Introduction to the Web

PGCertInfoTech
Programming with Web Technologies

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Auckland
ICT Graduate School

Course structure

1st half of the course : client-side development

- Focuses on what happens in the user's browser

Implementation using:

- HTML to describe page content
- CSS to define how the content is presented
- JavaScript to implement functional features

Course structure

2nd half of the course: server-side development

- Focuses on what happens on the web server

Implementation using:

- Sevlets + JSP to implement functional features,
- dynamically generate web page content,
- track user activity,
- integrate with Java DB databases

Course Assessment

5 practical tests:

- Alternating Fridays mornings (with *Programming for Industry)*

Assignment in final two weeks of the course:

- Develop your own Blogging software
- Pair programming
- Joint assessment with *Programming for Industry*

A long long time ago …

1962: Joseph Licklider at MIT envisioned and wrote about the concept of a “Galactic Network” of computers

1969: ARPANET goes Live

1972: The concept of open internet is floated

1982: ARPANET becomes the internet (a network of networks)

70s-80s: FTP (File Transfer Protocol)

: Archie Search Engine

: Gopher information retrieval system

A long long time ago …

1980: Tim Berners-Lee at CERN (European Centre for Nuclear Research)creates the first web prototype called ENQUIRE.

Loses everything in a hard drive crash

(Backup your work!)

1991: Tim Berners-Lee started again and this time released his work as World Wide Web

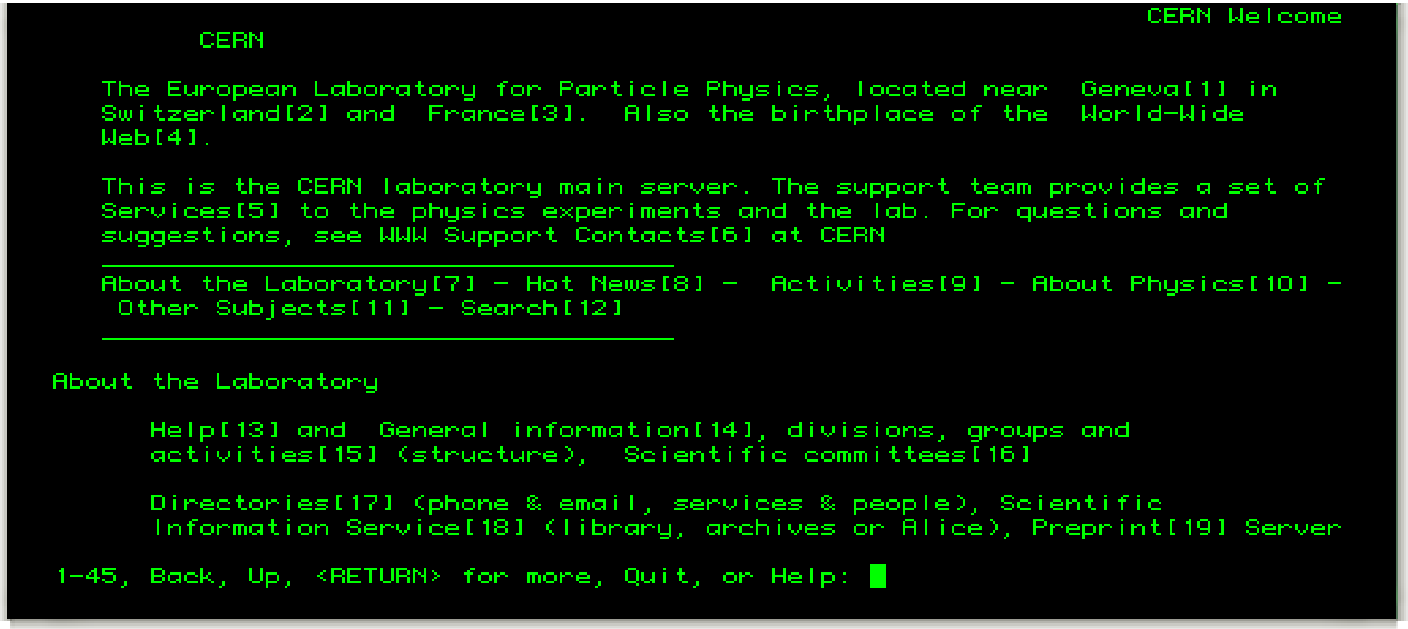
The first web site domain was:

- info.cern.ch

The first web page address was:

- http://info.cern.ch/hypertext/WWW/TheProject.html

First: A text only browser



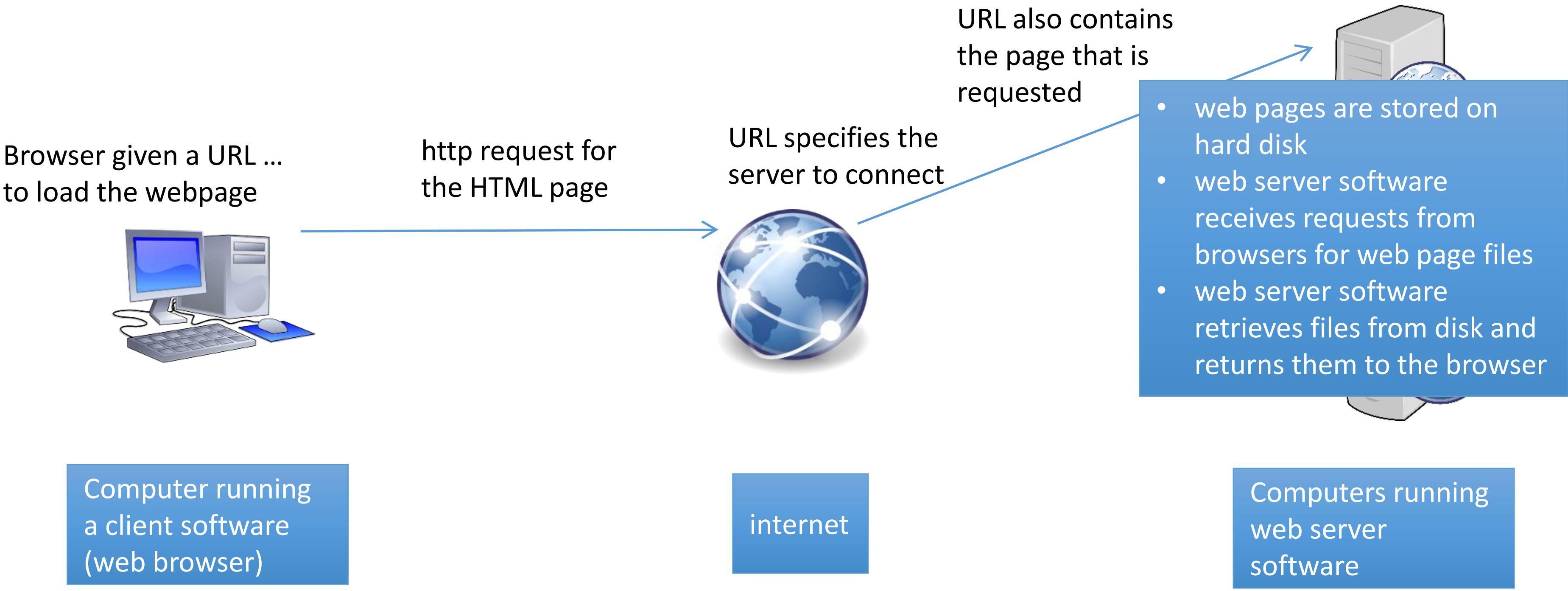
```
                                                                  CERN Welcome
           CERN

     The European Laboratory for Particle Physics, located near  Geneva[1] in
     Switzerland[2] and  France[3].  Also the birthplace of the  World-Wide
     Web[4].

     This is the CERN laboratory main server. The support team provides a set of
     Services[5] to the physics experiments and the lab. For questions and
     suggestions, see WWW Support Contacts[6] at CERN
     ___________________________________
     About the Laboratory[7] - Hot News[8] -  Activities[9] - About Physics[10] -
      Other Subjects[11] - Search[12]
     ___________________________________

  About the Laboratory

        Help[13] and  General information[14], divisions, groups and
        activities[15] (structure),  Scientific committees[16]

        Directories[17] (phone & email, services & people), Scientific
        Information Service[18] (library, archives or Alice), Preprint[19] Server

  1-45, Back, Up, <RETURN> for more, Quit, or Help: █
```

Not so long ago …

- 1993:NCSA (University of Illinois) released the first graphical web browser called Mosaic
- Browser wars started with IE, Netscape Navigator and Opera
- The Web …
  - … is now too big to measure (pages/sites/servers)
  - At least 250,000,000 sites, billions of pages
  - We'll be learning how to add more!
  - Although web technologies have developed rapidly, the basic principles of the web haven't changed since Berners-Lee created the first site

The web

- High-level view of the web



Browser given a URL … to load the webpage

http request for the HTML page

URL specifies the server to connect

URL also contains the page that is requested

- web pages are stored on hard disk
- web server software receives requests from browsers for web page files
- web server software retrieves files from disk and returns them to the browser

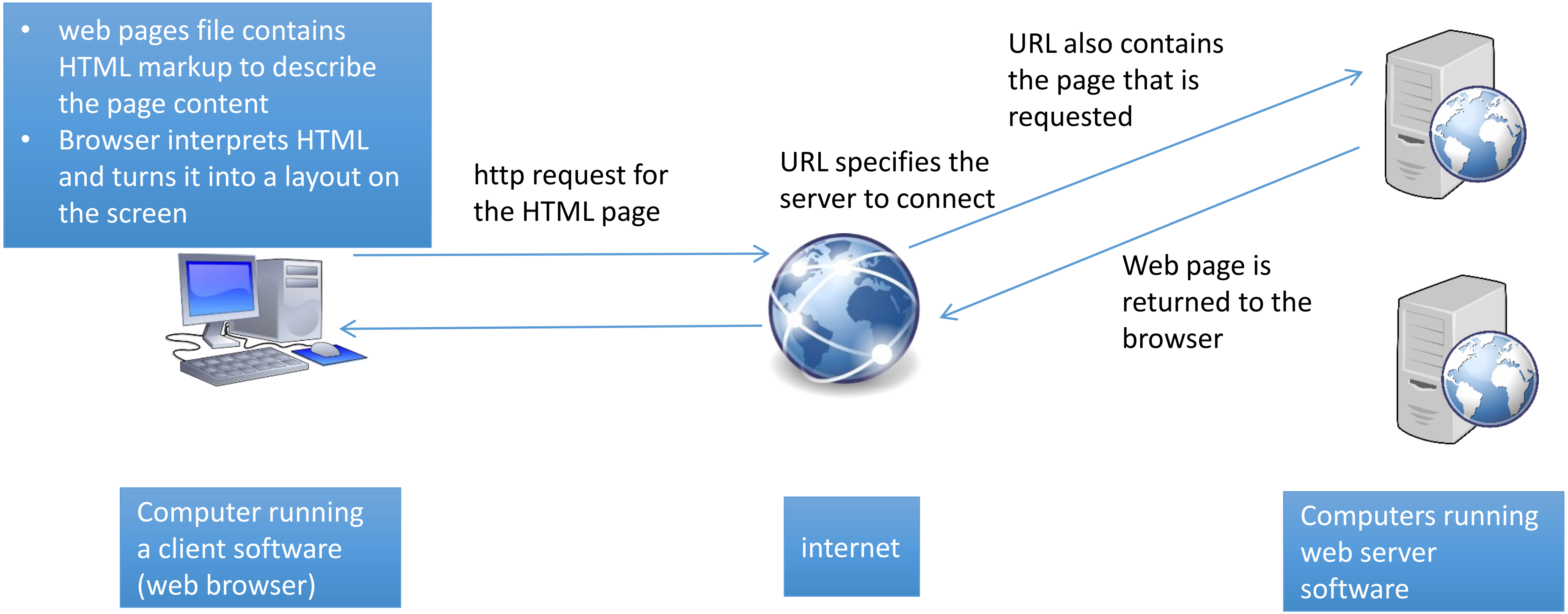
Computer running a client software (web browser)

internet

Computers running web server software

The web

- High-level view of the web



- web pages file contains HTML markup to describe the page content
- Browser interprets HTML and turns it into a layout on the screen

http request for the HTML page

URL specifies the server to connect

URL also contains the page that is requested

Web page is returned to the browser

Computer running a client software (web browser)

internet

Computers running web server software

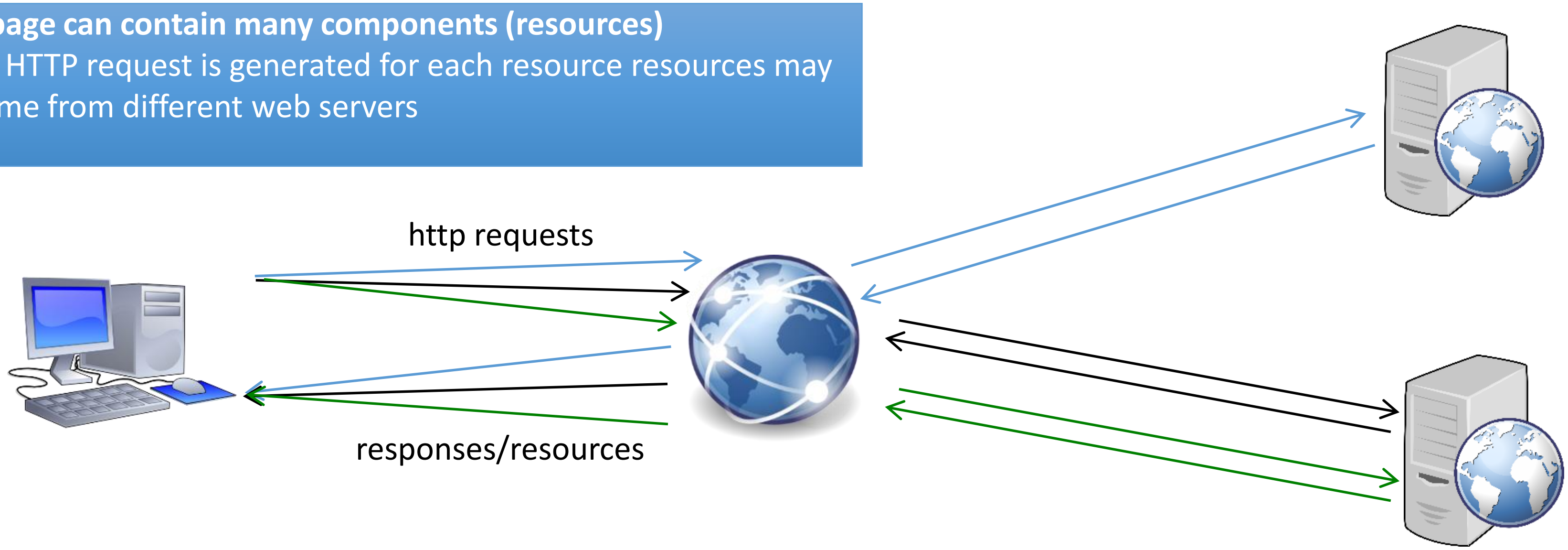
The web

- Actually a web page normally consists of many separate resources (files)
  - HTML markup
  - Images
  - JavaScript code
  - Presentation details using CSS
  - Movies
  - Audio
  - etc.
- And the resources that make up a page can come from different web servers
- So the browser needs to request each resource and manage the responses and construct the final result to display to the user

The web

a page can contain many components (resources)

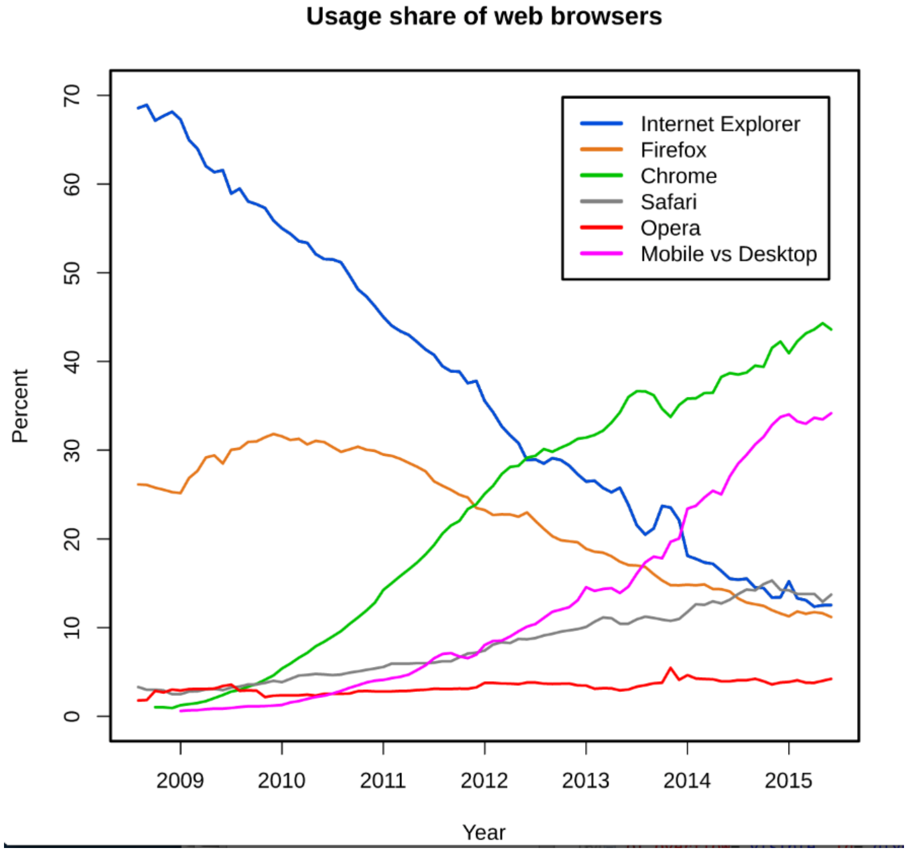
an HTTP request is generated for each resource resources may come from different web servers



Browsers

- Current browsers provide sophisticated functionality and GUIs
- Their core capabilities are essentially the same
  - issue requests over the network for files identified by a URL, and keep track of multiple concurrent requests and the responses
  - parse HTML describing page structure and build an internal representation (the Document Object Model)
  - generate graphical components corresponding to the document model, lay them out sensibly and render them in the window
  - apply default or specified styles (CSS) to the content to affect the look and feel
  - manage user interaction with page content/browser GUI
  - interpret and execute JavaScript code included in the page
  - handle security, cookies, caching

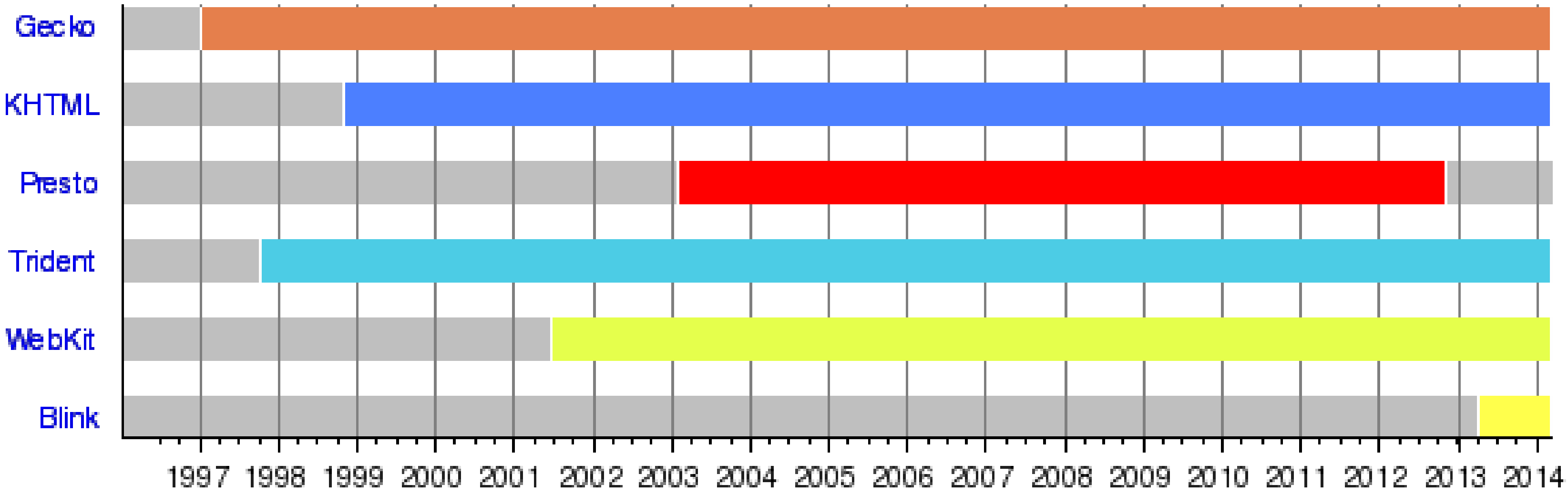
Browser usage share



By arichnad, Daniel.Cardenas, Litehacker - Own work. Licensed under CC BY 3.0 via Wikimedia Commons From https://en.wikipedia.org/wiki/Usage_share_of_web_browsers

Browsers

One significant point of difference between browsers is the **rendering engine** – the code that turns HTML markup into displayed web pages



https://en.wikipedia.org/wiki/Web_browser_engine

Gecko: Netscape, Firefox
KHTML: KDE Konqueror
Presto : Opera

Trident: Internet Explorer
WebKit: Chrome, Safari
Blink :Chrome, Opera

Browsers

Another significant point of difference between browsers is the **JavaScript engine**—the code that interprets and executes Javascript source code:

SpiderMonkey:	from Mozilla, used in Firefox
V8:	from Google, used in Chrome/Opera
Chakra:	from Microsoft, used in IE9
Nitro:	from WebKit group, used in Safari

Browser developers compete to create the fastest rendering and JavaScript engines:

- Better user experience, more users for the browser

Browsers

- Web browsers used to be monolithic
- Splitting into components (layout engine, JS engine etc.) means they can be deployed in other apps
- For example:
  - HTML formatted email
  - Help systems in software applications
  - eBooks

HTML

- **H**yper**T**ext **M**arkup **L**anguage, used to describe the content of documents on the web (web pages)
- The first version of HTML
  - text-only documents
  - define parts of a document content semantically (e.g. a title, paragraphs, headings, lists)
  - define links as *anchors*
- Now it is more complex – more than just text in web pages,
  - more complex document structures
  - enhanced user interaction with pages

HTML

Different versions as standards evolved:

- Current version is HTML5
  - https://en.wikipedia.org/wiki/HTML
- Standards are developed by W3C (World Wide Web Consortium)
  - http://w3.org
  - International body that defines HTML
  - And other standards (e.g., CSS, DOM, HTTP, XML, SVG etc.)

HTML

- An HTML document consists of two types of information
  - The content of the document
  - Markup that describes the content

- Markup is in the form of element tags
  - The content is surrounded by a start tag and an end tag

HTML example

A simple HTML Document:

```
<!DOCTYPE html>
<html>
    <head>
        <meta charset=“UTF-8”>
        <title>
          My First HTML page
        </title>
    </head>
    <body>
      <p>This is the content of my first HTML page</p>
    </body>
</html>
```

The <!DOCTYPE declaration is the very first thing in the HTML document. It is an instruction to the web browser about what version of HTML the page was written in

Start and end points of the HTML document

HTML example

A simple HTML Document:

```
<!DOCTYPE html>
<html>
    <head>
        <meta charset=“UTF-8”>
        <title>
          My First HTML page
        </title>
    </head>
    <body>
      <p>This is the content of my first HTML page</p>
    </body>
</html>
```

start and end points of information about the content

start and end points of the document content

HTML example

A simple HTML Document:

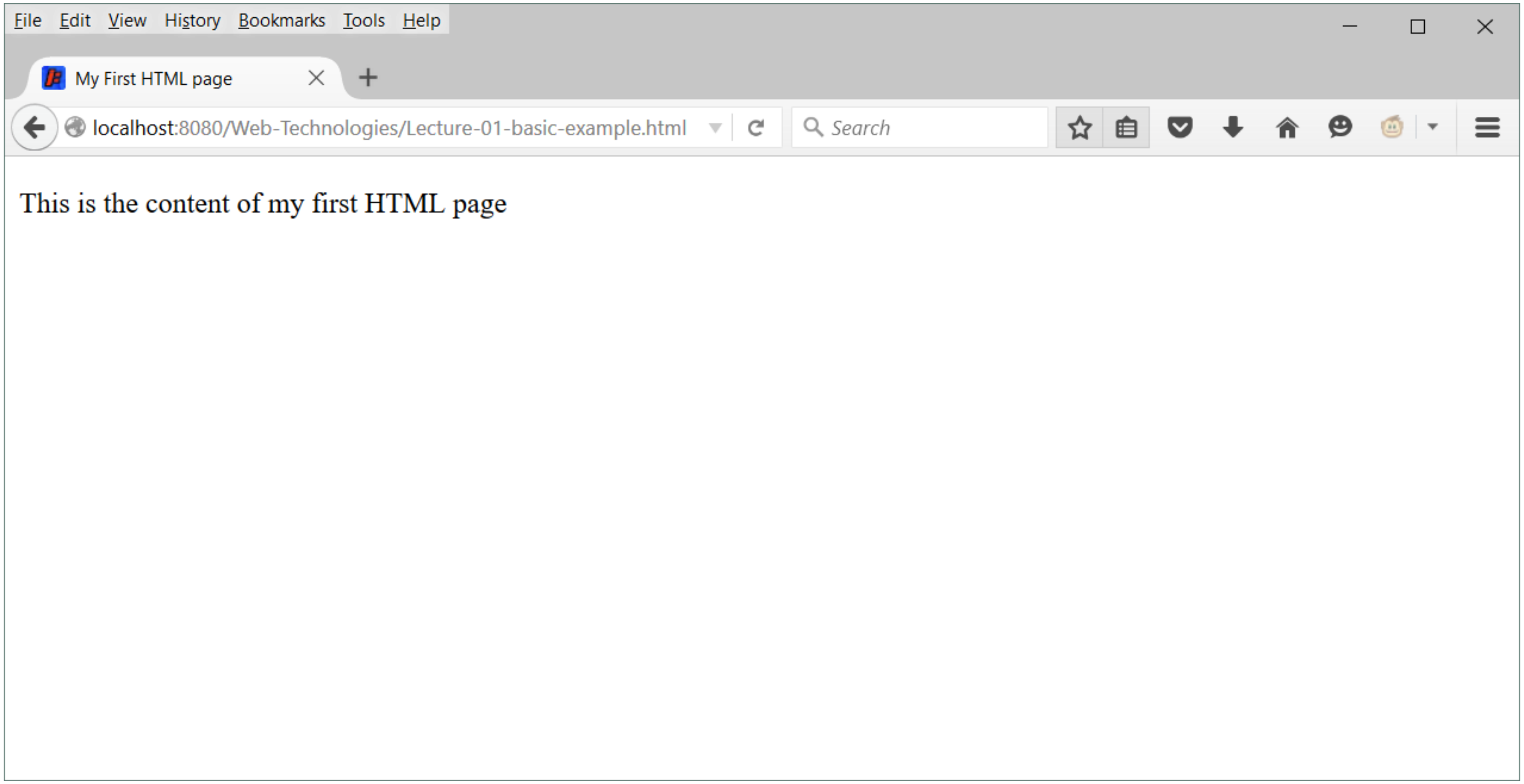
```
<!DOCTYPE html>
<html>
    <head>
        <meta charset=“UTF-8”>
        <title>
          My First HTML page
        </title>
    </head>
    <body>
      <p>This is the content of my first HTML page</p>
    </body>
</html>
```

contains information about the document such as page description, keywords, author etc. Information in the meta tag is not displayed in the browser

start and end points of the text that is the title of the document

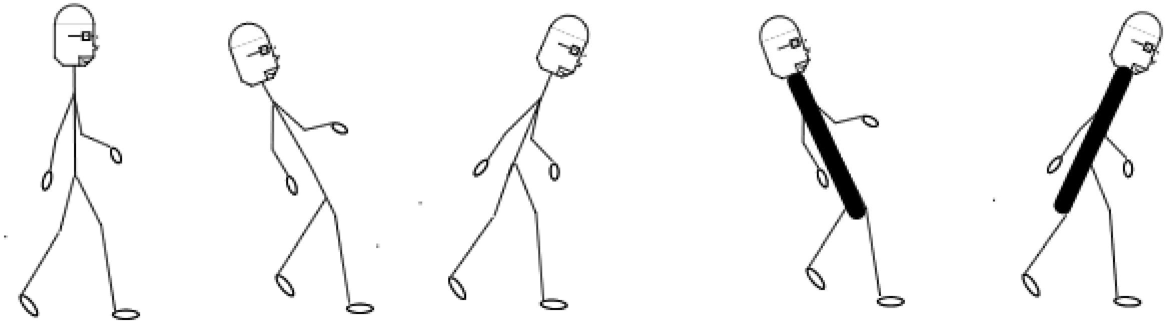
start and end points of a paragraph

HTML example: Rendered


File Edit View History Bookmarks Tools Help
My First HTML page
localhost:8080/Web-Technologies/Lecture-01-basic-example.html
Search

This is the content of my first HTML page

Backslash forward slash



HTML Example

HTML tags occur within angle brackets < >

If the start tag for an element is <foo> then the corresponding end tag is </foo>

Tags should be correctly nested

Whitespace has no meaning most of the time so the previous code is equivalent to this

```
<!DOCTYPE html><html><head><title>My First HTML
page</title></head><body><p>This is the content of my
first HTML page</p></body></html>
```

Code like this is impossible to debug and maintain

- indent your HTML sensibly
  (Many text editors and IDEs will do this for you)

Bad HTML

- When the web started to become popular with the general public, people with no experience/training began to create HTML web pages
- A good illustration of what the web used to be (bad HTML)
  - http://www.angelfire.com/super/badwebs/
- Modern browsers are resilient to some HTML errors
  - Even with an error something would be displayed

Bad HTML

```
<head>
</head>
<title>
        My First HTML page
</title>
        This is the content of my first HTML page
</body>
</html>
```

- No HTML version statement
- No <html> tag
- Nothing inside the <head> tag
- <title> element in the wrong place
- No meta tag
- No <body> tag

Good HTML

- We don't know what errors browsers can deal with or how they will deal with them
- **But, we do know that browsers deal correctly and predictably with valid HTML**
- So we should make sure that the HTML that we write is valid
- Validation tools
  - some IDEs have built-in validation
  - there are web sites that validate HTML
    - validator.waikato.ac.nz
    - https://validator.w3.org/nu/

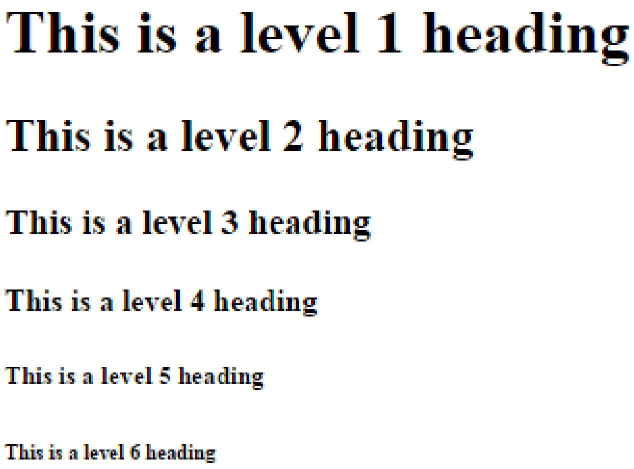
Two types of elements in HTML

- There are two types of elements which can be used for content in HTML
  - Block level elements
    - A block-level element always starts on a new line and takes up the full width available (stretches out to the left and right as far as it can)
    - e.g. <p>, <h1>….<h6>, lists, <table>, <form> etc.
    - https://developer.mozilla.org/en-US/docs/Web/HTML/Block-level_elements
  - Inline elements
    - An inline element does not start on a new line and only takes up as much width as necessary.
    - e.g. <a>, <img> etc.
    - https://developer.mozilla.org/en-US/docs/Web/HTML/Inline_elemente

HTML elements

Headings:

- <h1> This is a level 1 heading </h1>
- <h2> This is a level 2 heading</h2>
- <h3> This is a level 3 heading</h3>
- <h4> This is a level 4 heading</h4>
- <h5> This is a level 5 heading</h5>
- <h6> This is a level 6 heading</h6>



This is a level 1 heading

This is a level 2 heading

This is a level 3 heading

This is a level 4 heading

This is a level 5 heading

This is a level 6 heading

HTML elements

Text:

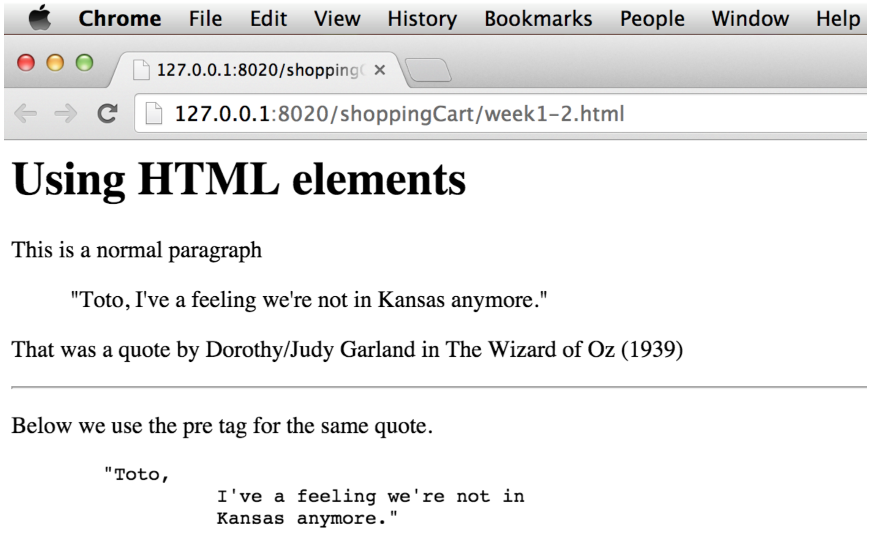
- Paragraphs: <p>Some text</p>
- Bold text: <b>Some text</b>
- Italics text: <i>Some text</i>
- Line break: <br> also sometimes written <br />
- Horizontal line: <hr> also sometimes written <hr />

Some other elements:

- <blockquote></blockquote> : indented quotation
- <pre> </pre> : pre-formatted text
- <!-- --> : comment

Mixing it up

```
<body>
   <h1> Using HTML elements</h1>
   <p>This is a normal paragraph</p>
   <blockquote>
   "Toto, I've a feeling we're not in Kansas anymore."
   </blockquote>
   <p> That was a quote by Dorothy/Judy Garland in The Wizard of Oz
(1939) </p>
   <hr />
   <p> Below we use the pre tag for the same quote. </p>

   <pre>       "Toto,
                     I've a feeling we're not in
                     Kansas anymore."
   </pre>
   <!-- This is a comment and would not be displayed in the browser -->
</body>
```



Using HTML elements

This is a normal paragraph

> "Toto, I've a feeling we're not in Kansas anymore."

That was a quote by Dorothy/Judy Garland in The Wizard of Oz (1939)

Below we use the pre tag for the same quote.

```
        "Toto,
                  I've a feeling we're not in
                  Kansas anymore."
```

Further elements

- Links (<a>)
- Lists (<ol>, <ul>, <dl>, <li>)
- Tables (<table>, <tr>, <td>, etc.)
- Blocks (<div>)
- Images and Video (<img> and <video>)
- Forms (<form>)
- Hyperlinks (<a>)

Will meet these later on. For now we will concentrate on practical work centred on paragraphs of text

HTML elements with attributes

- Many elements (<img>, <form>, form controls, <a>) require us to provide additional information for them to be useful

- Additional information is provided via element **attributes** which:
  - appear inside the start tag of the element
  - are given as name=value pairs
  - and can be a set of multiple attributes

```
<img src="http://www.waikato.ac.nz/images/coa_red.gif"/>
```

Tags, Attributes and Values

- Tags :- define what needs to be done
- Attributes: - modify the way tags are used
- Values:- define how the modification will be done

- Spaces
  - Absolutely necessary between a tag and an attribute
  - Not allowed between attributes and values