

Hypertext Markup Language

HTML

Chapter 2

Supporting Material for “Using Google App Engine” - O’Reilly and Associates
www.appenginelearn.com

open.michigan

Unless otherwise noted, the content of this course material is licensed under a Creative Commons Attribution 3.0 License.
<http://creativecommons.org/licenses/by/3.0/>.

Copyright 2009, Charles Severance



HTML

JavaScript

AJAX

CSS

HTTP

Response

POST

Request

GET

Python

Templates

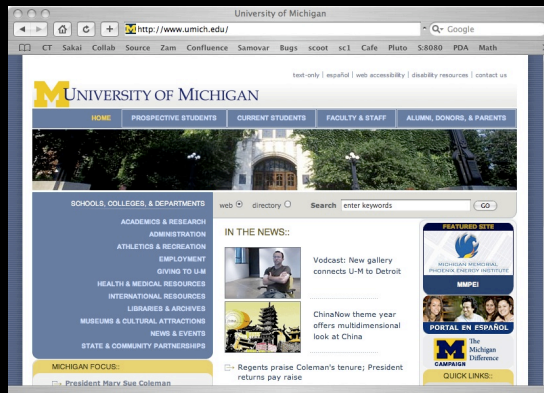
Data Store

memcache

The big picture...



```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
<head>
<title>University of Michigan</title>
....
```



A web server produces HTML which is handed to a browser which needs to lay it out in a blink of an eye and have it pixel perfect as good as a print brochure.

HTML

- A way of marking up text to indicate that some text is different than other text
- We “tag” portions of the text to communicate meaning

`<h1>Google App Engine:About</h1>`

`<p>Welcome to the site dedicated to learning the Google Application Engine. We hope you find www.appenginelearn.com useful.</p>`

```
<h1>Google App Engine:About</h1>
<p>Welcome to the site dedicated to
learning the Google Application
Engine.We hope you find
www.appenginelearn.com useful.</p>
```



Whitespace and
line wrapping.

My Workspace | SI 539 | SI 539 W08 | SI 539 F08 | SI502 | SI502Staff | - more sites -

powered by Google App Engine

Email Archive

Options Permissions

Home
Announcements
Resources
Podcasts
Polls
Chat Room
Email Archive
Calendar
Instructor IM
Software
Site Info

You are authorized to send email from: csev@umich.edu
Email sent to the following addresses will be archived and sent to participants:

si539@ctools.umich.edu
0ae836b9-7fc9-4060-006f-27b2066ac545@ctools.umich.edu

Viewing 1 - 20 of 446 items
show 20 items...

From	Subject	Date Received
SI 539 F08 Anthony Whyte <arwhyte@umich.edu>	Google Chrome comic book	Sep 2, 2008 10:58 AM EDT

<a href="mailto:si539@ctools.umich.edu"
title="Send mail to si539@ctools.umich.edu">

si539@ctools.umich.edu

Learning from
“View Source”

Evolution of HTML

The Web is a Young Technology

- Invented in early 1990's
- Popular in 1994
- Robert Cailliau -
coFounder of the World-
Wide-Web

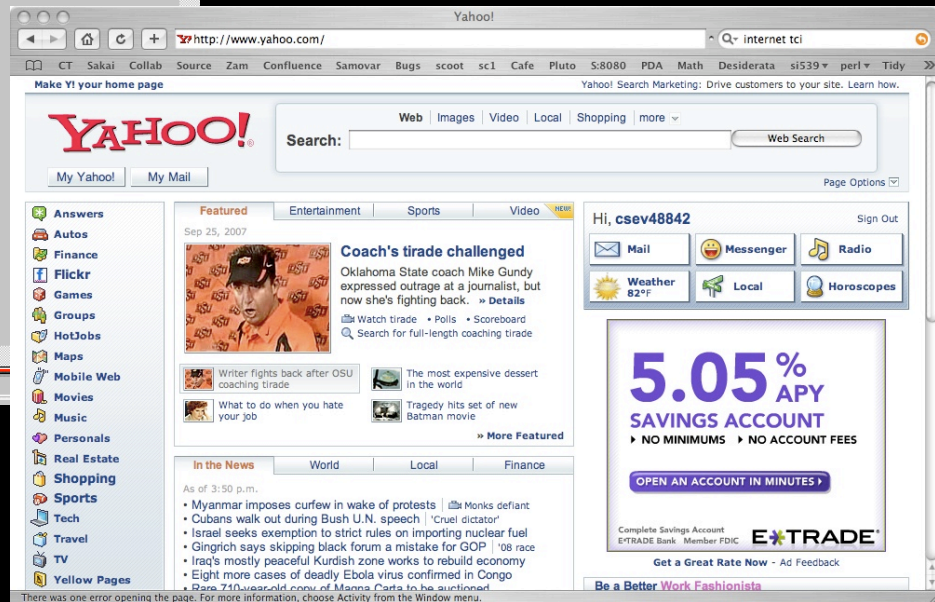


<http://www.dr-chuck.com/media.php?id=70>



1995

2007



HTML has evolved a *lot* over the years - as computers and networks have gotten faster.

An Iterative Process

- Designers want to do something
- They figure out how to do it with current generation browsers
- We figure out requirements - but HTML gets uglier and quirker - then the HTML starts to break
- We develop new standards that handle new requirements in an elegant manner
- New Browsers arrive in the marketplace with the new standard

History of HTML / CSS

- HTML 1.0 - 1993 - The Good Old Days - life was simple
- HTML 2.0 - 1995 - Some interesting layout features - abused
- CSS 1 - 1996
- HTML 3.2 - 1997
- HTML 4.0 - 1997 - Layout moving toward CSS
- CSS Level 2 - 1998
- HTML 4.01 - 1999 - What we use today

HTML has evolved a *lot* over the years - as computers and networks have gotten faster.

The Good Old Days

```
<h1>Hello World</h1>
```

Hi there.

```
<p>
```

A Paragraph

```
<ul>
```

```
<li>List one
```

```
<li>List 2
```

```
</ul>
```

In the good old days you wrote HTML and browsers displayed it - since we wrote it by hand - and modems were slow - it was never too long and never too complex. The browser was never the rate limiting factor.

Writing HTML was like using a simple, weak word processor. The tags acted as formatting commands to the browser.

The Ugly Middle Ages

- Web Designers designed to browser capabilities - down to particular minor versions of browsers.
- Extensive testing was needed on lots of browsers
- Designers used tables, nested tables, and chopped up graphics to gain control of the look and feel of web pages to produce a “print-like” experience.
- HTML was UGLY, Hard to develop, and brittle - what looked superb on one browser - often was broken on another browser - even a later release of the same browser.

The Modern Era

- HTML is clean and simple
- There is no presentation in HTML - font, color, spacing, etc etc
- No use of tables except for tabular data
- CSS controls all layout, and look and feel
- Still a bit challenging - but converging

1990-1994	HTML was simple and pages looked pretty ugly.
1995-1999	HTML became more complex and each browser was different.
2000-2005	Browsers slowly supported CSS to varying levels. HTML was still ugly to support multiple browsers.
2005-2008	New browsers supported CSS. Old browsers were still pretty pervasive but diminishing.
2009+	The last “pre-CSS” browser (IE5) is < 0.1%

http://www.w3schools.com/browsers/browsers_stats.asp

What does this mean for us?

- Don't bother with the intermediate steps :)
- Either keep it simple - or do it well - simple does work
- If you want a professional site use all of the best practices
 - Presentation in CSS + Semantic markup in HTML
 - <http://validator.w3.org/>
 - <http://jigsaw.w3.org/css-validator/>

```
<h1>Google App Engine:About</h1>
<p>Welcome to the site dedicated to
learning the Google Application
Engine.We hope you find
www.appenginelearn.com useful.</p>
```



Whitespace and
line wrapping.

Tags have a **beginning** and **end**...

```
<h1>Google App Engine:About</h1>  
<p>Welcome to the site dedicated to  
learning the Google Application  
Engine.We hope you find  
www.appenginelearn.com useful.</p>
```

HTML Tag Basics

Start tag

End tag

`<h1>Hello World</h1>`

Tags “mark up” the HTML document. The tags are read and interpreted by the browser - but now shown.

Attribute

``

A self closing tag does not need a corresponding end tag.

Self-closing tag

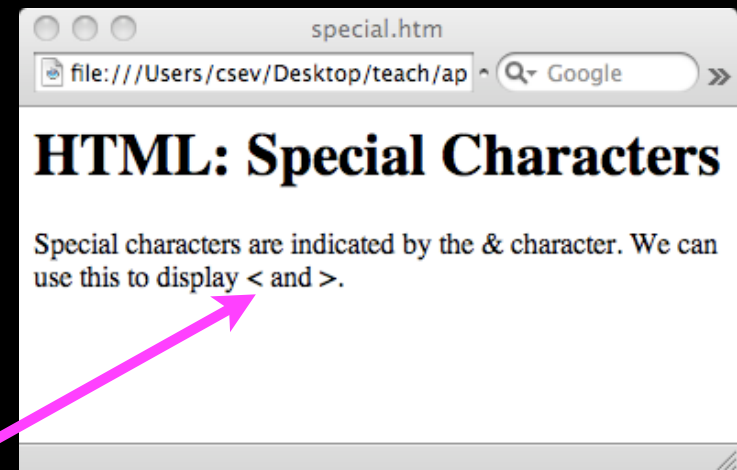
What about `<` 's in HTML?

`<h1>`HTML: Special Characters`</h1>`

`<p>`

Special characters are indicated by the `&` character. We can use this to display `<` and `>`;

`</p>`



Entity	Symbol
<code>&gt;</code>	<code>></code>
<code>&lt;</code>	<code><</code>
<code>&amp;</code>	<code>&</code>
<code>&pound;</code>	£
<code>&copy;</code>	©
<code>&trade;</code>	™

HTML Links

- One of the key things about HTML is making a set of pages and making “hypertext” links amongst those pages
- Links are what make the “web” a “web” - it is a web of interlinked documents.
- The interlinked nature of the web leads to the “knowledge” the search engines like Google appear to have

```
<h1>The First Page</h1>
```

```
<p>
```

If you like, you can switch to the

```
<a href="http://www.dr-chuck.com/page2.htm">
```

Second Page.

```
</p>
```

A link is a “hot spot” on the page. It can be text or an image. Often it is visually marked to make it easier to “notice” so as to engage users to click!

“a” is short for “anchor” and “href” is short for “hypertext reference”



```
<h1>The Second Page</h1>
```

```
<p>
```

If you like, you can switch back to the

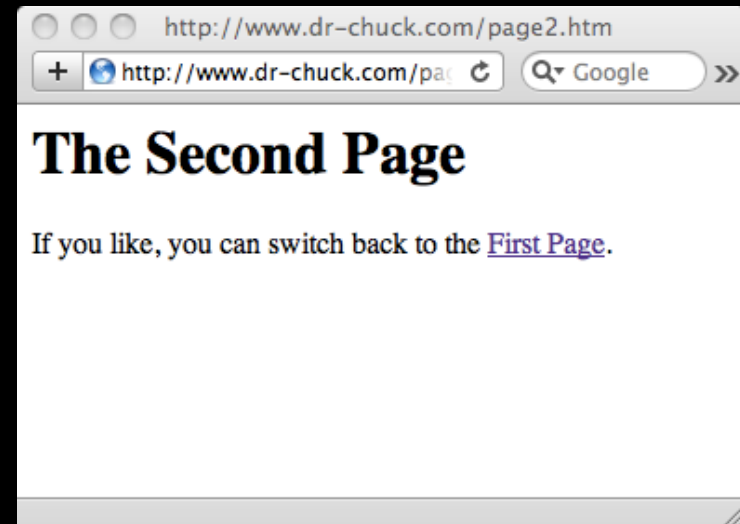
```
<a href="page1.htm">
```

First Page.

```
</p>
```

A link is a “hot spot” on the page. It can be text or an image. Often it is visually marked to make it easier to “notice” so as to engage users to click!

“a” is short for “**anchor**” and “href” is short for “**hypertext reference**”



Absolute Reference

```
<a href="http://www.dr-chuck.com/page2.htm">Second Page</a>
```

Start
tag

Where to
go if clicked.

Clickable
Text

End
tag

```
<a href="page1.htm">First Page</a>
```

Relative Reference

Absolute .vs. Relative

```
<a href="http://www.dr-chuck.com/page2.htm">Second Page</a>
```

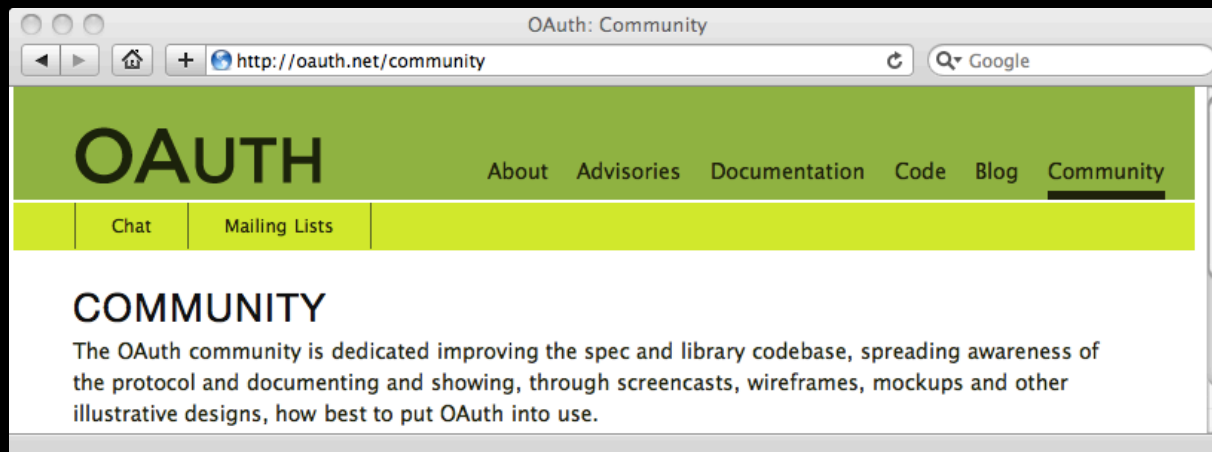
Hypertext references can be a full URL - and refer to some other page on anywhere on the Internet.

```
<a href="page1.htm">First Page</a>
```

Or the reference can be a file name that is assumed to be in the same folder as the current document (relative reference).

Navigation With Anchor Tags

- Sometimes we want to make a menu that provides our users consistent navigation across page.



A List ...

```
<h1><a href="index.htm">AppEngineLearn</a></h1>
```

```
<ul>
```

```
  <li><a href="sites.htm">Sites</a></li>
```

```
  <li><a href="topics.htm" >Topics</a></li>
```

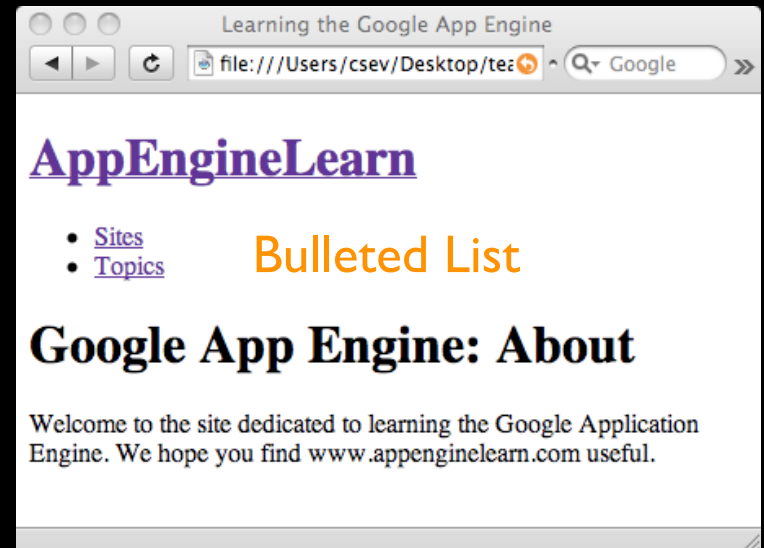
```
</ul>
```

```
<h1>Google App Engine:About</h1>
```

```
<p>
```

```
Welcome to the site dedicated to  
learning the Google Application Engine.  
We hope you find www.appenginelearn.com  
useful.
```

```
</p>
```



A List of Links...

```
<h1><a href="index.htm">AppEngineLearn</a></h1>
```

```
<ul>
```

```
  <li><a href="sites.htm">Sites</a></li>
```

```
  <li><a href="topics.htm" >Topics</a></li>
```

```
</ul>
```

```
<h1>Google App Engine:About</h1>
```

```
<p>
```

Welcome to the site dedicated to learning the Google Application Engine. We hope you find www.appenginelearn.com useful.

```
</p>
```



Multiple Files

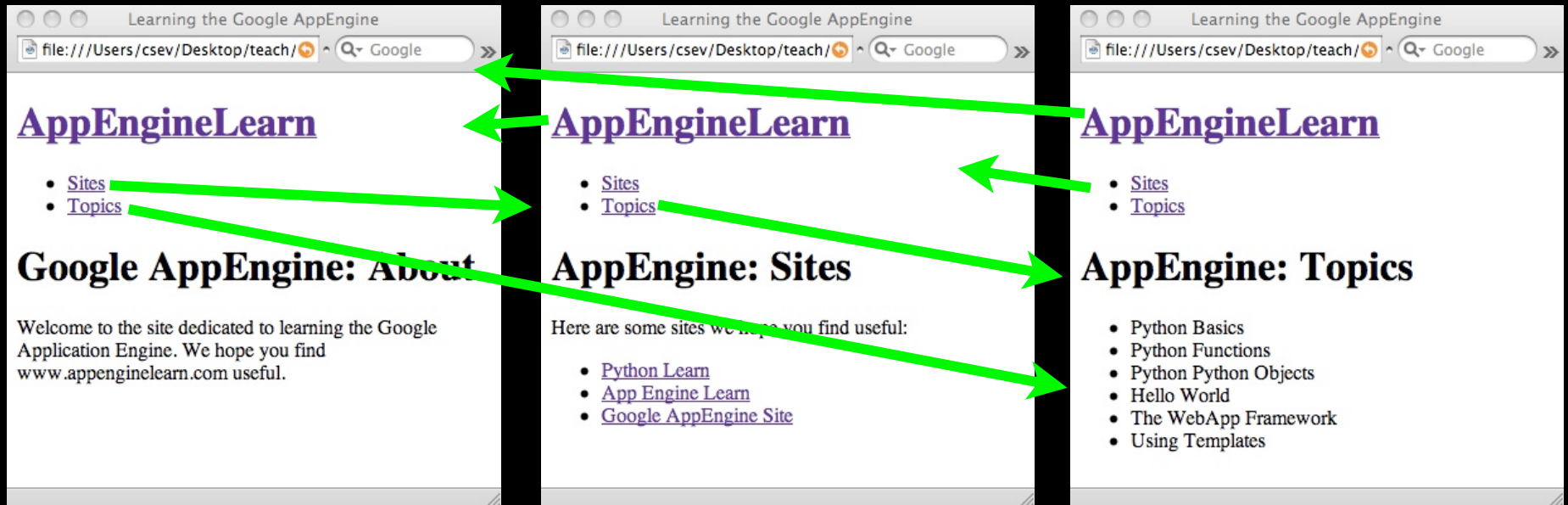
- We can put multiple files in the same directory and then use them in relative links.

```
csev$ ls -l
-rw-r--r-- | csev staff 618 Dec 18 22:56 index.htm
-rw-r--r-- | csev staff 883 Dec 18 22:57 sites.htm
-rw-r--r-- | csev staff 679 Dec 18 22:57 topics.htm
csev$
```

```
<h1><a href="index.htm">AppEngineLearn</a></h1>
<ul>
  <li><a href="sites.htm">Sites</a></li>
  <li><a href="topics.htm" >Topics</a></li>
</ul>
```

```
csev$ ls -l
-rw-r--r-- 1 csev staff 618 Dec 18 22:56 index.htm
-rw-r--r-- 1 csev staff 883 Dec 18 22:57 sites.htm
-rw-r--r-- 1 csev staff 679 Dec 18 22:57 topics.htm
csev$
```





```
csev$ ls -l
-rw-r--r-- | csev staff 618 Dec 18 22:56 index.htm
-rw-r--r-- | csev staff 883 Dec 18 22:57 sites.htm
-rw-r--r-- | csev staff 679 Dec 18 22:57 topics.htm
csev$
```

We create the appearance of a fixed “menu” by including the navigation at the same place in each page.

Special File Names

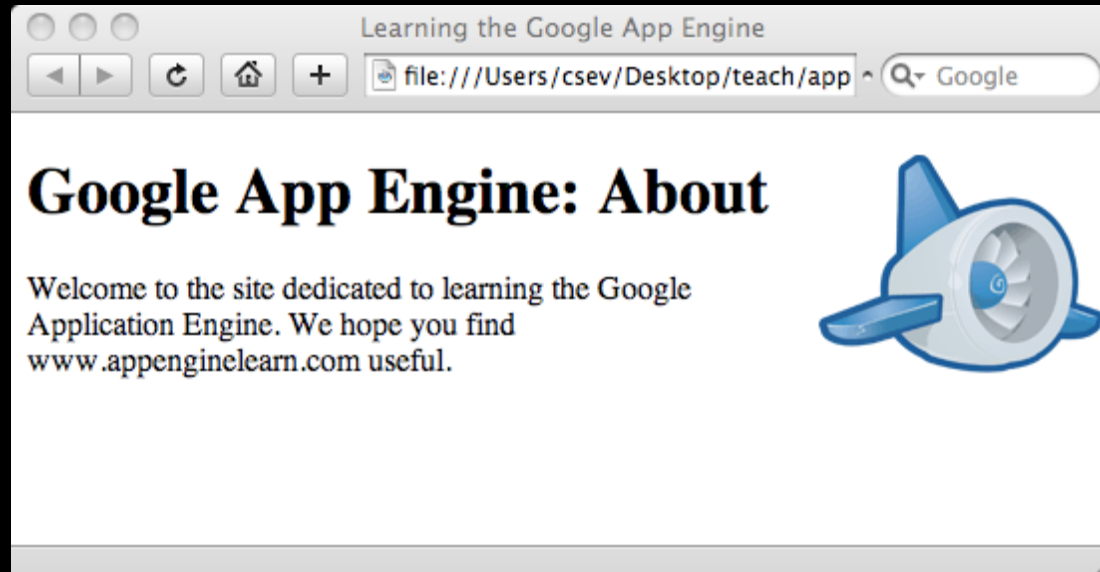
- When a URL points to a directory in your web server, it looks for a file with a special name:
 - `index.html`, `index.htm`, `index.php`, `default.htm`, etc..
- While there is a convention, the “default file” is configurable - so nothing is “sure”
- Usually `index.htm` or `index.html` is a safe bet
- This only works when viewing through a web server - when viewing from disk, you must view the file.

```
<body>
  <!-- Make sure to style the h1 -->
  <h1>App Engine: Topics</h1>
  <ul>
    <li>Python Basics</li>
    <li>Python Functions</li>
    <li>Python Python Objects</li>
    <!-- Leave these two out for a while
    <li>Hello World</li>
    <li>The WebApp Framework</li>
    -->
    <li>Using Templates</li>
  </ul>
</body>
```



HTML Comments

Images



Images

Which image to display

Begin Tag

```

```

Optional - makes
display quicker.
Will resize to fit.

End Tag

Put the image on the right
and wrap text around it.

Show this when hovering,
images are off, or for
screen readers.

All information is communicated through the attributes of the img tag.

Images

```
<h1>
```

```
  
```

```
  Google App Engine:About</h1>
```

```
<p>
```

```
Welcome to the site dedicated to  
learning the Google Application Engine.  
We hope you find www.appenginelearn.com useful.
```

```
</p>
```

In this case, the file `appengine.jpg` needs to be in the same directory as the file `index.html`.

```
<h1>
```

```

```

```
Google App Engine:About</h1>
```

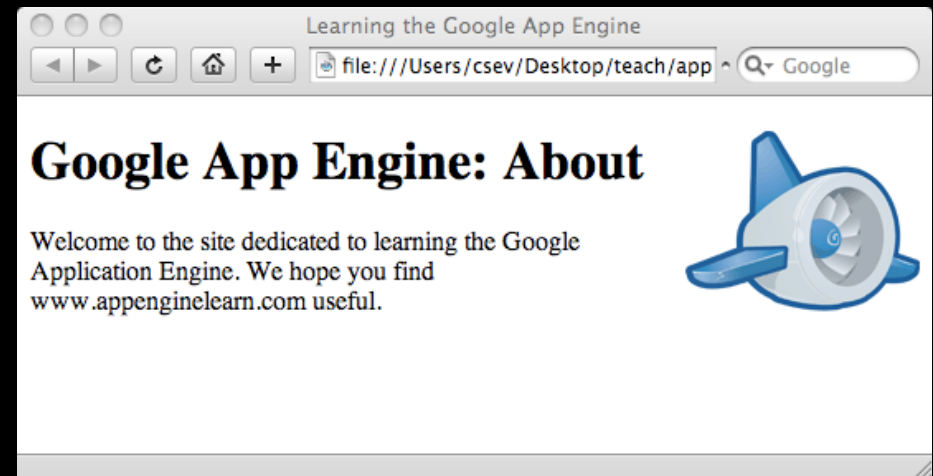
```
<p>
```

```
Welcome to the site dedicated to  
learning the Google Application Engine.
```

```
We hope you find
```

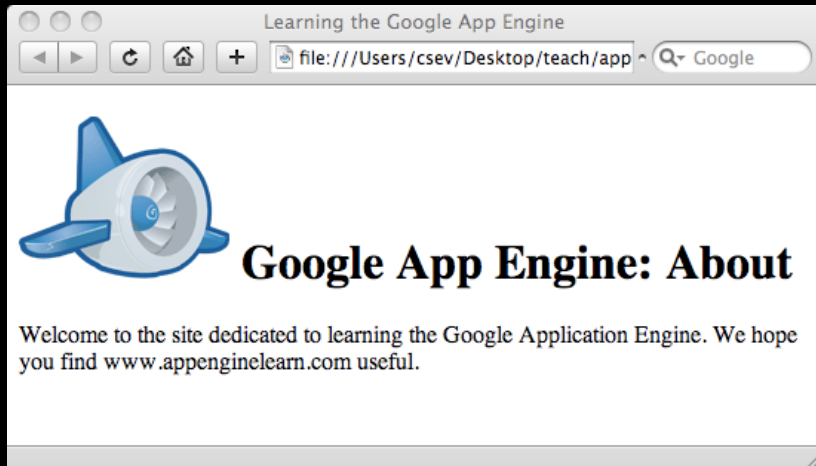
```
www.appenginelearn.com useful.
```

```
</p>
```



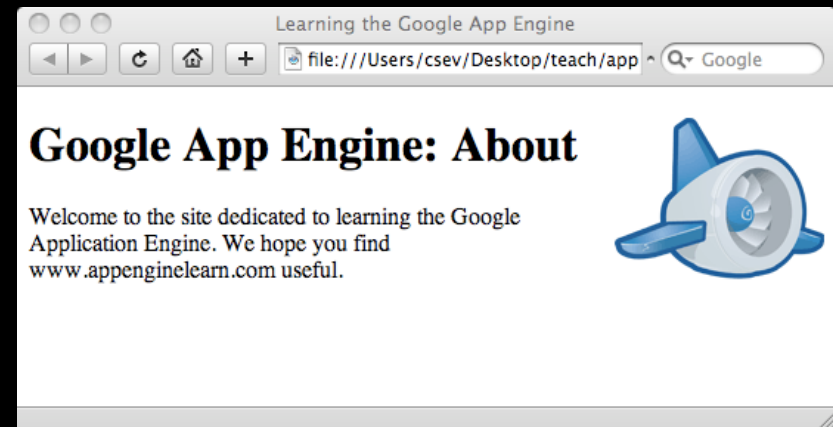
```
<h1>

Google App Engine:About</h1>
<p>
Welcome to the site dedicated to
learning the Google Application Engine.
We hope you find www.appenginelearn.com useful.
</p>
```



```
<h1>

Google App Engine:About</h1>
<p>
Welcome to the site dedicated to
learning the Google Application Engine.
We hope you find www.appenginelearn.com useful.
</p>
```



HTML Document Structure

Well-Formed HTML Documents

- Browsers need to know what contract you are agreeing to (doctype)
- We need a section for meta-data about the HTML document (head)
- And then the displayable content of the HTML document (body)

Outline of an HTML Document

```
<!DOCTYPE >
```

```
<html>
```

```
<head>
```

```
  meta-data....
```

```
</head>
```

```
<body>
```

```
  Page content....
```

```
</body>
```

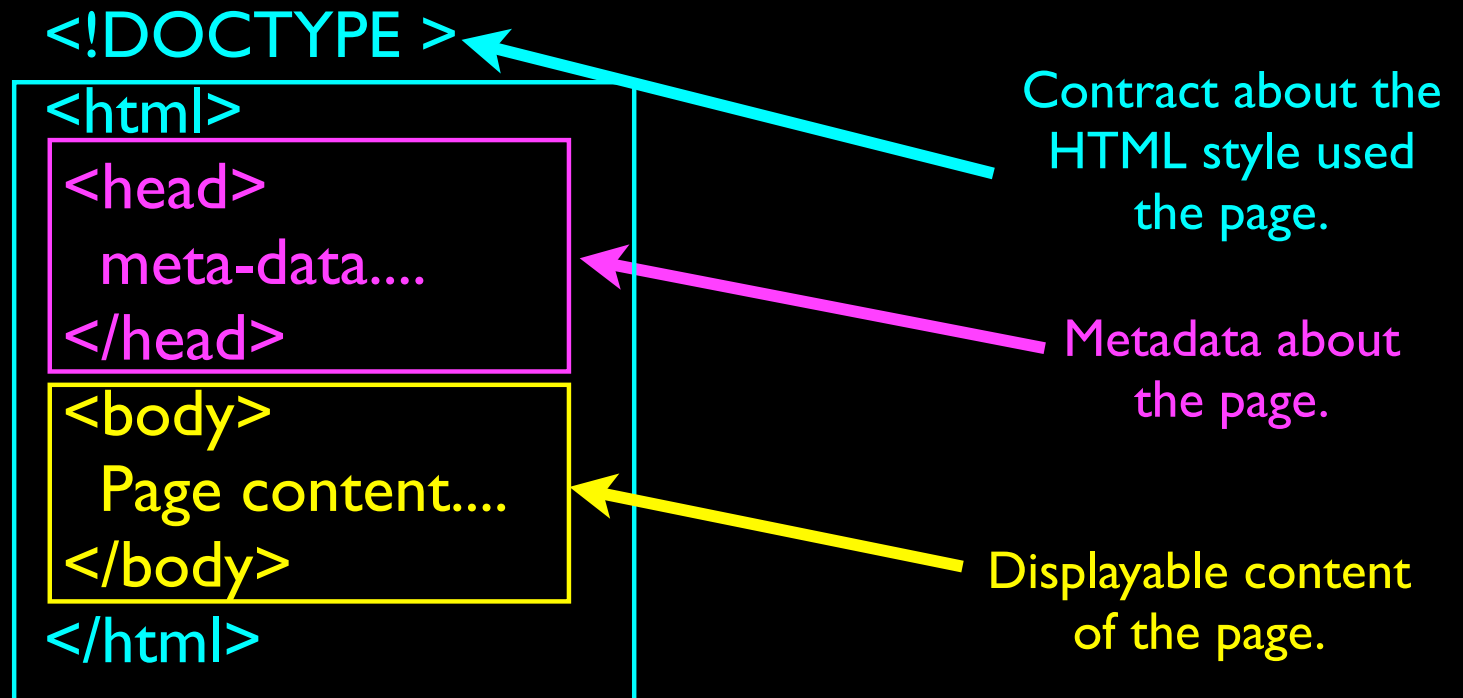
```
</html>
```

Contract about the
HTML style used
the page.

Metadata about
the page.

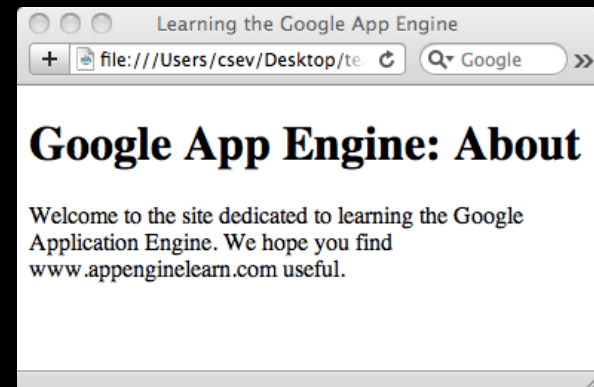
Displayable content
of the page.

Outline of an HTML Document



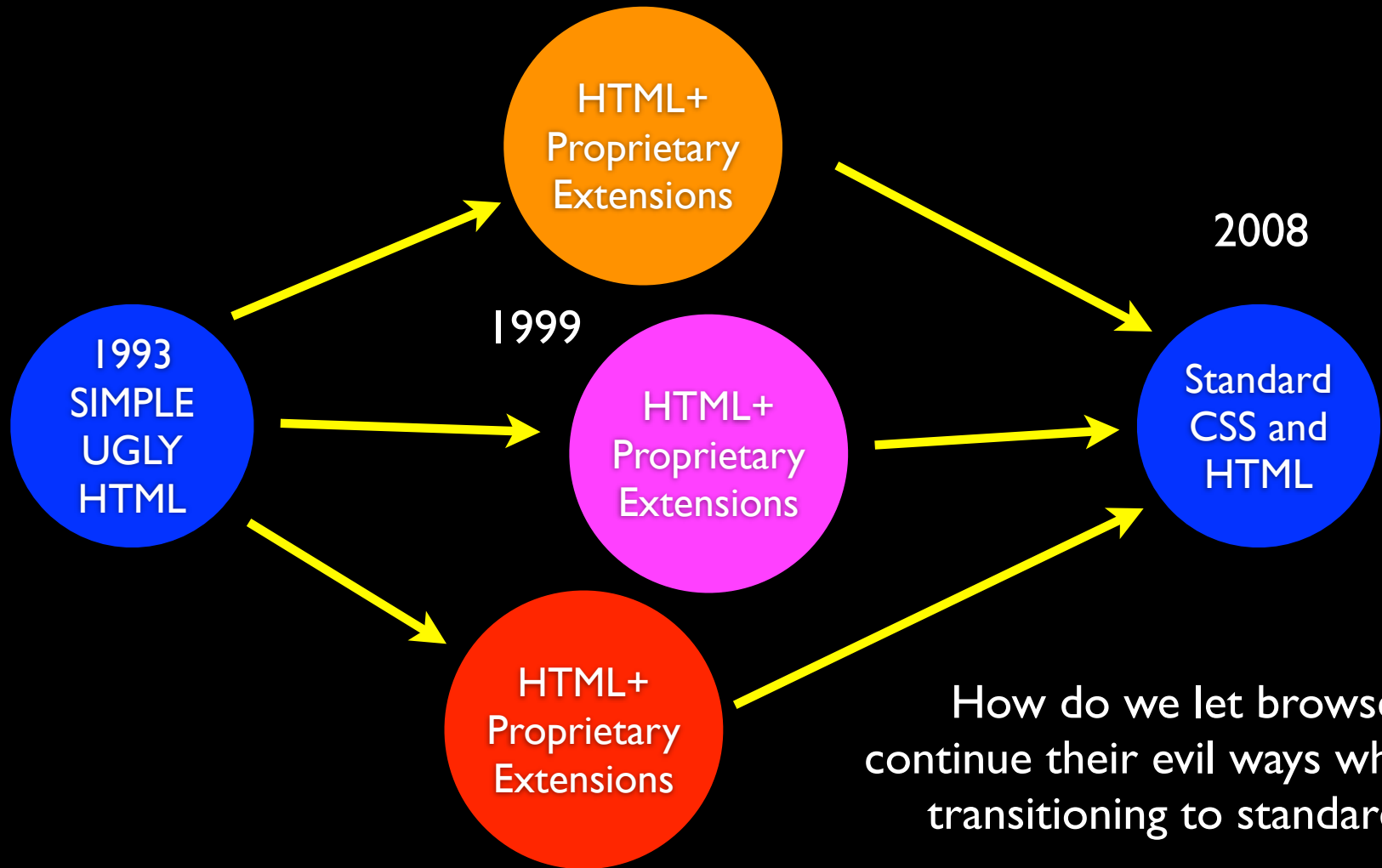
A Simple but Modern Page

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<html>
  <head>
    <title>Learning the Google App Engine</title>
  </head>
  <body>
    <h1>Google App Engine:About</h1>
    <p>
      Welcome to the site dedicated to
      learning the Google Application Engine.
      We hope you find www.appenginelearn.com useful.
    </p>
  </body>
</html>
```



Browsers: Lost in Transition

- Between 1995 and 2003 browsers were different - often on purpose to attempt to make their market share “sticky”
- At some point we knew there needed to be a future where all browsers did *exactly* the same thing based on CSS and HTML standards
- How to get from “here” to “there” - no one could “win”



Quirks Mode

1993
SIMPLE
UGLY
HTML

HTML+
Proprietary
Extensions

HTML+
Proprietary
Extensions

Inconsistency
is OK

HTML+
Proprietary
Extensions

1999

Standard Mode

2008

Standard
CSS and
HTML

Web developers and
browsers are expected
to be “perfect.”
Responsibility goes
both ways.

Designing for the Browsers out there...

2009	IE7	IE6	IE8	Firefox	Chrome	Safari	Opera
August	15.1%	13.6%	10.6%	47.4%	7.0%	3.3%	2.1%
July	15.9%	14.4%	9.1%	47.9%	6.5%	3.3%	2.1%
June	18.7%	14.9%	7.1%	47.3%	6.0%	3.1%	2.1%
May	21.3%	14.5%	5.2%	47.7%	5.5%	3.0%	2.2%
April	23.2%	15.4%	3.5%	47.1%	4.9%	3.0%	2.2%
March	24.9%	17.0%	1.4%	46.5%	4.2%	3.1%	2.3%
February	25.4%	17.4%	0.8%	46.4%	4.0%	3.0%	2.2%
January	25.7%	18.5%	0.6%	45.5%	3.9%	3.0%	2.3%
September	49.1%	34.4%	4.3%	1.3%	4.3%	2.2%	
July	44.4%	40.1%	3.5%	1.2%	3.5%	2.6%	0.5%
May	40.7%	46.0%	2.8%	1.2%	2.7%	3.4%	0.7%
March	36.7%	49.4%	3.0%	1.2%	2.4%	4.1%	0.7%
January	30.1%	55.7%	2.8%	1.3%	2.2%	4.4%	1.0%

2009	IE7	IE6	IE8	Firefox	Chrome	Safari	Opera	pt On
August	15.1%	13.6%	10.6%	47.4%	7.0%	3.3%	2.1%	95%
July	15.9%	14.4%	9.1%	47.9%	6.5%	3.3%	2.1%	94%
June	18.7%	14.9%	7.1%	47.3%	6.0%	3.1%	2.1%	90%
May	21.3%	14.5%	5.2%	47.7%	5.5%	3.0%	2.2%	89%
April	23.2%	15.4%	3.5%	47.1%	4.9%	3.0%	2.2%	92%
March	24.9%	17.0%	1.4%	46.5%	4.2%	3.1%	2.3%	89%
February	25.4%	17.4%	0.8%	46.4%	4.0%	3.0%	2.2%	88%
January	25.7%	18.5%	0.6%	45.5%	3.9%	3.0%	2.3%	81%
September	49.1%	34.4%	4.3%	1.3%	4.3%	2.2%		80%

www.w3schools.com/browsers/browsers_stats.asp

You get to make a choice...

- (a) Do you want to be lazy and sloppy and hope that your pages look good across multiple browsers and multiple versions of the same browsers
- (b) Do you want to be professional and have your pages to render identically across all browsers?
- If you choose (b) - you take on some additional responsibility.

DOCTYPE: Indicating your Choice

- Browsers look at the first line of your HTML file to see if you have agreed to comply to HTML standards

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"  
  "http://www.w3.org/TR/html4/strict.dtd">  
<html>  
  <head>  
    <title>Learning the Google App Engine</title>  
  </head>  
  <body>  
    <h1>Google App Engine: About</h1>
```

Which DOCTYPE?

```
<!doctype HTML PUBLIC "-//W3C//DTD HTML 4.0I Transitional//EN"  
"http://www.w3.org/TR/html4/loose.dtd">
```

```
<!doctype HTML PUBLIC "-//W3C//DTD HTML 4.0I Transitional//EN">
```

```
<!doctype HTML PUBLIC "-//W3C//DTD HTML 4.0I//EN"  
"http://www.w3.org/TR/html4/strict.dtd">
```

<http://www.webmasterworld.com/html/3061635.htm>

Which DOCTYPE?

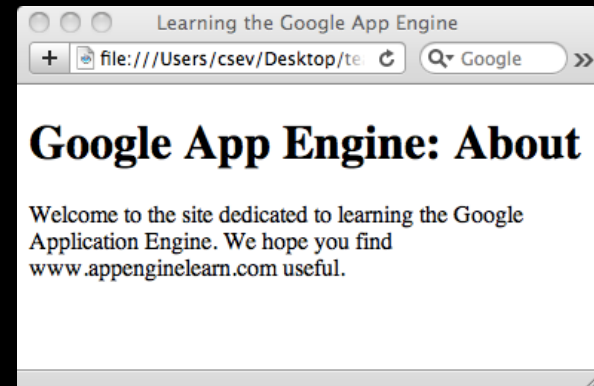
- Easy to start a fight - just Google “what is the best doctype”
 - Loose - legacy
 - Transitional - Mix of Developers
 - Strict - New Project which can enforce rules from the start
- <http://www.webmasterworld.com/html/3061635.htm> (and others)

Well-Formed HTML Documents

- Browsers need to know what contract you are agreeing to (doctype)
- We need a section for meta-data about the HTML document
- And then the displayable content of the HTML document

A Valid Web Document

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<html>
  <head>
    <title>Learning the Google App Engine</title>
  </head>
  <body>
    <h1>Google App Engine:About</h1>
    <p>
      Welcome to the site dedicated to
      learning the Google Application Engine.
      We hope you find www.appenginelearn.com useful.
    </p>
  </body>
</html>
```



A Valid Web Document

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<html>
  <head>
    <title>Learning the Google App Engine</title>
  </head>
  <body>
    <h1>Google App Engine:About</h1>
    <p>
      Welcome to the site dedicated to
      learning the Google Application Engine.
      We hope you find www.appenginelearn.com useful.
    </p>
  </body>
</html>
```

Contract about the
HTML style used
the page.

Metadata about
the page.

Displayable content
of the page.

Checking Validity

- We can check to see if a page “meets the rules” - does it live up to its doctype?
- <http://validator.w3.org>
- Enter, HTML, upload a file, or enter a web-accessible URL

Validation Output: 4 Errors

✖ Line 13, Column 7: end tag for "p" omitted, but OMITTAG NO was specified.

```
</body >
```

You may have neglected to close an element, or perhaps you meant to "self-close" an element, that is, ending it with "/>" instead of ">".

Success

Less-than-successful

Note: When the validator fails, only look at the first message and fix that - and then re-run. One simple error will often cascade into 30-40 "The validator is lost and confused errors".

The screenshot shows the W3C Markup Validation Service interface. At the top, it says "W3C® Markup Validation Service" and "Check the markup (HTML, XHTML, ...) of Web documents". Below that, there are links for "Jump To: Notes and Potential Issues" and "Congratulations · Icons". A green banner states "This document was successfully checked as HTML 4.01 Strict!". The "Result:" section shows "Passed, 2 warning(s)". The "Source:" section displays the HTML code:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
<html>
<head>
<title>Learning the Google App Engine</title>
</head>
<body>
<h1>Google App Engine: About</h1>
<p>
Welcome to the site dedicated to
learning the Google Application Engine.
We hope you find www.appenginelearn.com useful.
```

 The "Encoding:" is set to "utf-8" and the "Doctype:" is set to "HTML 4.01 Strict". The "Root Element:" is "HTML".

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
<html>
<head>
<title>Learning the Google App Engine</title>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" >
</head>
<body>
<h1>Google App Engine:About</h1>
<p>
Welcome to the site dedicated to
learning the Google Application Engine.
We hope you find www.appenginelearn.com useful.
</p>
</body>
</html>
```

To fix the warning - we add a header to indicate the character set of our document.

This site contains a bunch of stuff including software, television shows, and other miscellaneous material. This also includes a number of activities and sites which I have done with my good friend Richard Wiggins.

Recent Video: [Dr. Chuck Goes Motocross Racing](#)
Funny Video: [Mandy Birthday Project - Remix](#)
Updated: [Community Information Toolkit - Building Community Networks](#)

Quick Links

Chuck's TV and Media
 Chuck's Talks
 Chuck's Book
 Chuck's Columns
 Resume and Bio Blog (20-Sep 10:01)
 Chuck's Papers

Sakai Project
 Sakai Planet Blog

Software Tools
 PHP Image software for v300 and Treo-600 SMS
 ClipBoard-2000 - Distance Education Software
 Sync-O-Matic 2000 - Distance Education Software
 Other software Tools

Projects
 Practical Home Networking - A book that Chuck and Rich almost got written. (Version 1, Version 2, Version 3, and Version 4)
 The Community Information Toolkit - A project to provide public libraries and other organizations a start on using Internet in Community Networking.

Photo log last update: September 24 2007 09:21:39 PM
Blog last update: September 20 2007 10:01:30 PM

Now Playing on Netfact TV

Online Lecture: HTML Overview

Production date: 1997 This lecture provides a basic overview of HTML and its use. This is a lecture from a fully-online course on the Internet and technology. Lecturer: Charles Severance Details: 50 minutes, Real Media 28kbps.

This is a randomly selected show from our library of shows - more shows will be added as they are uploaded. You can view the entire collection here.

Sakai

W3C Markup Validation Service
 Check the markup (HTML, XHTML) of Web documents

Jump To: [Important Warnings](#) [Validation Output](#)

This page is not Valid (no Doctype found)!

Result: Failed validation, 47 Errors

Address:

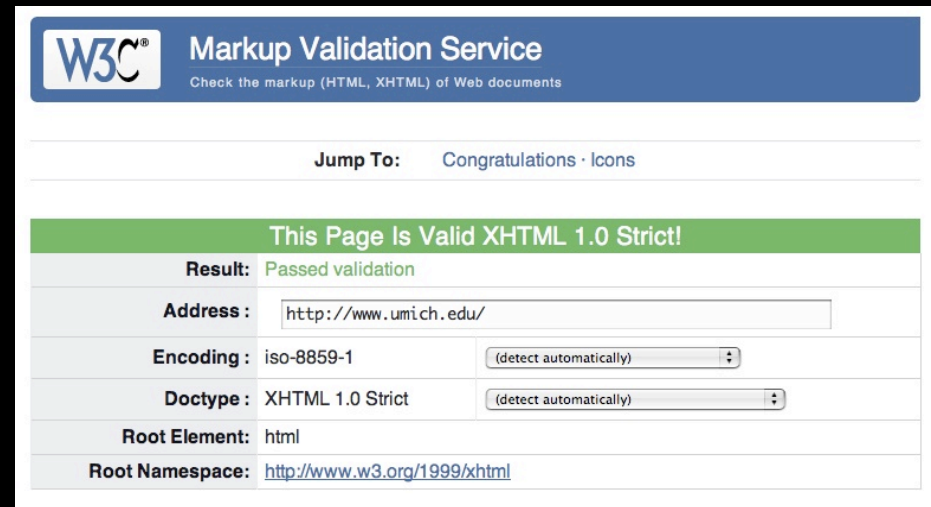
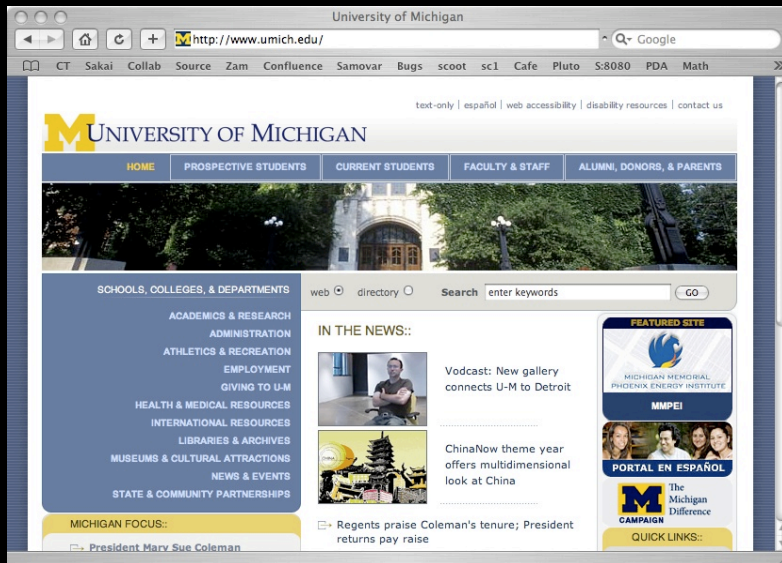
Encoding: utf-8

Doctype: (no Doctype found)

Root Element: html

```
<HTML>
<HEAD>
  <META NAME="GENERATOR" CONTENT="Adobe PageMill 3.0 Win">
  <TITLE>dr-chuck.com </TITLE>
</HEAD>
<BODY BGCOLOR="#000000" LINK="#AAAAAA" VLINK="#AAAAAA" ALINK="#AAAAAA">
<table Border=0>
<tr>
...

```



```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://
www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
<head>
<title>University of Michigan</title>
```

....

Summary

- HTML has gone through many changes and evolutions
 - It started clean and simple - then got ugly and nasty - now we are back to a clean and simple approach
- HTML Markup needs to focus on meaning - not formatting
- Formatting is handled using CSS - Cascading Style Sheets