HTML

Lab. Bases de Dados e Aplicações Web
MIEIC, FEUP 2010/11

Sérgio Nunes

Summary

- Quick Overview.
- A Brief History of HTML.
- The HTML Language.
- A Note on HTML5.

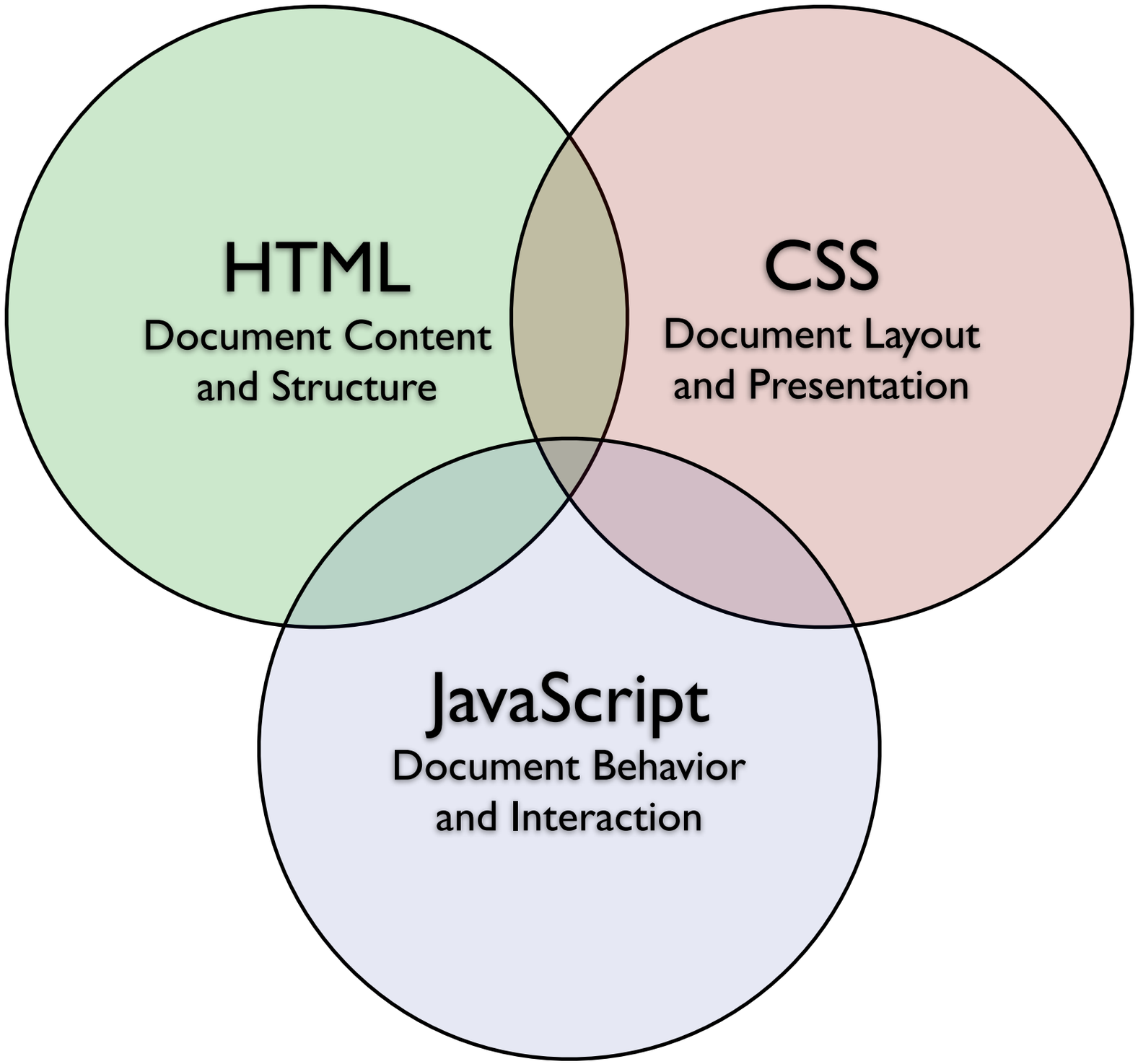
The Big Picture

Web browsers issue requests to web servers, which produce and return HTML documents for browsers to parse and display.

1. HTTP request

2. HTTP answer + HTML document

Client-Side Technologies



"The World Wide Web (W3) initiative links related information throughout the globe.

HTML provides one simple format for providing linked information, and all W3 compatible programs are required to be capable of handling HTML."

in HTTP RFC Draft, IETF (1993)

HTML

- Stands for HyperText Markup Language and is a format for providing linked information.
- HTML documents are simply text files containing marked-up text using tags.
- A HTML document is a hypertext node within a hypertext network.

Hypertext

Concept defined by Ted Nelson in the 1950s.

"Hypertext: Human-readable information linked together in an unconstrained way."

"HyperText is a way to link and access information of various kinds as a web of nodes in which the user can browse at will.

It provides a single user-interface to large classes of information (reports, notes, data-bases, computer documentation and on-line help)."

in WorldWideWeb: Proposal for a HyperText Project (1990)

simple.html

```
<!DOCTYPE html>
<html>
  <head>
    <title>A simple HTML5 document</title>
  </head>

  <body>
    <h1>Simple HTML5 document</h1>
    <p>This is a simple HTML5 document.</p>
  </body>
</html>
```

A simple HTML5 document

file:///...

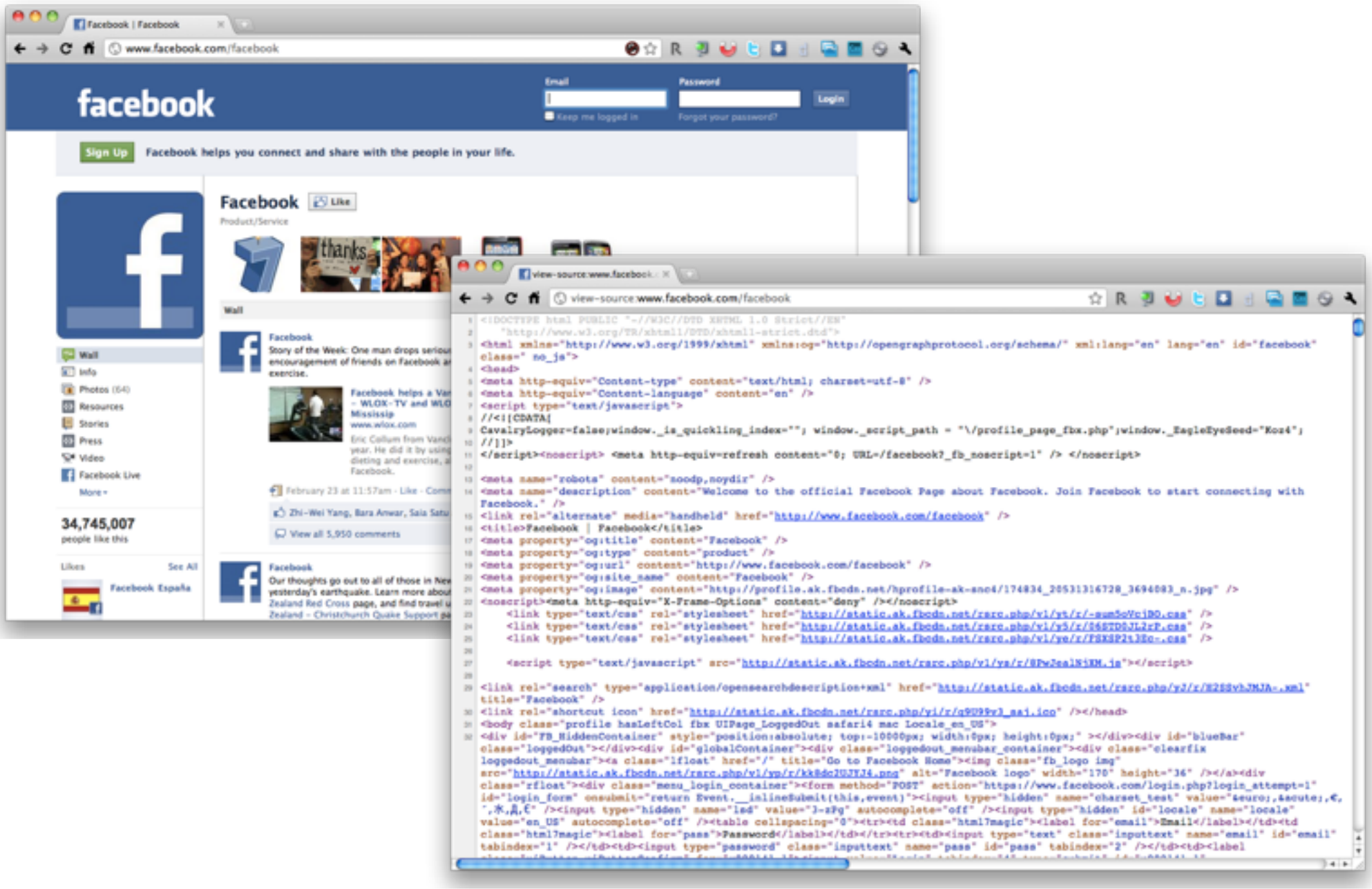
Simple HTML5 document

This is a simple HTML5 document.

view-source:web.fe.up.pt/~ssn/2010/lbaw/

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
<html>
<head>
  <title>Laboratório de Bases de Dados e Aplicações Web (LBAW), 2010/11</title>
  <meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">

  <link rel="stylesheet" href="/~ssn/web.css" type="text/css" media="screen">
  <style type="text/css">
    li { margin-top: 4px; }
  </style>
</head>

<body class="disciplina" lang="pt">

<p><a href="/~ssn">Início</a> : LBAW</p>

<h1>Laboratórios de Bases de Dados e Aplicações Web, 2010/11</h1>

<p>Nesta disciplina pretende-se dotar os estudantes da capacidade de projectar e desenvolver sistemas de informação
acessíveis através da web e suportados por sistemas de gestão de bases de dados.</p>

<p><strong>SIGARRA</strong>: <a href="https://www.fe.up.pt/si/disciplinas_geral.formview?
p_cad_codigo=EIC0085&p_ano_lectivo=2010/2011&p_periodo=2S">Ficha da Disciplina</a>, <a
href="https://www.fe.up.pt/si/alunos_inscricoes.fotos?
p_fotos=4&p_a_lectivo=2010/2011&p_dis=EIC0085&p_periodo=2S">Alunos</a>, <a
href="https://www.fe.up.pt/si/inscricoes_turmas_LISTAGEM.lista_turma_disciplina?
p_cod_curso=9459&p_cad_codigo=EIC0085&p_ano_lectivo=2010/2011&p_periodo=2S">Turmas</a>, <a
href="https://www.fe.up.pt/si/horarios_geral.disciplinas_view?
pv_dis_codigo=EIC0085&pv_periodos=2S&pv_ano_lectivo=2010/2011">Horário</a>, <a
href="https://www.fe.up.pt/si/sumarios_geral.ver?
pv_ano_lectivo=2010/2011&pv_cad_codigo=EIC0085&pv_periodo=2S">Sumários</a>.</p>

<p><em style="color:red;">Documento provisório em actualização.</em></p>

<h2>Datas Importantes</h2>
<ul>
<li>Relatório de Especificação de Requisitos (RER) - <strong>7 Março</strong></li>
<li>Relatório de Especificação da Base de Dados (REBD) - <strong>28 Março</strong></li>
<li>Relatório de Arquitectura e Protótipo Vertical (RAP) - <strong>25 Abril</strong></li>
<li>Produto e Apresentação (PA) - <strong>6 Junho</strong></li>
</ul>


<div id="documentos">
```

View Source


Facebook | Facebook
www.facebook.com/facebook
facebook
Email
Password
Login
Keep me logged in
Forgot your password?
Sign Up
Facebook helps you connect and share with the people in your life.
Facebook
Like
Product/Service
Wall
Info
Photos (64)
Resources
Stories
Press
Video
Facebook Live
More
34,745,007
people like this
Likes
See All
Facebook España
view-source:www.facebook.com/facebook
<!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" xmlns:og="http://opengraphprotocol.org/schema/" xml:lang="en" lang="en" id="facebook"
class=" no_js">
<head>
<meta http-equiv="Content-type" content="text/html; charset=utf-8" />
<meta http-equiv="Content-language" content="en" />
<script type="text/javascript">
//<![CDATA[
CavalryLogger=false;window._is_quickling_index=""; window._script_path = "\/profile_page_fbx.php";window._EagleEyeSeed="Kos4";
//]]>
</script><noscript> <meta http-equiv=refresh content="0; URL=/facebook?_fb_noscript=1" /> </noscript>
<meta name="robots" content="noodp,noydir" />
<meta name="description" content="Welcome to the official Facebook Page about Facebook. Join Facebook to start connecting with
Facebook." />
<link rel="alternate" media="handheld" href="http://www.facebook.com/facebook" />
<title>Facebook | Facebook</title>
<meta property="og:title" content="Facebook" />
<meta property="og:type" content="product" />
<meta property="og:url" content="http://www.facebook.com/facebook" />
<meta property="og:site_name" content="Facebook" />
<meta property="og:image" content="http://profile.ak.fbcdn.net/hprofile-ak-snc4/174834_28531316728_3694083_n.jpg" />
<noscript><meta http-equiv="X-Frame-Options" content="deny" /></noscript>
<link type="text/css" rel="stylesheet" href="http://static.ak.fbcdn.net/rsrc.php/v1/yt/r/-wum5oVrjBO.css" />
<link type="text/css" rel="stylesheet" href="http://static.ak.fbcdn.net/rsrc.php/v1/y5/r/S6STD6JL2rF.css" />
<link type="text/css" rel="stylesheet" href="http://static.ak.fbcdn.net/rsrc.php/v1/ye/r/PSXSP2t3Ec-.css" />
<script type="text/javascript" src="http://static.ak.fbcdn.net/rsrc.php/v1/ys/r/8PwZealNiXM.js"></script>
<link rel="search" type="application/opensearchdescription+xml" href="http://static.ak.fbcdn.net/rsrc.php/y2/r/H2S8vhJMJA-.xml"
title="Facebook" />
<link rel="shortcut icon" href="http://static.ak.fbcdn.net/rsrc.php/v1/r/q9U99v1_sai.ico" /></head>
<body class="profile hasLeftCol fbx UIPage_LoggedOut safari4 mac Locale_en_US">
<div id="blueBar"
class="loggedOut"> <div class="clearfix
loggedout_menubar"><img class="fb_logo img"
src="http://static.ak.fbcdn.net/rsrc.php/v1/yp/r/kk8dcJUJYJ4.png" alt="Facebook logo" width="170" height="36" /><div
class="rfloat"> <form method="POST" action="https://www.facebook.com/login.php?login_attempt=1"
id="login_form" onsubmit="return Event.__inlineSubmit(this,event)"><input type="hidden" name="charset_test" value="€,´,€,
´,水,Д,Є" /><input type="hidden" name="lsd" value="]-zPg" autocomplete="off" /><input type="hidden" id="locale" name="locale"
value="en_US" autocomplete="off" /><table cellspacing="0"><tr><td class="html7magic"><label for="email">Email</label></td><td
class="html7magic"><label for="pass">Password</label></td></tr><tr><td><input type="text" class="inputtext" name="email" id="email"
tabindex="1" /></td><td><input type="password" class="inputtext" name="pass" id="pass" tabindex="2" /></td><td><label

Why Learn HTML?

There are many editors available,
why learn to code in HTML directly?

- “HTML Editor X expert” is a narrower expertise.
- Editors get in the way of coding.
- Editors are not always up to date.
- Need to master HTML details to
  fully explore all possibilities.
- In dynamic web sites HTML is generated.

History

HTML

- Created by Tim Berners-Lee and Robert Caillau at CERN in the late 1980s.
- Main goal was to facilitate document sharing between researchers over the network.
- CERN released it as royalty free in 1993.
- First official version published by IETF in 1993.
- W3C was created to define common standards for browsers and developers to adhere to.

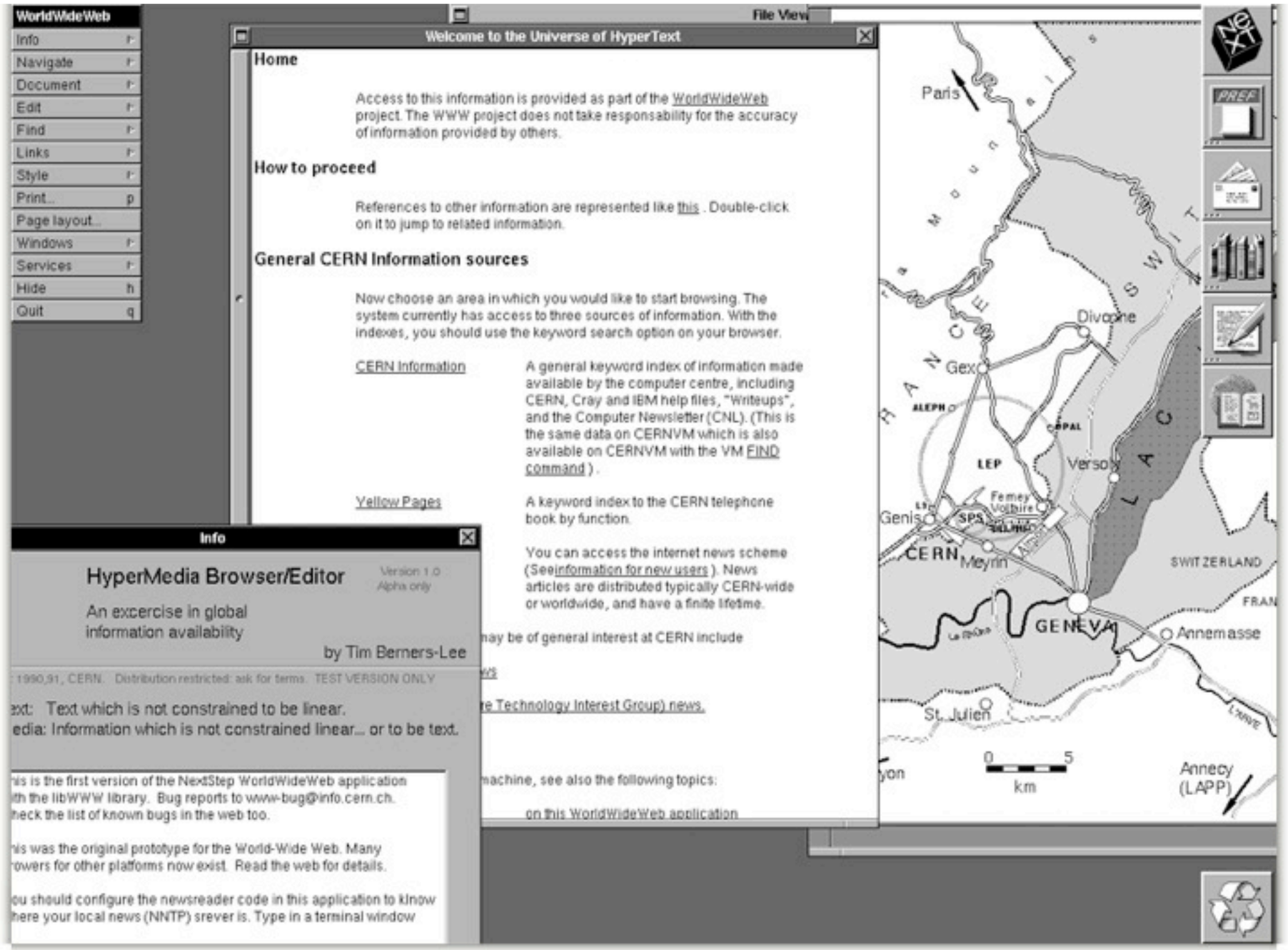
“Information Management: A Proposal”
Tim Berners-Lee, CERN (1990)

“This proposal concerns the management of general information about experiments at CERN.”

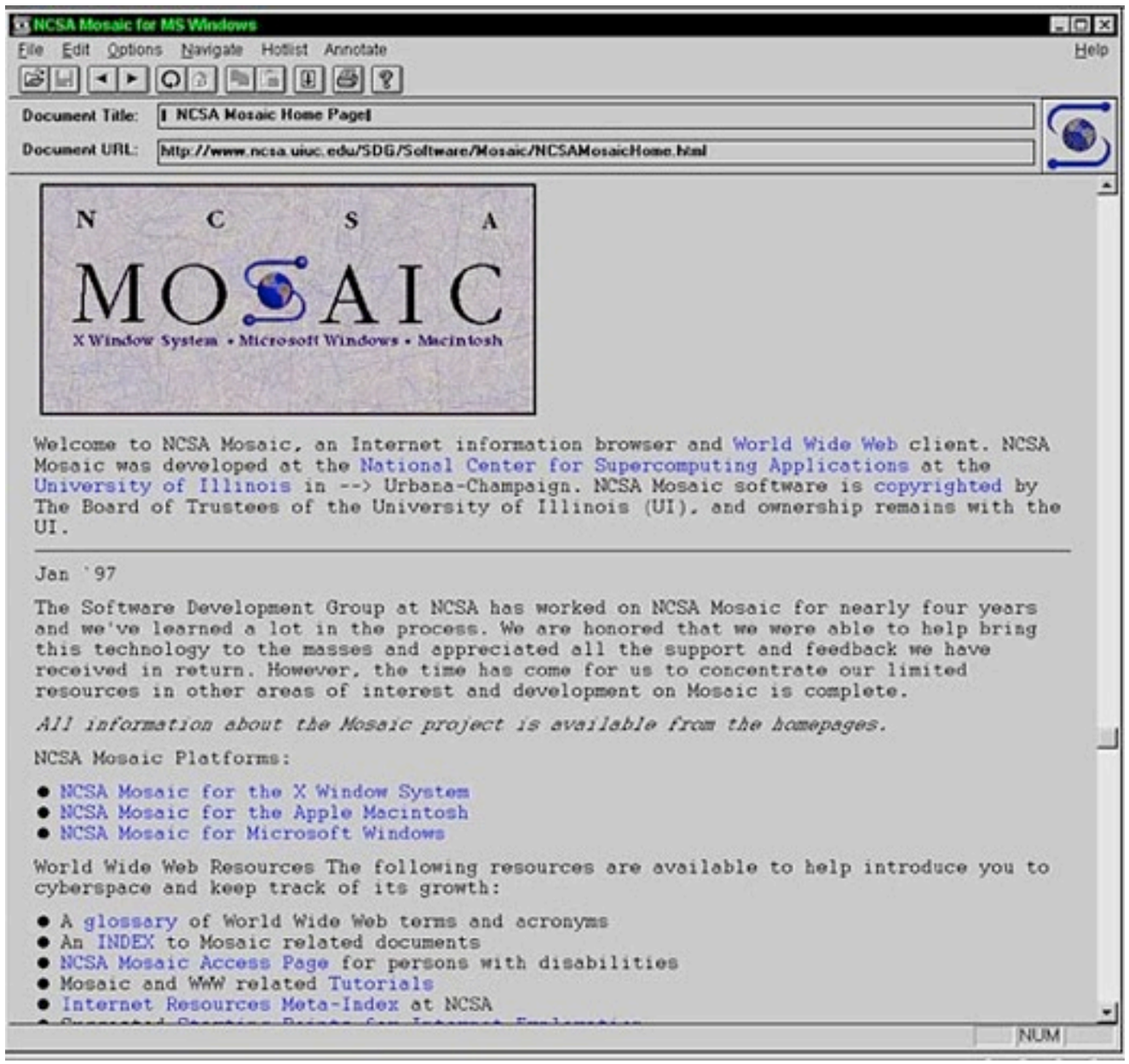
“It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.”

Some practical requirements: remote access, heterogeneity, non-centralization, text-based, “live links”.

WWW in 1993


WorldWideWeb
Info
Navigate
Document
Edit
Find
Links
Style
Print... p
Page layout...
Windows
Services
Hide h
Quit q
Welcome to the Universe of HyperText
Home
Access to this information is provided as part of the WorldWideWeb project. The WWW project does not take responsability for the accuracy of information provided by others.
How to proceed
References to other information are represented like this . Double-click on it to jump to related information.
General CERN Information sources
Now choose an area in which you would like to start browsing. The system currently has access to three sources of information. With the indexes, you should use the keyword search option on your browser.
CERN Information
A general keyword index of information made available by the computer centre, including CERN, Cray and IBM help files, "Writeups", and the Computer Newsletter (CNL). (This is the same data on CERNVM which is also available on CERNVM with the VM FIND command).
Yellow Pages
A keyword index to the CERN telephone book by function.
You can access the internet news scheme (Seeinformation for new users). News articles are distributed typically CERN-wide or worldwide, and have a finite lifetime.
Info
HyperMedia Browser/Editor
Version 1.0 Alpha only
An excercise in global information availability
by Tim Berners-Lee
1990,91, CERN. Distribution restricted: ask for terms. TEST VERSION ONLY
ext: Text which is not constrained to be linear.
edia: Information which is not constrained linear... or to be text.
his is the first version of the NextStep WorldWideWeb application
ith the libWWW library. Bug reports to www-bug@info.cern.ch.
heck the list of known bugs in the web too.
his was the original prototype for the World-Wide Web. Many
owers for other platforms now exist. Read the web for details.
ou should configure the newsreader code in this application to know
here your local news (NNTP) srever is. Type in a terminal window
Paris
Divonne
Gex
ALEPH
LEP
Versoix
Genis
SPS
CERN
Meyrin
SWITZERLAND
GENEVA
Annemasse
St. Julien
0 5
km
Annecy (LAPP)

NCSA Mosaic


NCSA Mosaic for MS Windows
File Edit Options Navigate Hotlist Annotate Help
Document Title: NCSA Mosaic Home Page
Document URL: http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/NCSAMosaicHome.html
N C S A
MOSAIC
X Window System • Microsoft Windows • Macintosh
Welcome to NCSA Mosaic, an Internet information browser and World Wide Web client. NCSA Mosaic was developed at the National Center for Supercomputing Applications at the University of Illinois in --> Urbana-Champaign. NCSA Mosaic software is copyrighted by The Board of Trustees of the University of Illinois (UI), and ownership remains with the UI.
Jan '97
The Software Development Group at NCSA has worked on NCSA Mosaic for nearly four years and we've learned a lot in the process. We are honored that we were able to help bring this technology to the masses and appreciated all the support and feedback we have received in return. However, the time has come for us to concentrate our limited resources in other areas of interest and development on Mosaic is complete.
All information about the Mosaic project is available from the homepages.
NCSA Mosaic Platforms:
● NCSA Mosaic for the X Window System
● NCSA Mosaic for the Apple Macintosh
● NCSA Mosaic for Microsoft Windows
World Wide Web Resources The following resources are available to help introduce you to cyberspace and keep track of its growth:
● A glossary of World Wide Web terms and acronyms
● An INDEX to Mosaic related documents
● NCSA Mosaic Access Page for persons with disabilities
● Mosaic and WWW related Tutorials
● Internet Resources Meta-Index at NCSA
NUM

W3C

The mission of the World Wide Web Consortium (W3C) is to lead the World Wide Web to its full potential by developing common protocols that promote its evolution and ensure its interoperability.

W3C Process

- A new ‘topic’ is introduced by a member, either as a note or as result of a workshop.
- A new working group is formed or the topic is assigned to an existing group.
- Work on specifications or guidelines progresses from Technical Reports to W3C Recommendations.

HTML Timeline

- During its first years (1990-1995), HTML revisions and extensions where first hosted at CERN and then IETF.
- Development was moved to the W3C after its creation in 1994. HTML development stopped in 1998 with the publication of HTML4.
- W3C decided to migrate to a XML-based equivalent, named XHTML. No wide adoption by web authors.
- HTML development continued outside W3C, with the WHATWG, whose work is now the basis for HTML5.

HTML Ages

The Early Days

- From proposal (1989) to Mosaic (1993).
- Web users were mostly from academia.
- Few browsers, most of them text-based.
- HTML documents were simple
  and usually written by hand.

Growth Years

- Wide adoption of the Web - dot.com bubble (1995-2000).
- Companies dispute the web browser market -
  aka “browser wars”.
- Browser development focused on new features,
  less on standards support.
- Wide differences between rendering engines.
  Many web pages “designed for browser version x.x”.
- Extensive use of tables and sliced graphics to achieve
  “pixel perfect” layouts - “print-like design”.
  Resulted in ugly and complex HTML code.

Modern Era

- Wide adoption of modern web browsers.
- Separation of content and structure from layout and presentation.
- HTML controls content and structure.
- CSS controls layout and presentation.
- Clean and simple code (again!).
- CSS (2003), AJAX (2005), mobile (2007).

HTML

HTML Document

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
   "http://www.w3.org/TR/html4/strict.dtd">

<html>
  <head>
    <title>Document Title</title>
  </head>

  <body>
    <h1>Header</h1>
    <p>Document Body.</p>
  </body>
</html>
```

HTML Elements

- HTML documents consist of a tree of elements and text.
- Each HTML element has three parts: start tag, content and end tag. There are exceptions.

```
<p>HyperText Markup Language</p>

<h1>HTML Elements</h1>
```

Nesting Elements

- HTML tags must be properly nested.
- Tags have to be nested such that elements are all completely within each other, without overlapping.

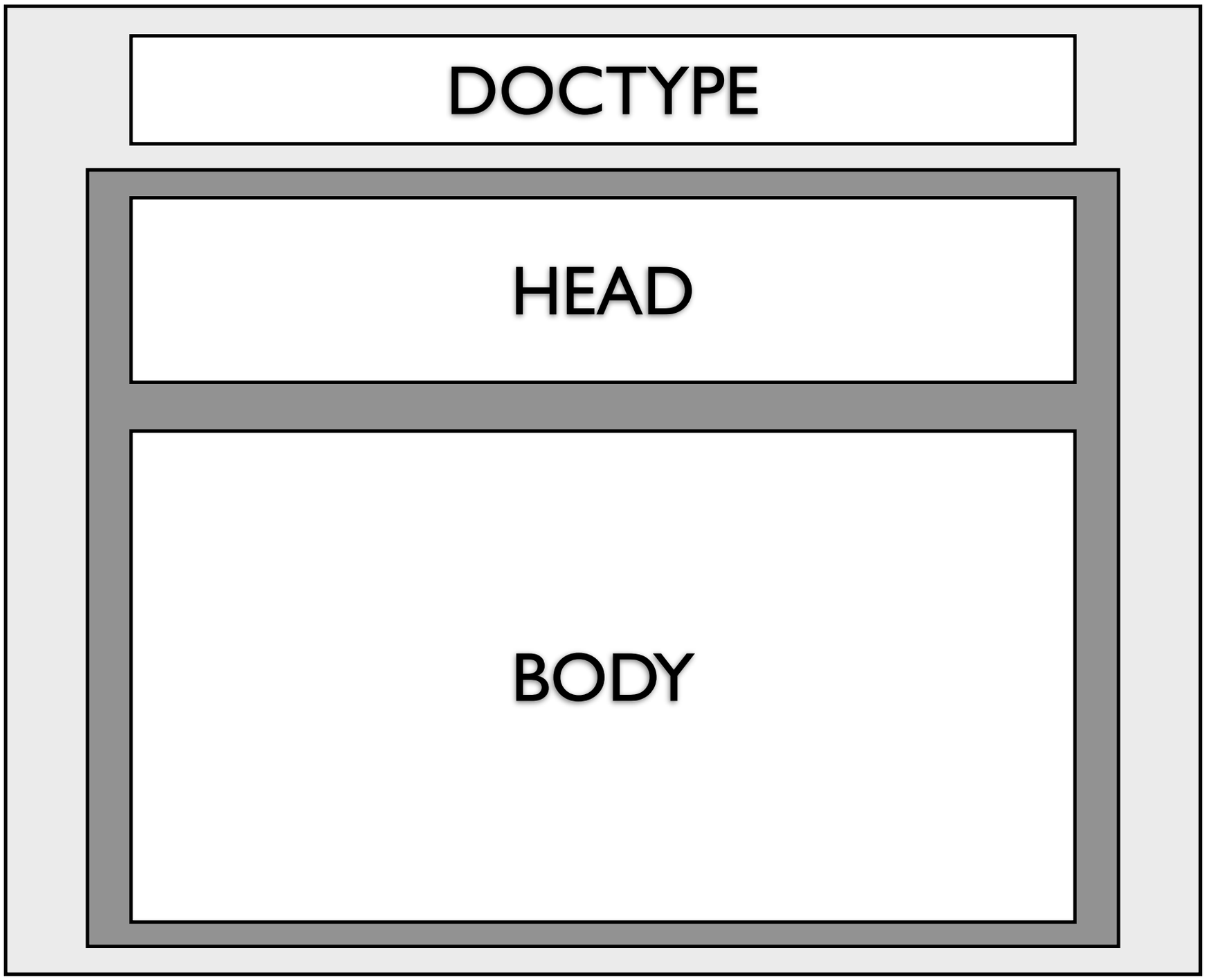
```
<p>This is <strong>correct</strong>!</p>

<p>This is <strong>wrong</p>!</strong>
```

Element Attributes

- HTML elements can have attributes, which control how the elements work.
- Attributes are placed inside the start tag, and consist of a name and a value.
- A single start tag can have multiple attributes.
- There are mandatory and optional attributes.

```
<a href="file.html">this is a link</a>

<img src="logo.png" alt="Company Logo">
```

HTML Document



HTML Document

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
   "http://www.w3.org/TR/html4/strict.dtd">

<html>
  <head>
    <title>Document Title</title>
  </head>

  <body>
    <h1>Header</h1>
    <p>Document Body.</p>
  </body>
</html>
```

DOCTYPE

The document type declaration is a contract about the HTML version used.

```
<!DOCTYPE html PUBLIC "-//IETF//DTD HTML 2.0//EN">
```

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
   "http://www.w3.org/TR/html4/strict.dtd">
```

```
<!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 1.1//EN"
   "http://www.w3.org/Graphics/SVG/1.1/DTD/svg11.dtd">
```

<!DOCTYPE html>

HEAD

- The head part of a HTML document contains information about the document itself. This information is used by the browser for rendering.
- The <head> must contain a <title> element.
- Optional elements <meta> and <link>.
- The <meta> element provides additional information about a document itself.
  - E.g. <meta name="author" value="John Doe">
  - Other meta properties: revised, generator, description, keywords, encoding, etc.

<link>

- The <link> element connects the HTML document to other resources, e.g. style sheets, icons, navigation, etc.
- The location of the resource is given by the href attribute.
- The attribute rel (relationship) is mandatory.

```
<link rel="stylesheet" href="base.css">
```

BODY

The body part of a HTML document contains the actual content of the web page. There is only one <body> element.

HTML Comments

- Comments in HTML code are ignored by browsers when rendering the document.
- Useful for documentation or for structuring the document.

```
<!-- this is a comment -->

<!-- this is
also a comment -->
```

Hyperlinks

- One of the key features of HTML is hypertext support.
- Links are the basic hypertext construct. A link is a connection from one Web resource to another.
- HTML links are created with the <a> element (anchor). Together with the href attribute specifies the destination resource.

```
<a href="http://www.google.com">Google</a>

<a href="image.jpg">A flower picture</a>
```

Naming Files

- A URL points to a directory and a file in a web server. Pay attention to special characters.
- The web server returns a default index file when none is requested, e.g. http://www.google.com
- Usually the default index file is named index.html. Other possibilities include: index.htm, default.htm, index.php, etc. This can be configured.

Best Practices

- Start with simple code.
- Reuse available templates and examples.
- Use a modern standard compliant browser for development and debugging.
- Handle browser ‘bugs’ later.

HTML Elements

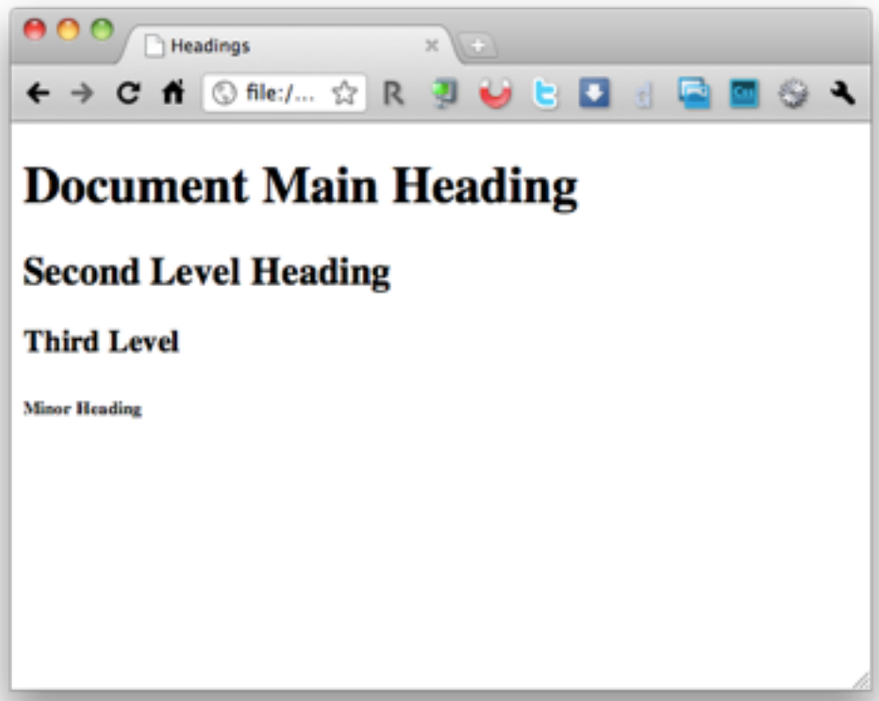
Block vs. Inline

- HTML elements can be displayed either in block or inline style.
- A block-level element spans the full width of the space available, starting a new line in the flow of HTML. Examples: headings, paragraphs.
- An inline element doesn't break the flow, fitting within the flow of the document. Examples: links, emphasis, images.
- Inline elements can contain other inline elements.
- Block elements cannot be nested inside inline elements.

Headings

- HTML supports several levels of headings for structuring documents.
- H1 is the highest level of heading.
  H6 is the lowest.

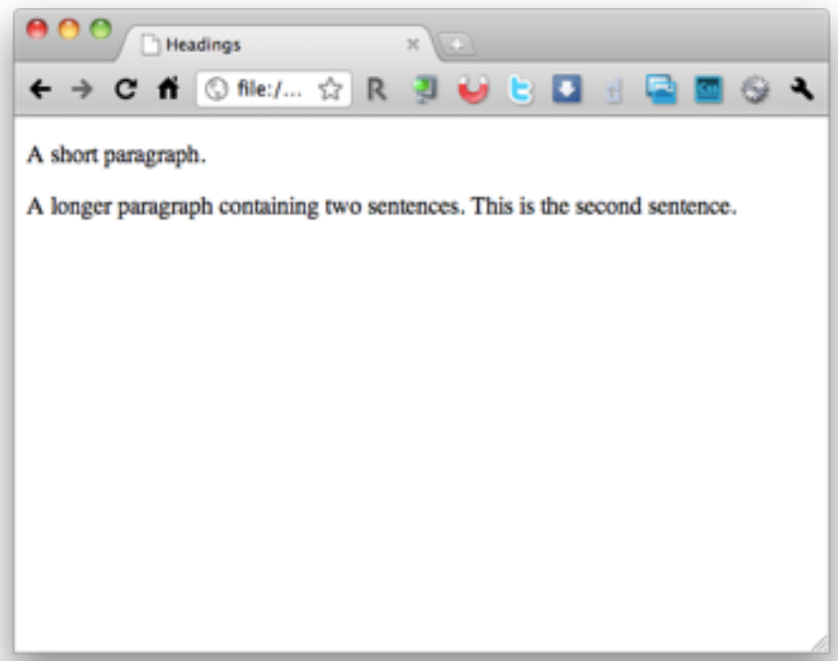
```
<h1>Document Main Heading</h1>
<h2>Second Level Heading</h2>
<h3>Third Level</h3>

<h6>Minor Heading</h6>
```



Paragraphs

- The <p> element represents a paragraph.
- Line breaks can be controlled with
.

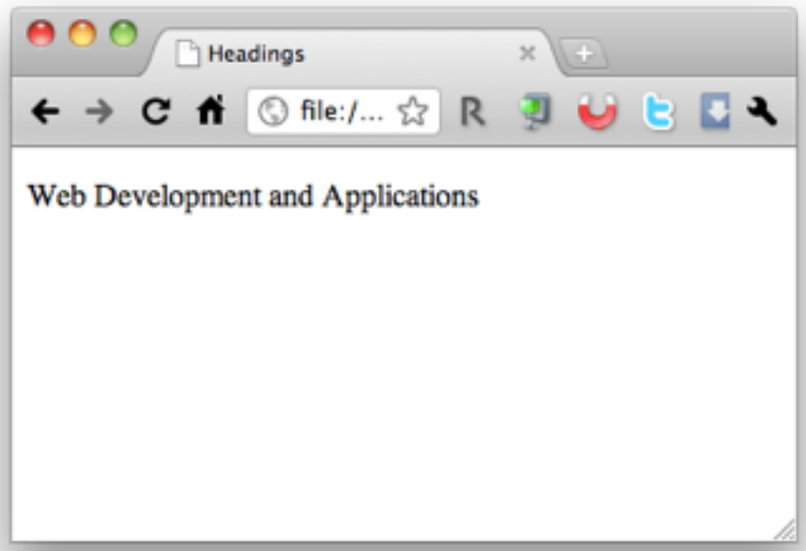
```
<p>A short paragraph.</p>
<p>A longer paragraph containing
two sentences.

This is the second sentence.</p>
```



Text

- In HTML documents, white spaces and newlines are collapsed into a single white space.
- Newlines can be forced with
.
- White spaces can be introduced with   (non-breaking space).

```
<p>Web
Development
and
        Applications</p>
```



Character References

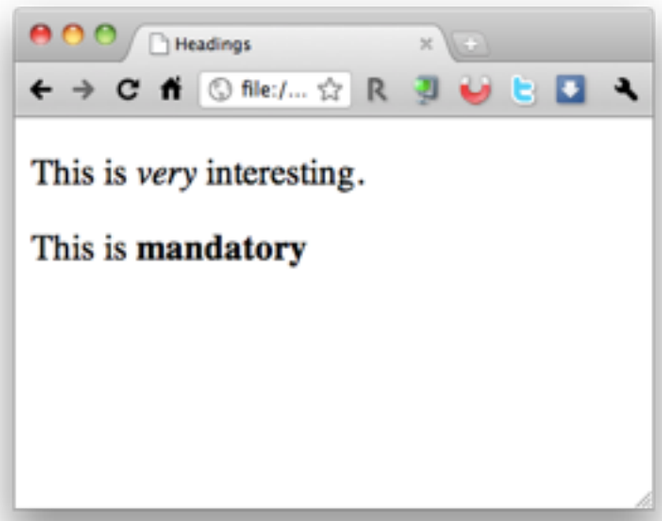
- Character references are numeric or symbolic names for characters that may be included in a HTML document.
- Useful for rarely used characters, or those that authoring tools make difficult or impossible to enter.

< - <	& - &
> - >	á - é
© - ©	à - á
ç - ç	ã - ã
	ä - ä

Text Highlighting

- There are two HTML elements for highlighting isolated portions of text.
- The element is intended to convey emphasis. Typically rendered as italics.
- The element is intended to convey importance. Typically rendered as bold.

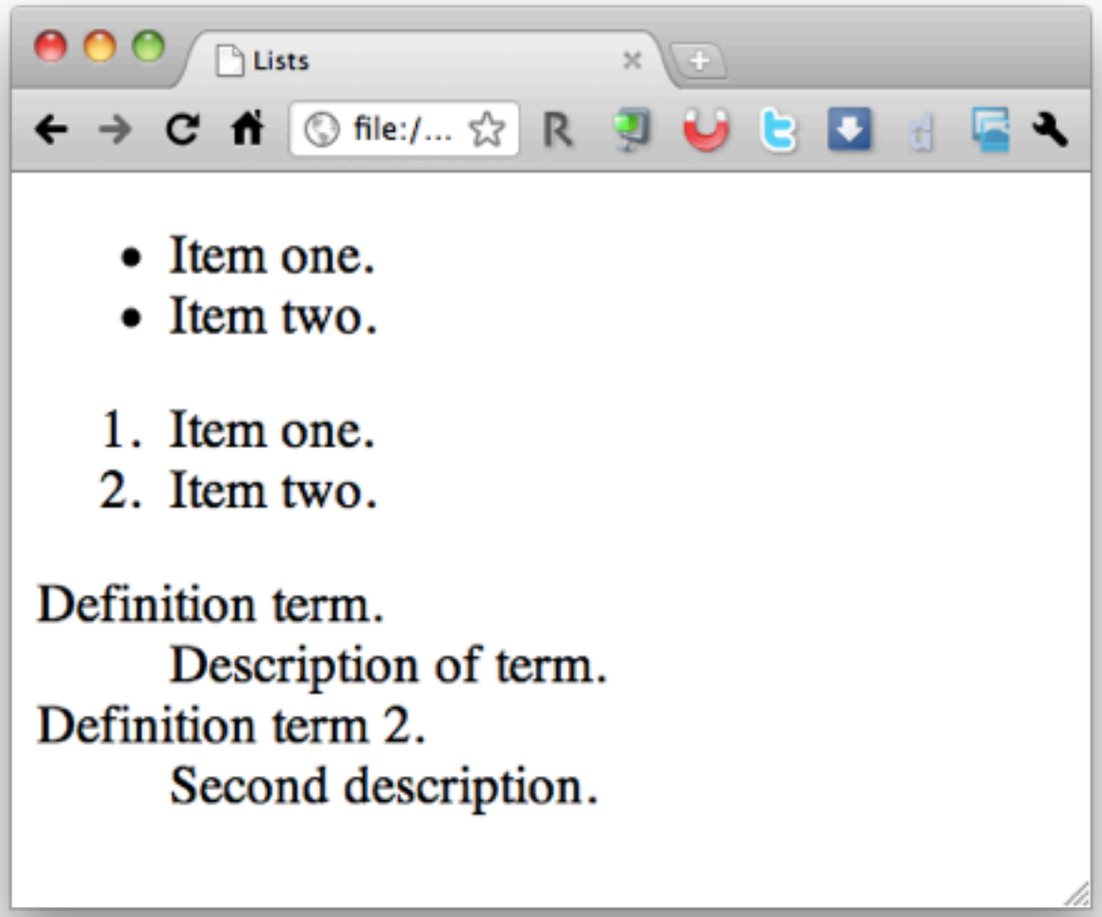
```
<p>This is <em>very</em> interesting.</p>

<p>This is <strong>mandatory</strong>.</p>
```



Lists

- Three types of lists: unordered, ordered and association lists.
- The <ol> element represents a list of items, where the items have been intentionally ordered.
- The <ul> element represents a list of items, where the order of the items is not important.
- Each list item is defined with the <li> element.
- The <dl> element represents an association list, consisting of name-value groups. Used together with <dt> and <dd>.

```
<ul>
<li>Item one.</li>
<li>Item two.</li>
</ul>


<ol>
<li>Item one.</li>
<li>Item two.</li>
</ol>


<dl>
<dt>Definition term.</dt>
<dd>Description of term.</dd>
<dt>Definition term 2.</dt>
<dd>Second description.</dd>
</dl>
```

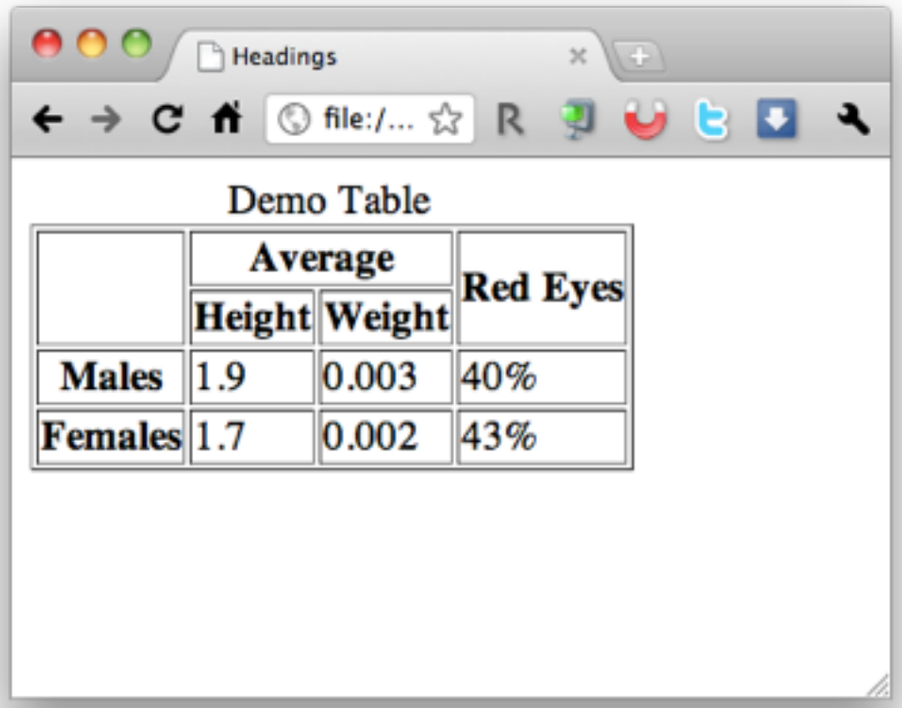


Tables

- Used to present tabular information.
- Tables are not for web page layout.
- Table elements:
  - <table> - root element.
  - <tr> - table row.
  - <th> - table heading.
  - <td> - table division.

Table Example

```
<table border="1">
<caption>Demo Table</caption>
<tr>
  <th rowspan="2"></th>
  <th colspan="2">Average</th>
  <th rowspan="2">Red Eyes</th>
</tr>
<tr>
  <th>Height</th>
  <th>Weight</th>
</tr>
<tr>
  <th>Males</th><td>1.9</td><td>0.003</td><td>40%</td>
</tr>
<tr>
  <th>Females</th><td>1.7</td><td>0.002</td><td>43%</td>
</tr>
</table>
```



Links

- The link, hyperlink or Web link, is the basic hypertext construct. A link is a connection from one Web resource to another.
- A link has two ends (anchors) and a direction. The link starts at the source anchor and points to the destination anchor.
- The <a> element defines an anchor.
  The href attribute defines the destination.

```
<a href="http://www.up.pt">University of Porto</a>

<a href="photo.jpg">A photo</a>
```

Hyperlink References

- Hyperlinks can either be absolute or relative.
- Absolute links always resolve to the same destination regardless of the originating document.

```
<a href="http://www.google.com">Google</a>
```

- Relative links are resolved according to the base address of the document.

```
<a href="aboutus.html">About Us</a>

<a href="../../docs/index.html">Documentation</a>

<a href="./team.html">Team</a>
```

Destination Anchors

- When the name or id attributes of the <a> element are set, the element defines an anchor that may be the destination of other links.
- Can be used to establish links within the same document (e.g. TOC).

```
<a name="section1">Section One</a>
```

```
<a href="document.html#section1">Section 1</a>
```

Images

- The <img> element embeds an image in the current document at the location of the element's definition.
- The location of the image is given by the src attribute.
- The <img> element has no content and no end tag.

```
<img src="photo.jpg">

<img src="http://www.google.com/logo.png">
```

Image Formats

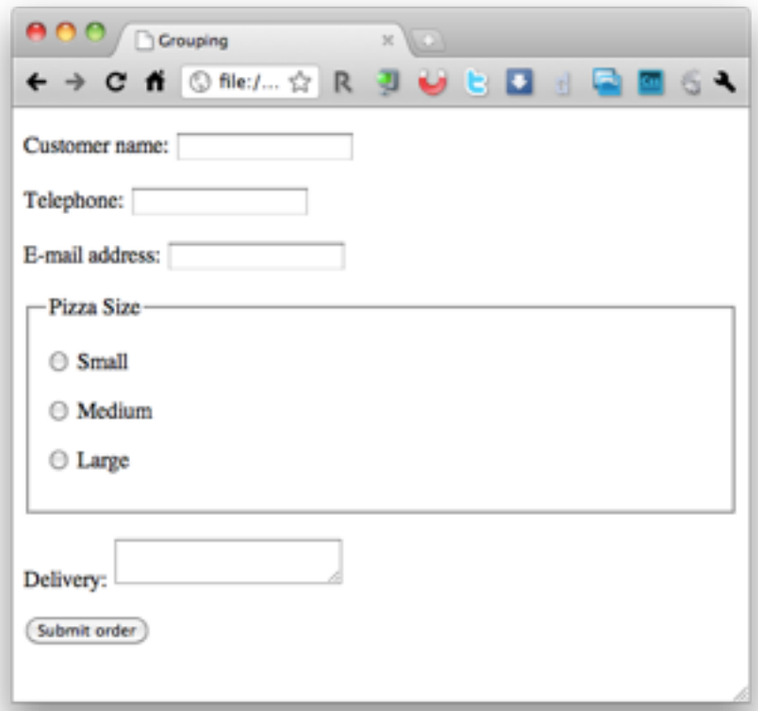
- JPEG is the format of choice for photographic images.
- GIF was the original format for drawn images (e.g. logos, diagrams). The GIF format should be avoided.
- PNG is the new format of choice.
  The PNG format supports transparency.
- SVG is a vector format. Only recently browsers have started to support it.

Forms

- HTML forms provide a way for users to send information to the servers.
- A form is a section of a web document that contains special elements called controls (e.g. checkboxes, radio buttons, input fields, etc).
- A user can interact with a form, providing data that can then be sent to the server for further processing.
- Form submissions are sent to servers either as GET or POST requests. Each control is exposed to the server using the name given in the HTML document.

Form Example

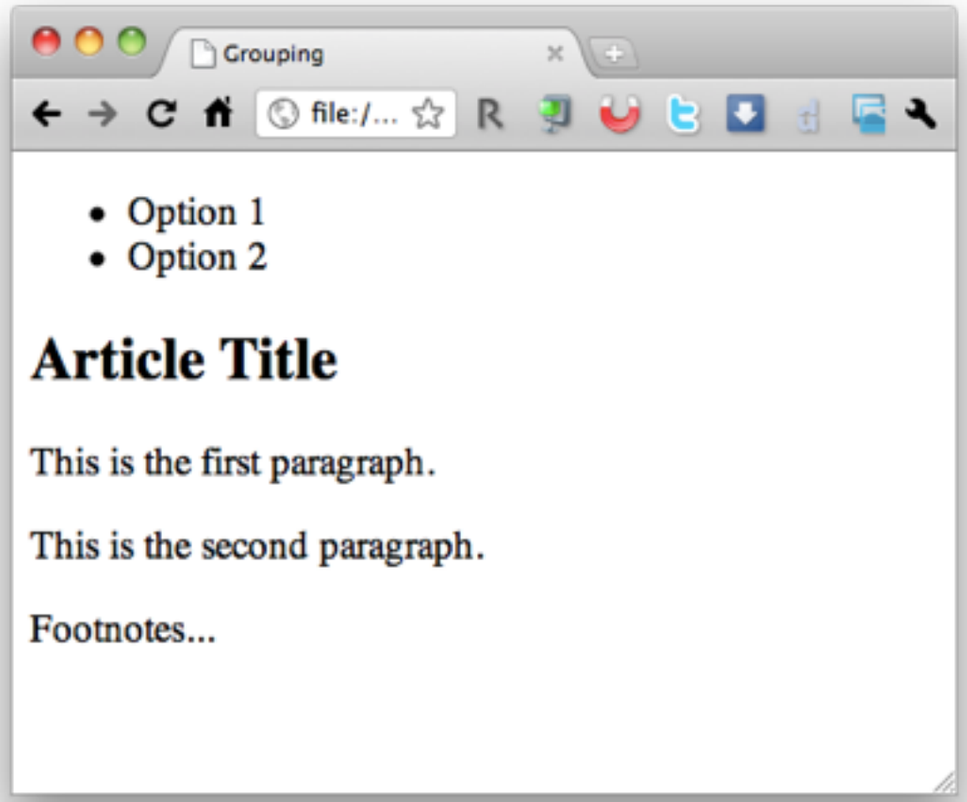
```
<form method="POST" action="script.php">

 <p><label>Customer name:<input></label></p>

 <p><label>Telephone:<input type=tel></label></p>

 <p><label>E-mail address: <input type=email></label></p>

 <fieldset>
  <legend> Pizza Size </legend>
  <p><label> <input type=radio name=size>Small</label></p>
  <p><label> <input type=radio name=size>Medium</label></p>
  <p><label> <input type=radio name=size>Large</label></p>
 </fieldset>

 <p><label>Delivery: <textarea></textarea></label></p>

 <p><button>Submit order</button></p>

</form>
```



Grouping Elements

- The and elements can be used to provide additional structure to HTML documents.
- These elements can be used to group other document elements.
- is an inline element and is a block-level element.
- Typically used in conjunction with stylesheets. These elements are invisible by default.

```
<div id="navigation">
<ul>
<li>Option 1</li>
<li>Option 2</li>
</ul>
</div>

<div id="article">
<h2>Article Title</h2>
<p>This is the first paragraph.</p>
<p>This is the second paragraph.</p>
</div>

<div id="footer">
<p>Footnotes...</p>
</div>
```



More

- There are many more HTML elements.
  We have just seen the most important ones.
- See References and Further Reading at the end.

XHTML

- In 1998 W3C decided abandon HTML development and focus on a XML-based equivalent, named XHTML.
- XHTML 1.0 was completed in 2000.
- W3C moved to XHTML 2.0, introducing several new features and less backward compatibility.
- Real world adoption of XHTML was small.
- A proposal to refocus on HTML was discarded, leading to outside development of HTML.

HTML5

HTML

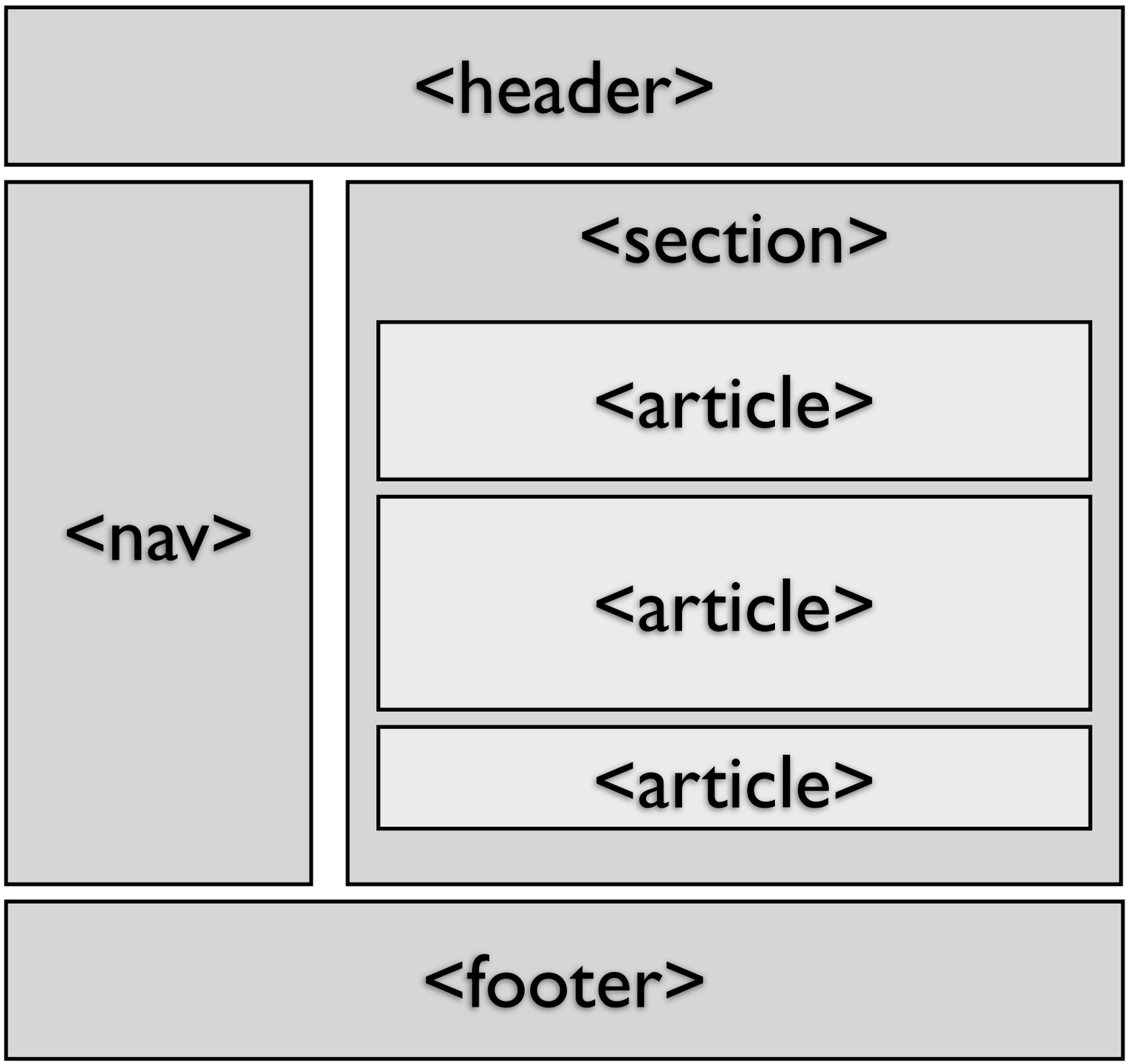
- HTML5 was initially developed outside W3C, by the WHATWG as “Web Applications 1.0”.

- The Web Hypertext Application Technology Working Group (WHATWG) was formed by W3C members that disagreed with W3C’s roadmap for HTML and XHTML.

- It was later adopted by W3C and is currently on the process of becoming a W3C Recommendation. It is still a work in progress.

New Elements

- Several new elements where introduced in HTML5 for better structure, with improved semantics.
- `<section>` — represents a generic document section.
- `<article>` — represents an independent piece of content of a document (e.g. blog entry).
- `<header>` — represents a group of introductory or navigational aids.
- `<footer>` — represents a footer for a section.
- `<nav>` — represents a section of the document intended for navigation.


<header>
<nav>
<section>
<article>
<article>
<article>
<footer>

Video and Audio

- Native support of multimedia content with the new <audio> and <video> elements.
- Web browsers that support these elements can play audio or video without external plugins (e.g. no Flash required).

Web Storage

- HTML5 includes an API for persistent data storage of key-value pairs in web clients.
- Unlike cookies, which can be accessed by both the server and the client, web storage can only be accessed in the client.
- Already supported in Safari, Chrome and Opera. Interesting for web applications (e.g. GMail).

Web Workers

- Web Workers allow an application to spawn tasks for the browser to work in the background.
- Allows for thread-like operations with message-passing as the coordination mechanism.

Much More

- Web Sockets
- Geolocation
- Notifications
- New Form Fields
- Canvas 2D & 3D
- SVG
- CSS3

References

- Using Google App Engine
  Charles Severance, O'Reilly (2009)
- HTML5 (Edition for Web Authors)
  http://dev.w3.org/html5/spec-author-view/
- HTML 4.01 Specification (1999)
  http://www.w3.org/TR/html401

Further Reading

- Dive Into HTML5, Mark Pilgrim
  http://diveintohtml5.org
- HTML5 — Edition for Web Developers
  http://developers.whatwg.org