

HTML

Developing Static Websites



Other precursors

- Janet Walker's (1985)
Symbolics Document Examiner
- Xerox (1985)
NoteCards a LISP based document system
- OWL-Guide (1986)
Application for the Macintosh
- Bill Atkinson (of Apple) (1987)
HyperCard clickable Hypermedia system

The Web

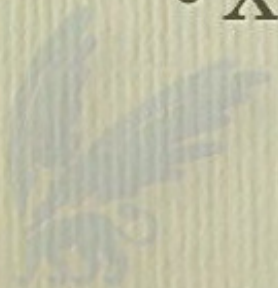


- Invented in 1989 by Tim Berners-Lee
 - Invented at CERN for documentation
 - Created HTTP, URLs and HTML (sort of, SGML as precursor)
 - 1991 Publicly available on the internet
 - He now directs W3C
 - Professor at MIT labs



SGML: Standard Generalized Markup Language

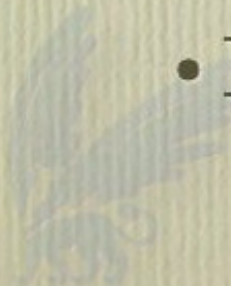
- ISO Standard for document markup
- Influenced the development of HTML
- XML is a successful profile of SGML
- HTML is NOT usually valid SGML
- XHTML unifies XML and HTML and is valid



NCSA Mosaic



- Developed in 1992 by Marc Andreessen a student at the National Center for Supercomputer Applications (USA)
- First cross platform browser
- First browser to support inline images
- Founded Mosaic Communications Corp
- Renamed to Netscape Later Due to Lawsuit

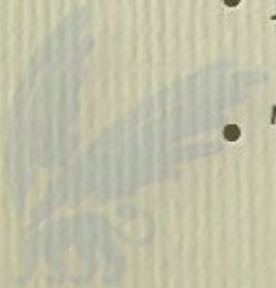


Microsoft & The Browser Wars

- 1996 75% of users use Netscape Navigator
- Internet Explorer released in 1995
- Microsoft bundles IE with Windows
(Later ruled in violation of Anti-Trust law!)
- Netscape is bought by AOL and open sources Navigator
- Navigator Rebranded Mozilla and then Forked to make Firefox
- Apple forks KDE browser for WebKit rendering engine and builds Safari
- Google builds Chrome using WebKit

Rendering Engines

- Three Major Client Rendering Engines
 - Internet Explorer (with several versions)
 - Gecko (In Firefox)
 - WebKit (In Safari, iPhone & now Chrome)
 - Opera (Hardly used, a distant 4th)
- All have their quirks making web development harder
 - Much better than it used to be!
 - Thanks to open standards and compliance tests!



HTML

Structuring the Web



HTML

- Language for building Hypertext documents
- Standardized by the World Wide Web Consortium (W3C) run by Tim Berners Lee
- Philosophy:
Author writes structure.
Browser decides layout



HTML: Recent Versions

- HTML 4.0: 1997
 - Introduce many new features
 - Deprecated lots of bad features
- HTML 4.01 1999
 - Cleanup of HTML 4.0
- XHTML 1.0 2000
 - XML conversion of HTML 4.01
- XHTML 1.1 2001
 - Modularize 1.0 and drop frames

HTML: Basic Syntax

- Defined using open and closing tags
 - Open: `<name>`
 - Close: `</name>`
- Text between the tags is the content the tags mark up
- Example:
 - `<h1>First Heading</h1>`
 - `<p>Paragraph about something...</p>`

Browsers Ignore

- Comments: `<!-- This is a comment -->`
- Unrecognized Tags
- White Space
 - Line Breaks
 - Tabs
 - Multiple Spaces



Simplest HTML Page

```
<html>  
  <body>  
    Hello World!  
  </body>  
</html>
```



Simple HTML Example

```
<html>
  <head>
    <title>Hello World!</title>
  </head>

  <body>
    <h1>Hello World!</h1>

    <p>Here is a paragraph. Would you like to visit my
      <a href="http://www.cs.vu.nl/~palmer">webpage</a>?
      Or would you rather look at my picture?
      
    </p>
  </body>

</html>
```



HTML Headings

- Comes from HTMLs history as a scientific documentation system
- Comes in 6 levels, `<h1>` to `<h6>`
- Browsers usually render them in different sizes



HTML Headings



HTML Headings

- Don't use them to resize text!
- Used by search engines to determine structure of your document
- Search Engine Optimization (SEO) is key to getting traffic to your site.



HTML Paragraphs

- Used to tell browsers where paragraphs start and end
- `<p>This is a paragraph</p>`
`<p>This is another one</p>`
- Required since browsers ignore multiple white spaces including line breaks!



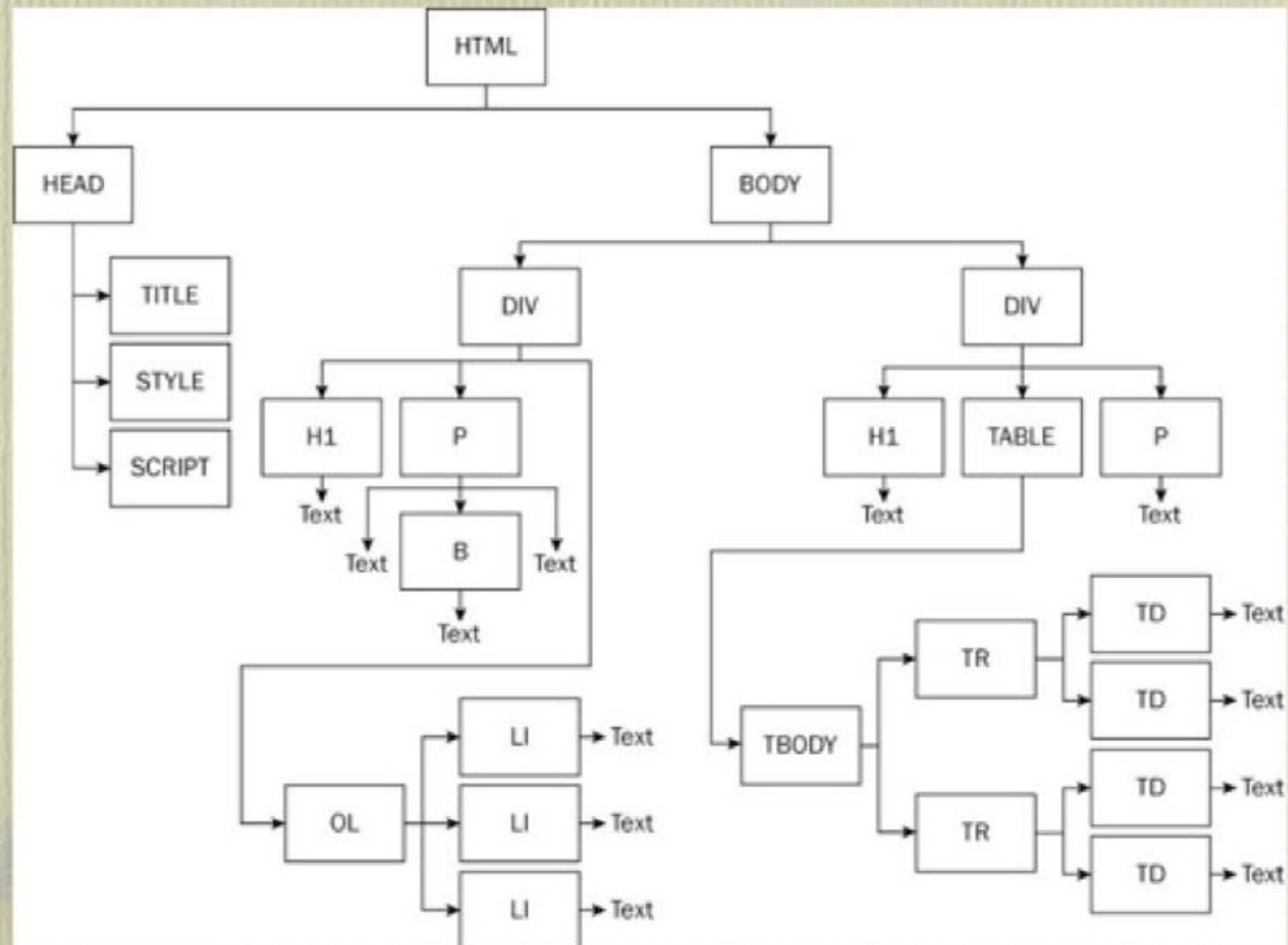
HTML Links

- The key to making HTML Hypertext!
- Also called “Anchors” hence the tag
- Specify a “hyper referenced” document
- `This is a link`
- The href inside the tag is an attribute

HTML Elements

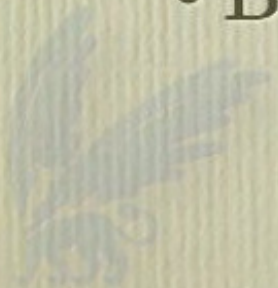
- Elements run from open tag to close tag
- Element content is everything between
- Some elements have no content
- Some have no close tag in HTML or are self closed in XHTML:
e.g. `
` or `
` instead of `
</br>`
- All elements can have attributes
- Most can be nested inside each other

HTML Entities: Nesting



HTML Entities

- Tags are NOT case sensitive
- W3C recommends lower case tags
- HTML 5 demands lower case tags
- Browsers will understand uppercase
- But maybe not forever!



HTML Entities

- Browsers try to interpret poorly formed HTML

- `<p>` These are bad paragraphs

- `</p>` Because they are not closed

- Remember to close your tags!

- XML is strict about this and will not validate if tags are not closed.

HTML Attributes

- Attributes provide additional information about elements
- Attributes always appear in start tags
- Attributes are name value pairs
- Values are usually quoted (“ or ‘)
- attribute=”value” attribute=‘value’
 - Watch out for “ in value!
 - Either use ‘a”b’ or “a\”b”

HTML Attributes

- Some attributes are available on most tags:
 - class (more on this in CSS section)
 - id (more on this in CSS and Javascript)
 - style (Again a CSS attribute)
 - title (For tooltip text)

• Other attributes are element specific

HTML Character Entities

- Some characters have to be escaped
- Expressed either as `&name;` or `&#num;`

Char.	Name	Number	Meaning
&	<code>&amp;</code>	<code>&#38;</code>	Ampersand
<	<code>&lt;</code>	<code>&#60;</code>	Less Than
>	<code>&gt;</code>	<code>&#62;</code>	Greater Than
€	<code>&euro;</code>	<code>&#8634;</code>	Euro
©	<code>&copy;</code>	<code>&#169</code>	Copyright
Space	<code>&nbsp;</code>	<code>&#160;</code>	Non-breaking space

HTML Document Structure

```
<!DOCTYPE ...>  
<html>  
  <head>  
    [HEAD ELEMENTS]  
  </head>  
  
  <body>  
    [BODY ELEMENTS]  
  </body>  
</html>
```



HTML Doctype

- Defines what tags are allowed in your document
- Used by validators to tell you if you are breaking the rules
- Comes before the `<html>` tag



Common Doctypes

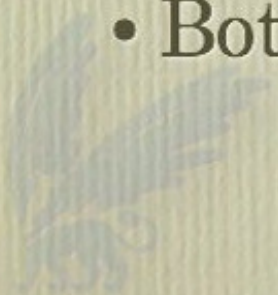
- HTML 4.01 Strict:
`<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">`
- HTML 4.01 Transitional:
`<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML Transitional
4.01//EN" "http://www.w3.org/TR/html4/loose.dtd">`
- XHTML 1.0 Strict:
`<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">`
- XHTML 1.0 Transitional:
`<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0
Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-trasitional.dtd">`

HTML Head Elements

- `<title>` Used to specify the title of the window!
Very important to search engines
- `<script>` Used to link in external Javascript
- `<style>` Used to link in external CSS
- `<meta>` Used to specify metadata. Some are important to search engines.
- `<link>` A resource reference. Subsumes `<script>` and `<style>`
- `<base>` The base URL for relative URLs

HTML Meta Elements

- Meta Elements For Search Engines
 - `<meta name="keywords" content="some,keywords,for,this,site" />`
 - `<meta name="description" content="Some short description of the page. Often shows as the summary of the page on search engines." />`
- Both should be unique per page for best rank!



HTML Body Elements

- Too many to tell you all of them!



All About Anchors

- Anchor tags make HTML Hypertext
- Defines both links and anchors
- Links point to pages, anchors or both
 - `Link Text`
 - `Link Text`
 - `Link Text`
- Anchors specify places within a document
 - `Anchor Text`

Images

- Images are specified with the `` tag
- The `src` attribute gives the URL
- The `alt` attribute specifies alternative text for text only browsers (or the blind!)
- Each image must be fetched **AFTER** the HTML is fetched and parsed by the browser!

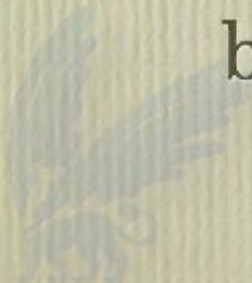


Image Formats

- GIF (Graphic Interchange Format)
 - 8-bit color (256 different colors)
- JPEG (Joint Photographic Experts Group)
 - 24-bit color (16 million different colors)
- PNG (Portable Network Graphics)
 - 24 or 32 bit color + transparency
 - Relatively new and will replace above

Formatting Tags

- Use of these is more uncommon now due to the rise of CSS for formatting text

Tag	Description
<code></code>	bold
<code><big></code>	big text
<code></code>	emphasized (usually italics)
<code><i></code>	italic
<code><small></code>	small text
<code></code>	stronger text (usually bold)
<code><sub></code>	subscript
<code><sup></code>	superscript
<code><ins></code>	inserted text
<code></code>	deleted text
<code><u></code>	underline (Deprecated)

HTML Tables

- Used to display tabular data
- Abused before the advent of CSS to place things on the page
 - Required the use of 1 pixel transparent images!



HTML Tables by Example

```
<table border="1">  
  <tr>  
    <td>row 1, cell 1</td>  
    <td>row 1, cell 2</td>  
  </tr>  
  <tr>  
    <td>row 2, cell 1</td>  
    <td>row 2, cell 2</td>  
  </tr>  
</table>
```

row 1, cell 1	row 1, cell 2
row 2, cell 1	row 2, cell 2

HTML Tables: Headings

```
<table border="1">  
  <tr>  
    <th>Heading</th>  
    <th>Another Heading</th>  
  </tr>  
  <tr>  
    <td>row 1, cell 1</td>  
    <td>row 1, cell 2</td>  
  </tr>  
</table>
```

Heading	Another Heading
row 2, cell 1	row 2, cell 2

HTML Tables: Empty Cells

- Some browsers do not display borders on empty cells!
- Add a ` ` to an empty cell to get a border in all browsers



HTML Lists

- Comes in 3 flavors
 - Ordered `` (Usually with numbers)
 - Unordered `` (Usually with bullets)
 - Definition `<dl>` (Usually inset)



HTML Lists: Examples

```
<ul>
```

```
  <li>Coffee</li>
```

```
  <li>Milk</li>
```

```
</ul>
```

- Coffee

- Milk



HTML Lists: Examples

```
<ol>
```

```
  <li>Coffee</li>
```

```
  <li>Milk</li>
```

```
</ol>
```

1. Coffee

2. Milk



HTML Forms

- Allows users to send information to a server
- `<form> FORM ELEMENTS </form>`
- Usually uses POST or GET methods
 - POST for actions that will modify server state
 - GET for actions that do not modify server state



Form Elements

- `<input>` tag is used to enter for data
 - text
 - radio
 - checkbox
 - button
 - hidden
- `<textarea>` a multiline text input
- `<submit>` submit button (also `<button>`)
- `<select>` `<optgroup>` `<option>` for menus

Form Example

```
<form>
```

```
Name: <input type="text" name="name">
```

```
<br/>
```

```
Male: <input type="radio" name="Sex" value="Male" checked="checked" />
```

```
<br/>
```

```
Female: <input type="radio" name="Sex" value="Female" />
```

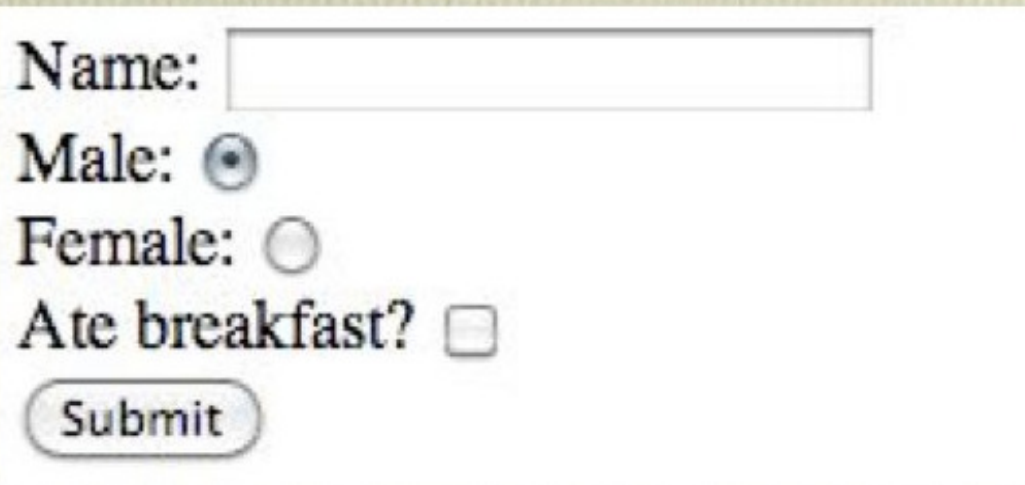
```
<br/>
```

```
Ate breakfast? <input type="checkbox" name="atebreakfast" />
```

```
<br/>
```

```
<input type="submit" value="Submit" />
```

```
</form>
```



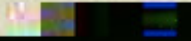
Name:

Male:

Female:

Ate breakfast?

Form Method and Action



Form Data Encoding

- enctype attribute specifies MIME content type for POST data sent to the server
- Default value is:
"application/x-www-form-urlencoded".
- "multipart/form-data" required for forms with `<input type="file" />`
 - accept attribute can specify acceptable mime file types so browsers can filter at file selection time.



Validating HTML

- <http://validator.w3.org>



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

Jump To: [Congratulations](#) · [Icons](#) [Source Listing](#)

This document was successfully checked as HTML 4.01 Transitional!

Result:	Passed	
Address :	<input type="text" value="http://www.cs.vu.nl/~palmer/"/>	
Encoding :	utf-8	<input type="text" value="(detect automatically)"/>
Doctype :	HTML 4.01 Transitional	<input type="text" value="(detect automatically)"/>
Root Element:	HTML	



The W3C CSS validator is developed with assistance from the Mozilla Foundation, and supported by community donations.

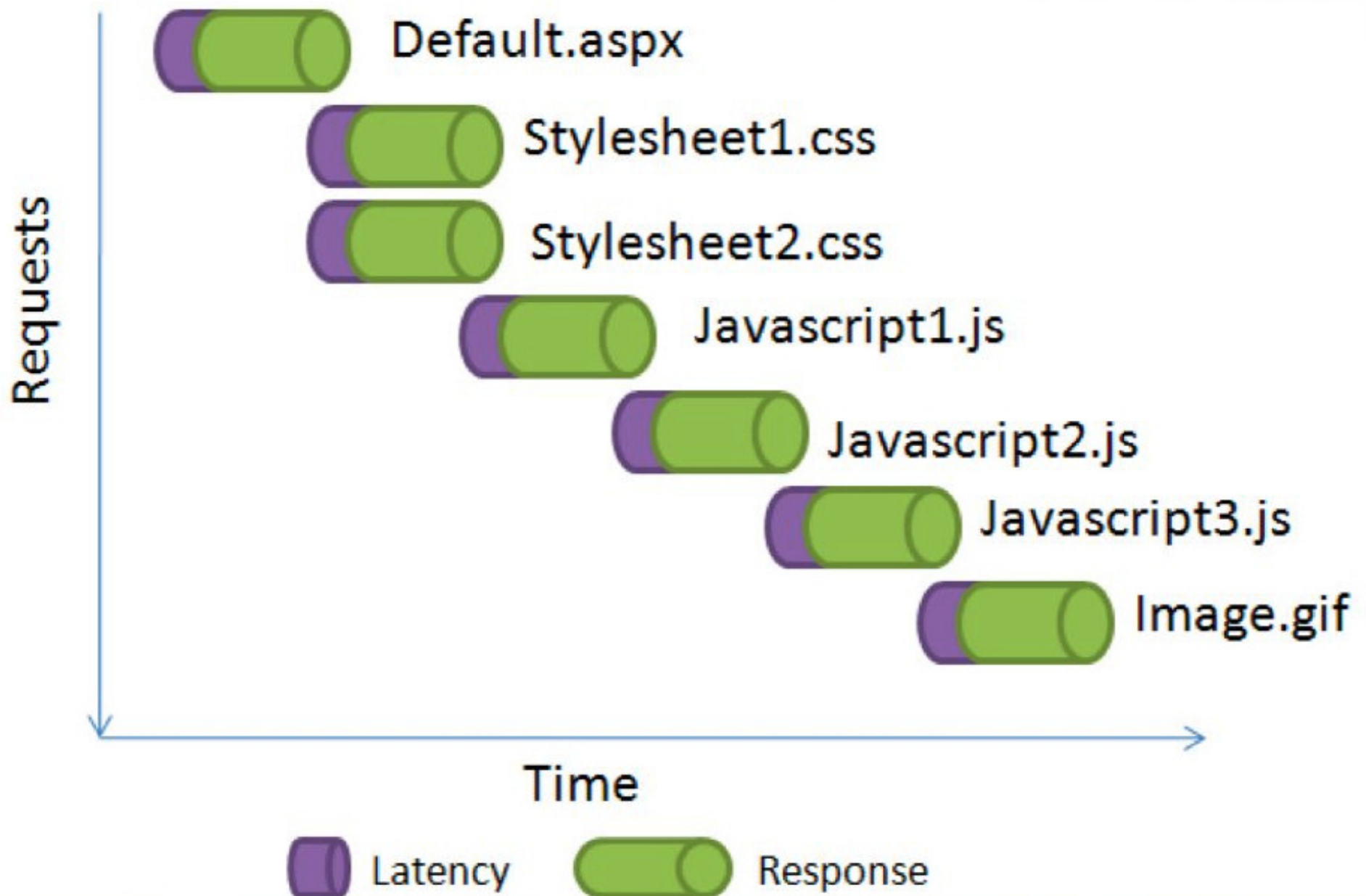
[Donate](#) and help us build better tools for a better web.

HTML

That's all!



A Note on Loading



Online References

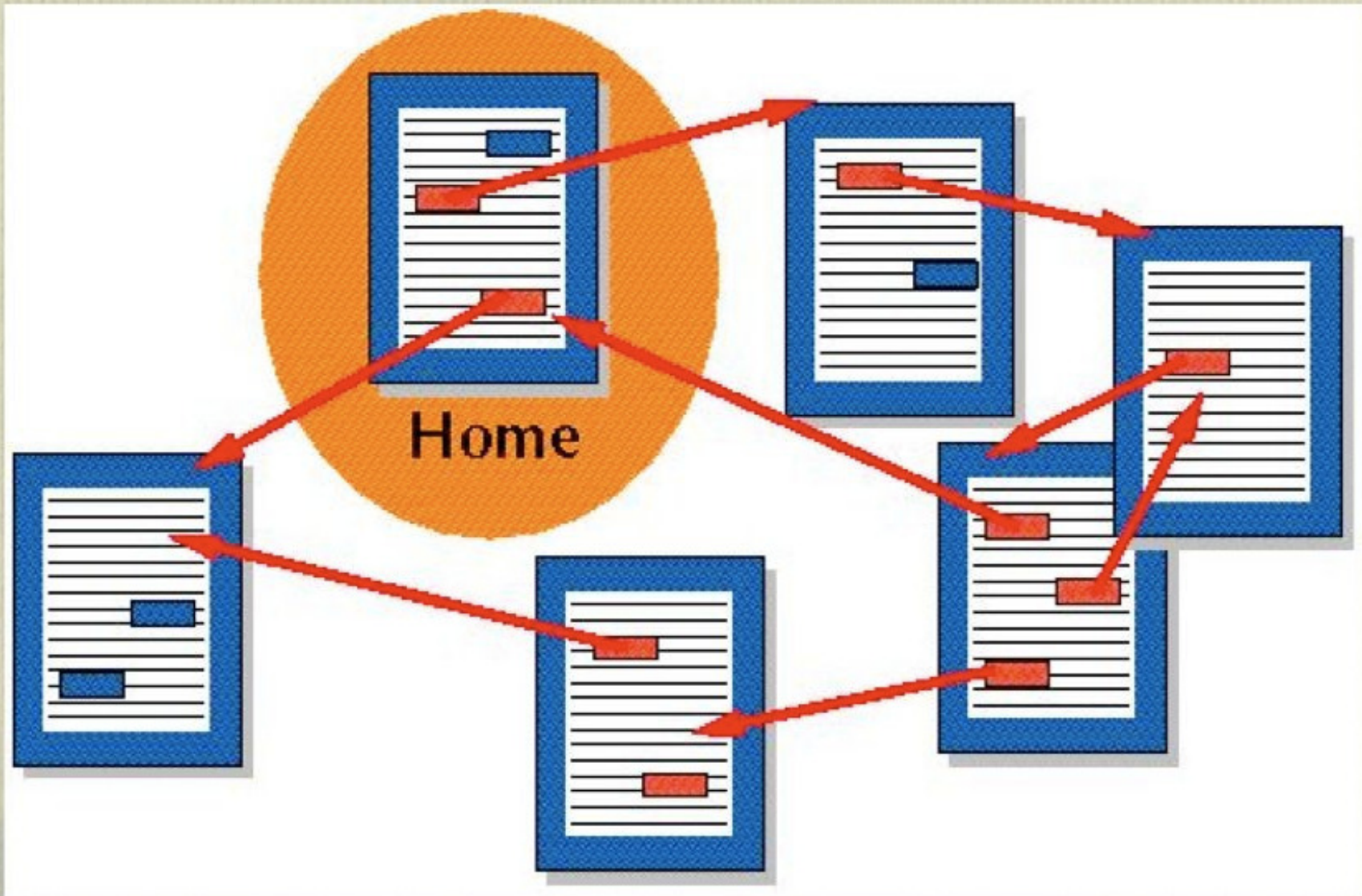
- <http://www.w3schools.com>
- <http://htmlhelp.com/reference/html40>



Important Concepts



Hypertext



HTML Document Structure

```
<!DOCTYPE ...>  
<html>  
  <head>  
    [HEAD ELEMENTS]  
  </head>  
  
  <body>  
    [BODY ELEMENTS]  
  </body>  
</html>
```

