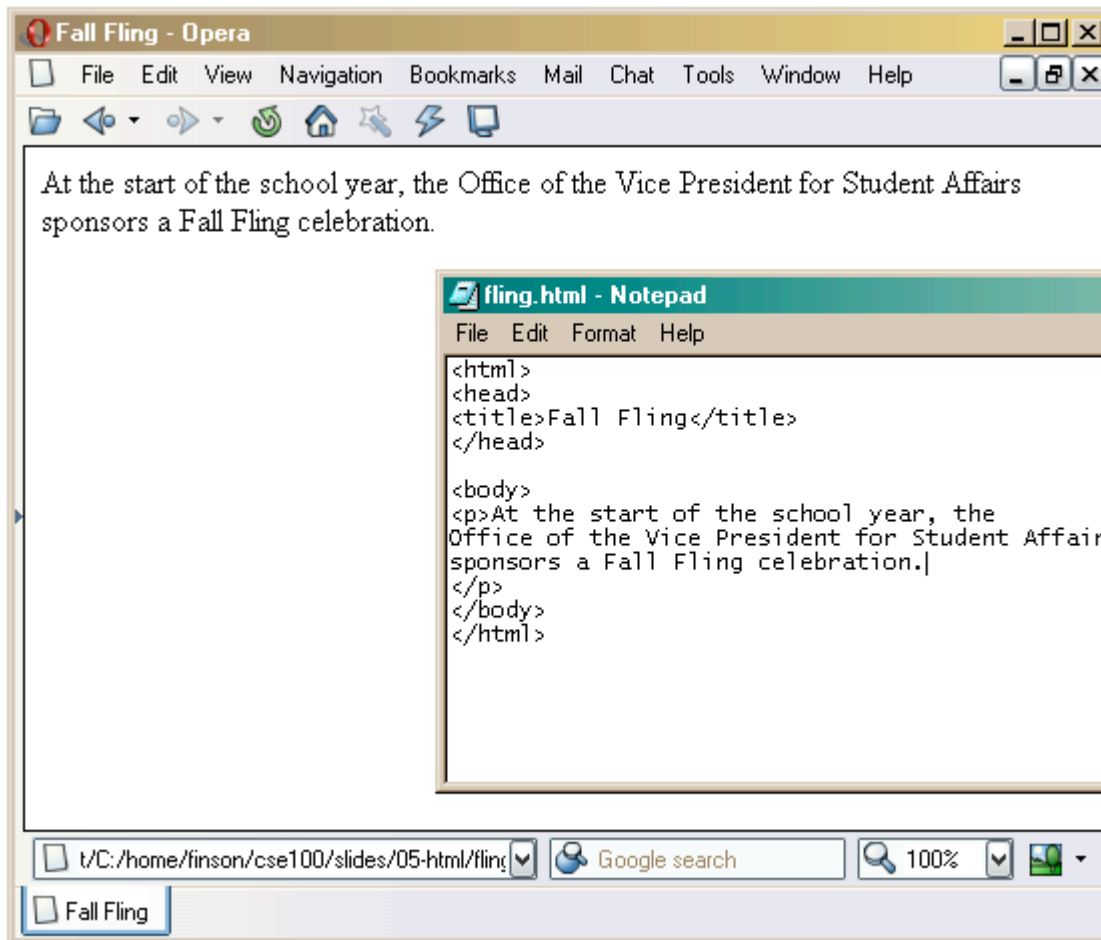
👉 Avoid
Confusion



The Source

The HTML code producing a page is the *source* which can always be *viewed*, either in the browser or in an editor





HTML Structure

- All HTML files use the same basic structure:

```
▶ <html>
  ▶ <head>
    ▶ <title>Name of Page Goes Here
    ▶ </title>
    Heading content goes here
  ▶ </head>
  ▶ <body>
    Body content goes here
  ▶ </body>
▶ </html>
```

HTML Page Structure

```
<!DOCTYPE definition... >
<html>
  <head>
    <title>Title text</title>
  </head>
  <body>
    Body text goes here...
  </body>
</html>
```

HTML Tags

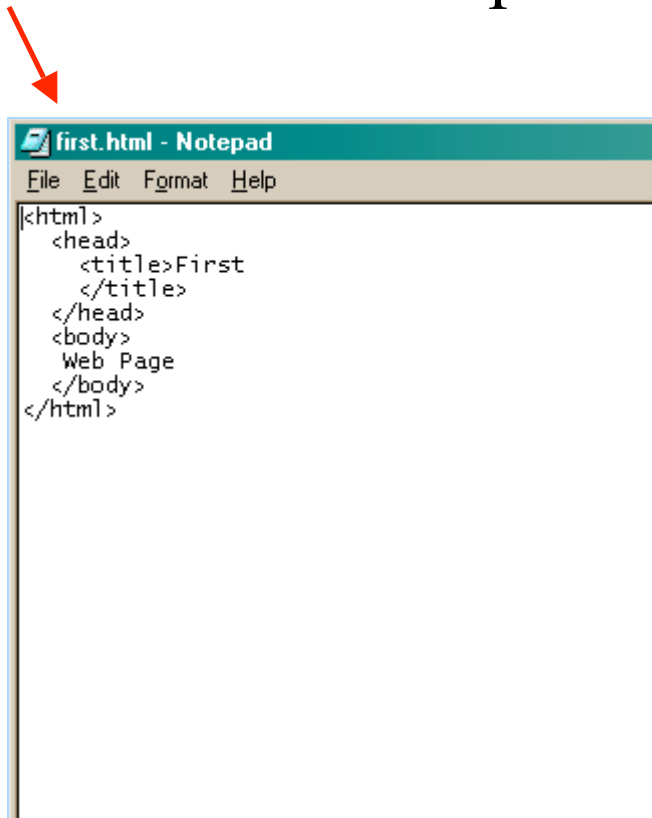
- Anatomy of an HTML tag
 - » `<element attribute="value"></element>`
 - » ``
- Types of tags
 - » Normally has an open AND a close element
 - » Open and close at the same time
 - » Some tags do not close at all (ex. Comment tag)
- Attributes
 - » Name-value pairs, values in quotes
 - » Some are required, some optional

More HTML Tags

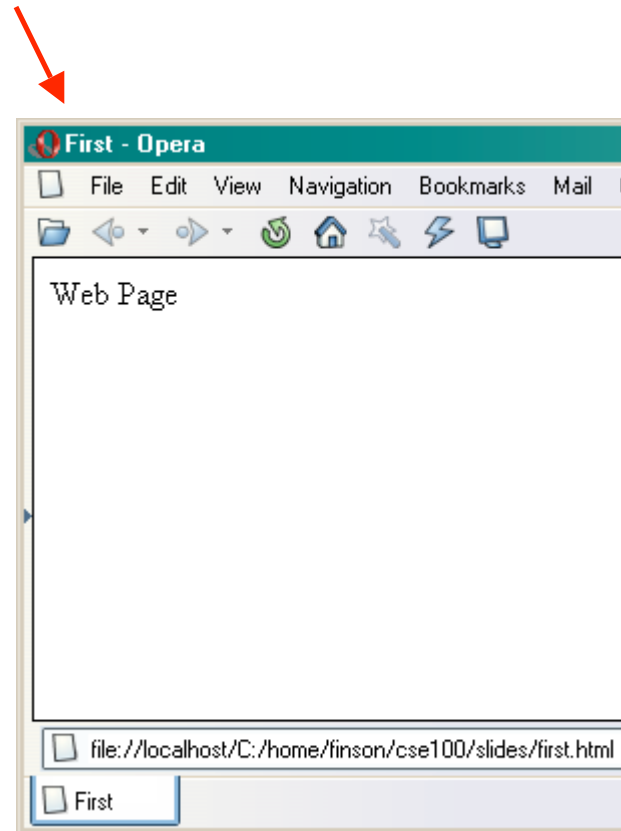
- Styles Tags
 - » `` bold, `<i>` italic, `<big>` big, `<small>` small
- Spacing Tags
 - » `<p>` paragraph, `
` line break, `<hr />` horizontal rule
- Heading tags
 - » `<h1>`, `<h2>`, `<h3>` header format
- List tags
 - » `` ordered list, `` list item
 - » `` unordered (bulleted) list, `` list item
- References
 - » `` anchor reference
 - » `` image reference

First HTML Web Page

This HTML code produces this result



```
<html>  
  <head>  
    <title>First  
  </title>  
  </head>  
  <body>  
    Web Page  
  </body>  
</html>
```



Add An Image

- Images are encoded many ways:
 - » GIF -- Graphics Interchange Format -- is for diagrams and simple drawings
 - » PNG -- Portable Network Graphics - GIF without patent problems
 - » JPEG -- Joint Photographic Experts Group -- is for high resolution photos, complex art
- Image tags for placing images

» ``

tag attribute name attribute value

File Locations

- The path must say how to reach the file
 - » When the file is in the directory as the web page, just give the file name, **ski.jpg**
 - » If the file is in a subdirectory, say how to navigate to it, **pix/ski.jpg**
 - » If the file is in a superdirectory, move up using dot-dot notation, **../ski.jpg**

The most common reason that an image is not displayed is the path is wrong ... check!

Anchors

- The link starts with an **a** tag...

```
<a href=" ... ">
```

- The value of the **href** attribute is a URL.

```
<a href="http://www.cs.washington.edu">
```

- The link encloses some text, and ends with...

```
<a href=" ... ">CSE home page</a>
```

Types of URLs


- <http://www.site.com/dir/file.html>
- <http://www.site.com/dir/>
- <ftp://ftp.site.come/pub/>
- <news:soc.culture.religion>
- <mailto:suzka@u.washington.edu>
- <file:///hardisk/path/file.html>

Fall Fling Foto - Opera

File Edit View Navigation Bookmarks Mail Chat Tools Window Help

I happened to have my camera with me on the day of the [Fall Fling](http://depts.washington.edu/sao/fallfling.php) this year and took this photo of a cool spinner.

Doug Johnson



fling.html - Notepad

File Edit Format Help

```
<html>
<head>
<title>Fall Fling Foto</title>
</head>

<body>
<p>
I happened to have my camera with me
on the day of the
<a href="http://depts.washington.edu/sao/fallfling.php">Fall Fling</a>
this year and took this photo of a cool spinner.
</p>
<p>Doug Johnson</p>
</body>
</html>
```

file://localhost/C:/home/finson/cse100/slides/fling.html

Fall Fling Foto


Simple HTML

- What we've seen here is *very simple* HTML
- HTML is changing
 - » Each document should start with a “**DOCTYPE**” comment telling which version of HTML it follows
 - <http://www.w3.org/QA/2002/04/valid-dtd-list.html>
 - We used HTML 4.01 Transitional
 - » There is a validator service that will check your page
 - <http://validator.w3.org/>


Fall Fling Foto - Opera

File Edit View Navigation Bookmarks Mail Chat Tools Window Help

I happened to have my camera with me on the day of the [Fall Fling](#) this year and took this photo of a cool spinner.



Doug Johnson



```
flingfoto.html - Notepad
File Edit Format Help
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
<title>Fall Fling Foto</title>
</head>
<body>
<p>
I happened to have my camera with me
on the day of the
<a href="http://depts.washington.edu/sao/fallfling.php">Fall Fling</a>
this year and took this photo of a cool spinner.
</p>
<p>Doug Johnson</p>
<p>
<a href="http://validator.w3.org/check?uri=referer"></a>
</p>
</body>
</html>
```

i/education/courses/cse100/04au/slides/05-html/flingfoto.html Google search 100%

Fall Fling Foto

Wrap-up

- HTML forms the basic structure for web pages
 - » Modern trends are to separate
 - the structural description of the data
 - the presentation of the data to the user
 - » process the information according to the content
 - » display the information as appropriate to the user
- Programming languages make information management and display dynamic
 - » languages come in many flavors but they share many basic concepts and much of their basic syntax

XHTML vs. HTML 4

- <http://www.w3.org/TR/xhtml1/#diffs>

Summary

- Web pages are written in HTML
 - The files must be text
 - The file extension must be .html
- Tags enclose content like parentheses
 - Control look with attributes on tags
- Use a change-and-test process
- Images have various formats and you must specify the explicit path to the picture