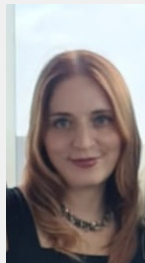# COM1008 Web and Internet Technologies

Lecture 1: Introduction to WWW

Dr Steve Maddock
s.maddock@sheffield.ac.uk
https://staffwww.dcs.shef.ac.uk/people/S.Maddock

Dr Vitaveska Lanfranchi
v.lanfranchi@sheffield.ac.uk
https://www.sheffield.ac.uk/dcs/people/academic/vitaveska-lanfranchi

Learning Objectives

- During this lesson you will learn about
    - Introduction to the Internet and WWW
    - Introduction to network protocols
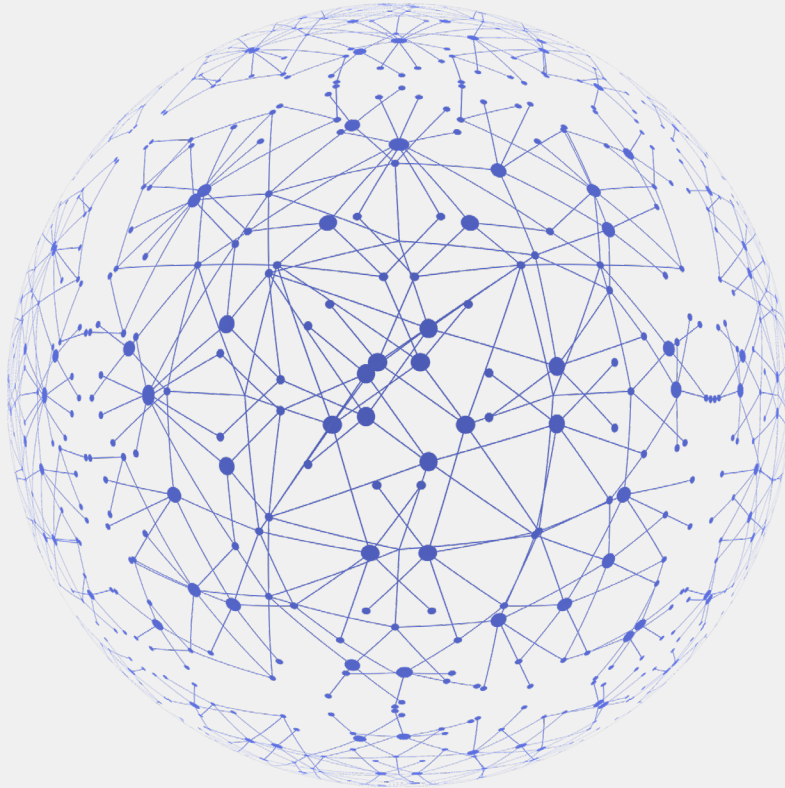   - Introduction to website development process

A BRIEF HISTORY OF WWW
&
INTERNET PROTOCOLS
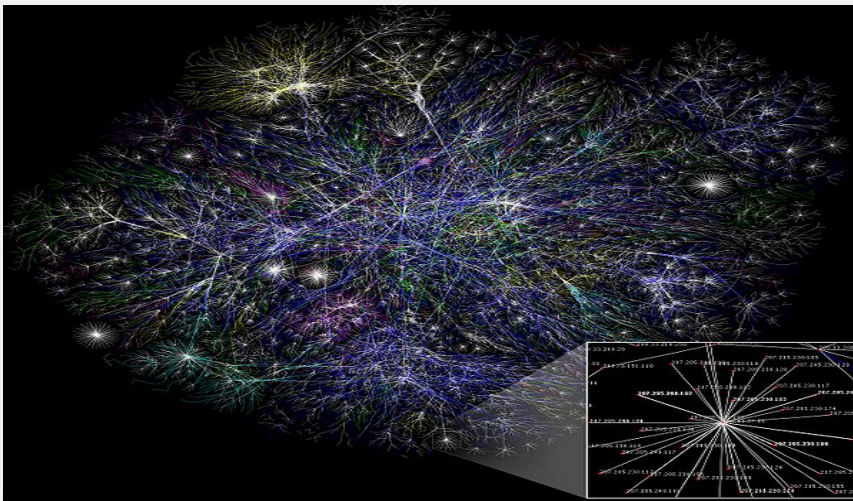
What is the Internet?

- A global network of interconnected networks and devices

- Grows exponentially

- Provides information and communication resources

- Uses the Internet protocol suite (TCP/IP) to communicate between networks and devices

# The Internet and the World Wide Web *They are **not** the same thing*

- The "Net"

  - Global interconnected collection of computer networks

  - Uses standard communications protocols

- The "Web"

  - An information sharing model built on top of the Internet

  - One of the services communicated via the Internet

- (Other services: e-mail, ftp, VOIP, instant messaging, …)

Partial map of the Internet @ Jan 15, 2005 based on data at opte.org
[showing 30% of "Class C" networks]
http://en.wikipedia.org/wiki/File:Internet_map_1024.jpg
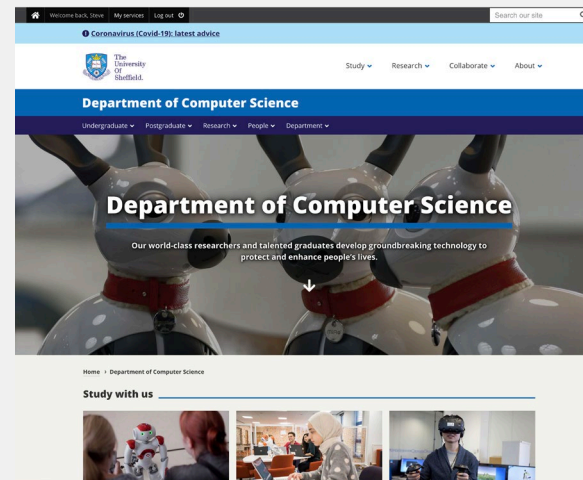

Department website, Sep 2021
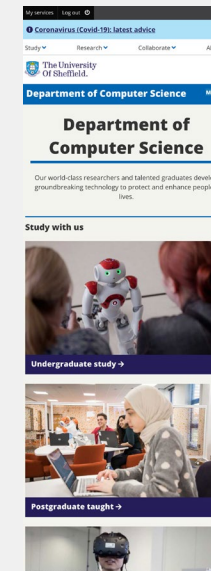
Precursors

- Whilst everyone recognises Tim Berners Lee with inventing what we know nowadays as Internet

- It is important to know the precursors
  - ARPANET
  - Hypertext

ARPANET

- Established by the Advanced Research Projects Agency (ARPA) of the United States Department of Defense

- First Computers connected in 1969

- Defense purpose
  - Communication network free from interference
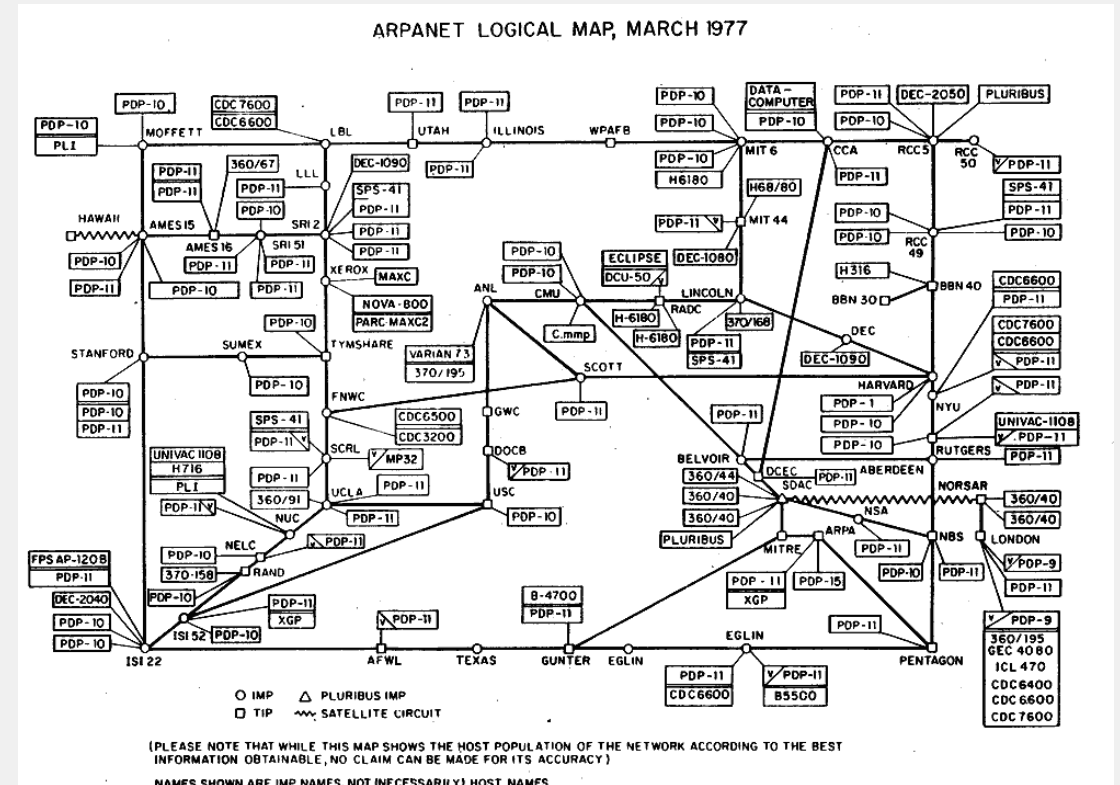
TCP Protocol

- In the 1970s as part of ARPANET Vint Cerf and Bob Kahn designed the TCP Protocol

  - Transmission Control Protocol

- Adopted as a standard for ARPANET in 1983

- Enables every Internet-connected device to communicate with every other such device simultaneously.

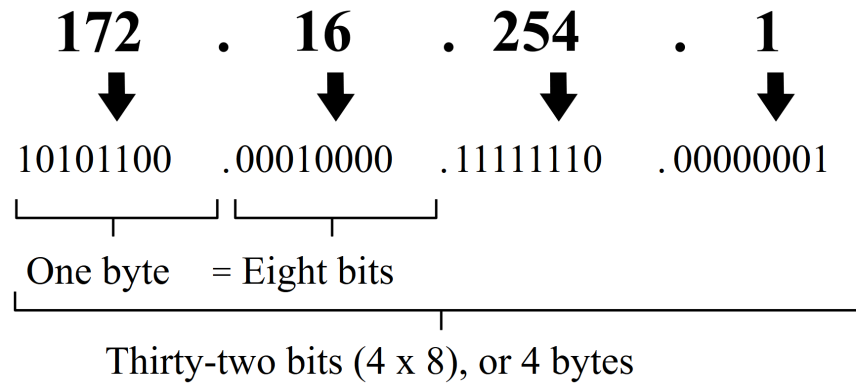- Later on they added the IP protocol
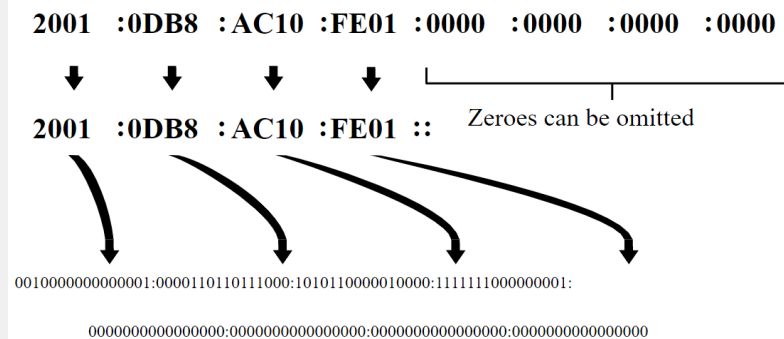
Internet: Packet Switching

- Transmitted data is grouped into suitably sized blocks called packets

- Packets are sent (and queued) across a network

  - Address: IPv4, 32bit; IPv6, 128bit

  - Best route chosen by intermediate nodes (e.g. routers) based on network traffic reports

  - Different packets, different routes, variable delay

- A packet-switching protocol governs the message transmission

  - Transmission Control Protocol and Internet Protocol (**TCP/IP**)

An IPv4 address   (dotted-decimal notation)

172 . 16 . 254 . 1

10101100   .00010000   .11111110   .00000001

One byte   = Eight bits

Thirty-two bits (4 x 8), or 4 bytes

An IPv6 address   (in hexadecimal)

2001 :0DB8 :AC10 :FE01 :0000 :0000 :0000 :0000

2001   :0DB8   :AC10   :FE01   ::   Zeroes can be omitted

0010000000000001:0000110110111000:1010110000010000:1111111000000001:

0000000000000000:0000000000000000:0000000000000000:0000000000000000
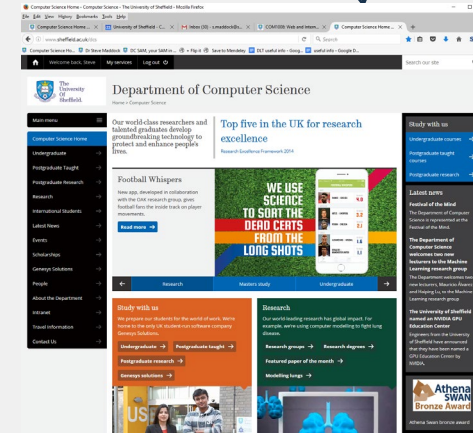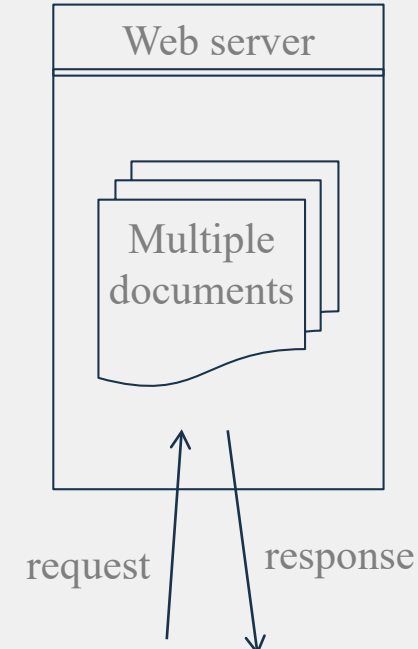
How does the WWW work?

- User runs a browser on a PC, Mac, mobile phone, tablet, games console, TV, wristwatch, car, fridge, …

- Browser (client) makes a request for a 'document', using a Uniform Resource Locator (URL):

http://staffwww.dcs.shef.ac.uk/people/S.Maddock/index.html
https://www.dcs.shef.ac.uk/dcs

Prefix	host name	path

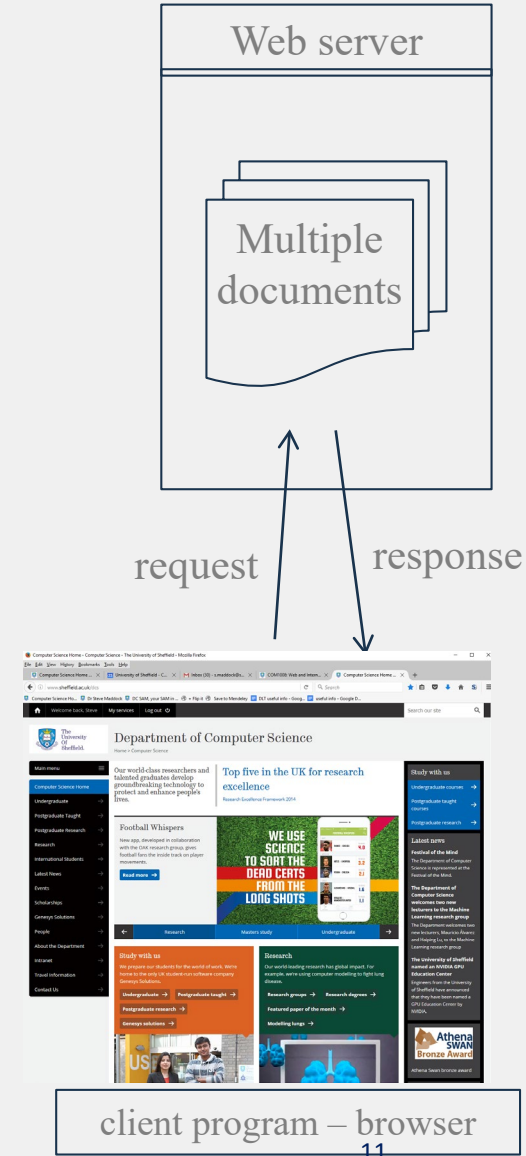Web server

Multiple documents

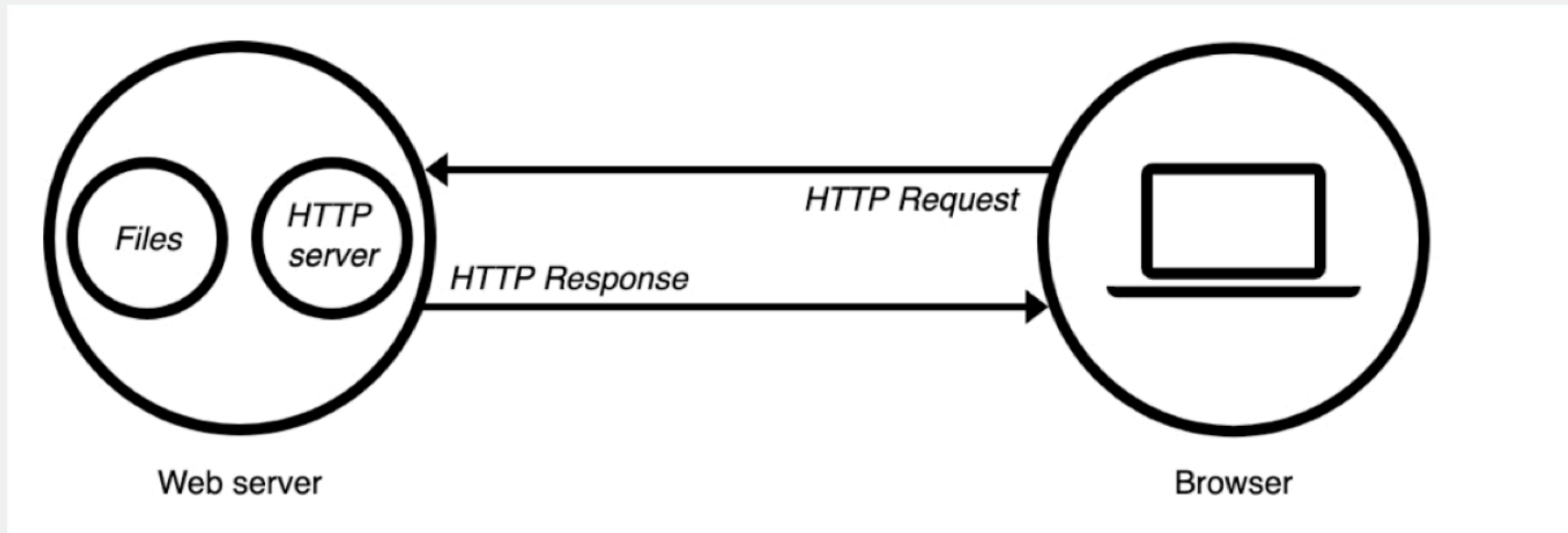request   response

client program – browser

How does the WWW work? – cont.

- Server responds with relevant (possibly dynamically-created) document(s)

- Communication governed by a *protocol*
  - **HTTP - H**yper**t**ext **T**ransmission **P**rotocol
  - Security: https and shtml

Web server

Multiple documents

request   response

client program – browser

A basic client server architecture

Client-server communication

- Web browsers communicate with web servers using

-   the **H**yper **T**ext **T**ransfer **P**rotocol (HTTP).

- HTTP is is the primary method used to convey information on the World Wide Web
  - http://en.wikipedia.org/wiki/Http_protocol

- It is a generic stateless object-oriented protocol,

- May be used for many similar tasks
  - E.g. name servers, and distributed object-oriented systems,

- by extending the commands, or "methods", used.

Transactions in HTTP

- A transaction consists of:

  - Connection

    The establishment of a connection by the client to the server

  - Request

    The sending, by the client, of a request message to the server;

  - Response

    The sending, by the server, of a response to the client;

  - Close

How does HTTP work?

- When the user performs an action on a web page
  - E.g. clicks on a link, submits an HTML form
- The browser sends an *HTTP Request* to the server.
- The Web server waits for request messages
- When a request arrives it is processed
- The server replies to the client with an HTTP Response message.

What does a request include

- The request includes
  - The URL of the file or of the web app
  - A method (get, post, put, delete etc)

Request methods

http://www.w3.org/Protocols/rfc2616/rfc2616-sec9.html

- The GET method means "retrieve whatever information (...) is identified by the Request-URI".

    - e.g. Your browser requires a page (e.g. containing a form) from a server using a GET method

- The POST method is used to <u>request</u> that the origin server accepts the entity enclosed in the request [and acts upon it].

    - e.g. the browser POSTs the values for a form to the server

What does a response include

- The response contains an HTTP <u>Response status code</u> indicating whether or not the request succeeded
    - E.g. 404 page not found

HTTPS

- *HTTP over TLS:* HTTP + connection encrypted by Transport Layer Security (or Secure Sockets Layer – *HTTP over SSL*)
  - Creates a secure channel over an insecure network

- Makes use of certificate authorities
  - Uses public-private keys and signing

- Authentication of server and website
  - Protects against man-in-the-middle attacks

- Bidirectional encryption of communications
  - Protects against eavesdropping

https://support.mozilla.org/en-US/kb/page-info-window-view-technical-details-about-page
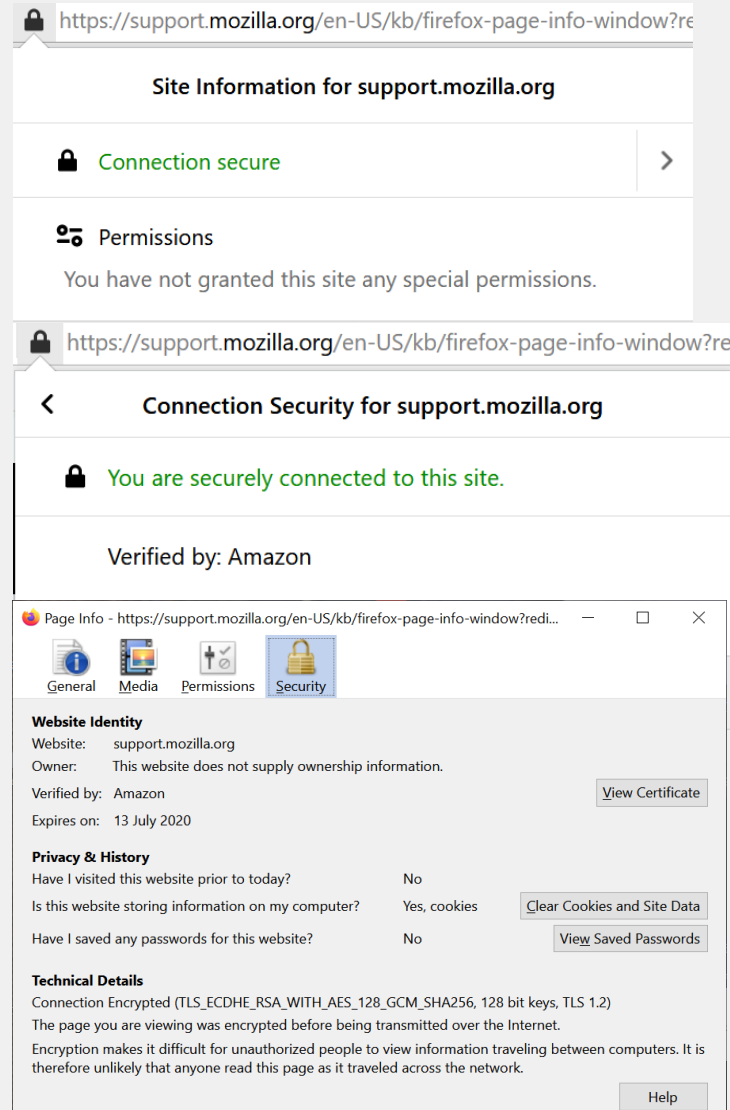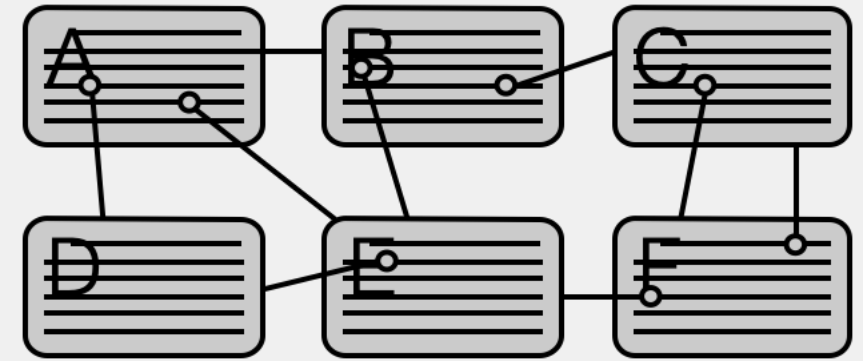
Hypertext

- Vannevar Bush (1890–1974) first described the concept of hypertext in drafts papers in the 1930s and then published a famous article "As we may think" in 1945

  - A device that would create links between different documents/images etc

- https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/
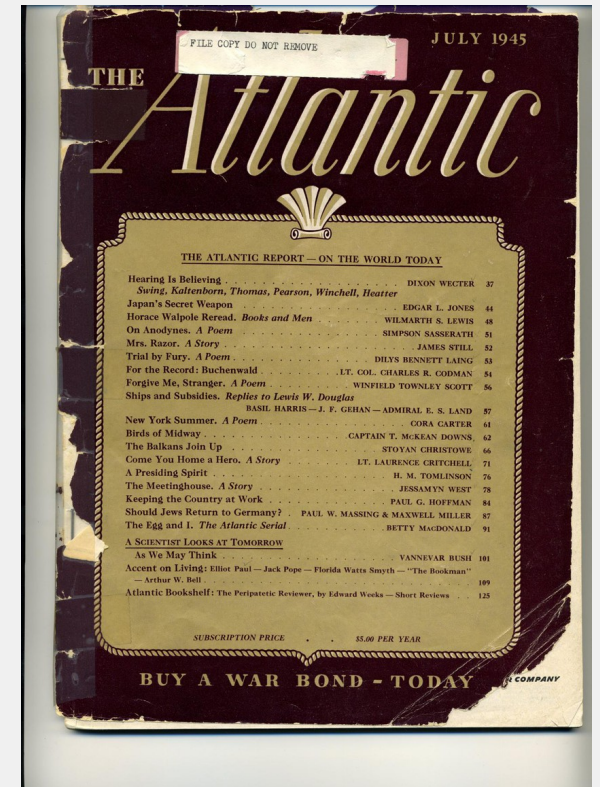
A proposal for the Internet

- In 1989 Tim Berners-Lee wrote a paper with a proposal for an information management system based on Hypertexts
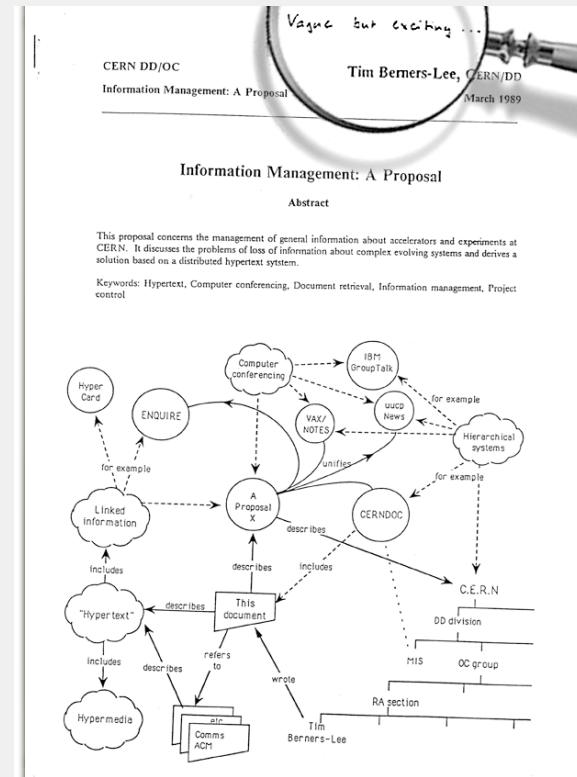
HTML, URI and HTTP

- In 1990, Berners-Lee and colleagues created the specifications for:

  - **HTML (HyperText Markup Language)**

  - **URI (Uniform Resource Identifier)** – a kind of 'address' that is unique and used to identify each resource on the web. It is also commonly called a URL.

  - **HTTP (Hypertext Transfer Protocol)** – allows for the retrieval of linked resources from across the web.
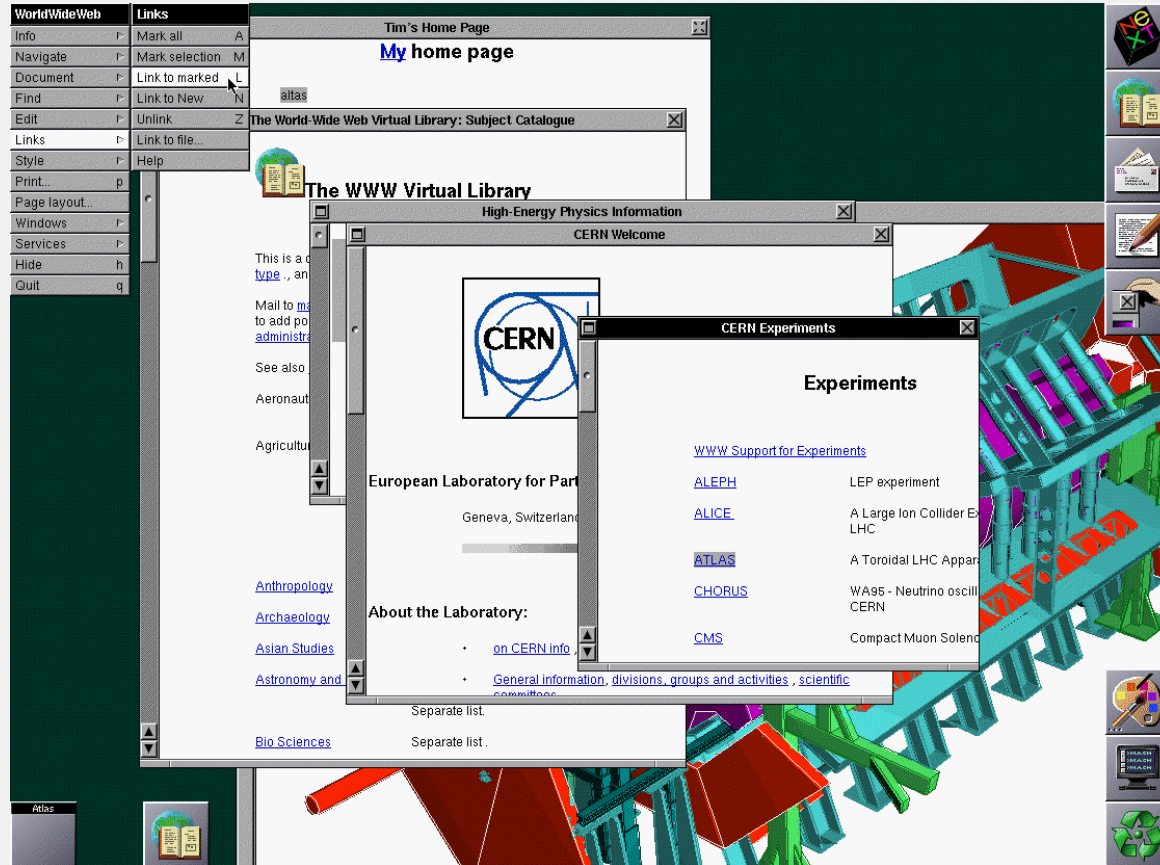
The first browser

WorldWideWeb by Berners-Lee and colleagues

A history of browsers
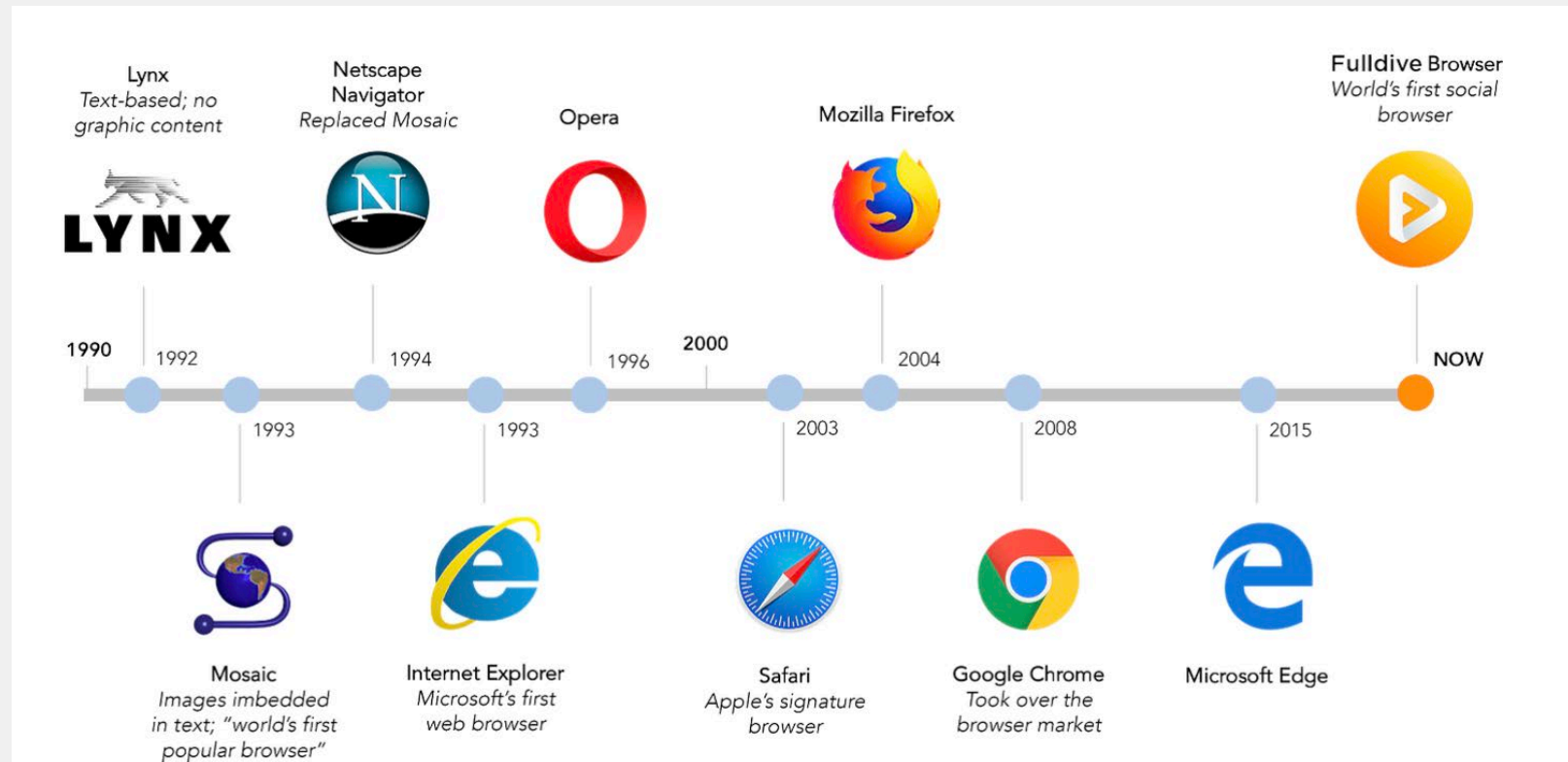
WEB DEVELOPMENT PROCESS

What is web development?

- Web development is the process of making websites and web-applications for the World Wide Web

- A website is normally split in two main sections
  - Client-side/Front-end
  - Server-side/Back-end

Front-end development

- Coding and creation of elements and features of a website that will be seen by the user.

- Focuses on the visual aspects and the interaction

- For example as part of front-end development you will build

  - Layouts

  - Navigation menus

  - Buttons

  - Graphics

  - Animations

Back-end development

- Coding and creation of features that the users will not see

- For example as part of back-end development you will build

  - Data storage solutions, i.e. database

  - A server to handle data requests and responses

Front-end web development languages

- You can use different languages to create websites

- We will look at the most popular ones
  - HTML
  - CSS
  - JavaScript

Process

- We can consider the creation of web sites as similar to a waterfall approach to software engineering

- It involves a series of stages

  - Disagreement on how many stages

  - Disagreement in which stage a particular step is taken

  - Larger companies may use more stages

  - Smaller companies may compress the stages, e.g. planning, design, develop, delivery

- There is iteration between the stages

Requirements

Design

Implementation

Testing

Maintenance
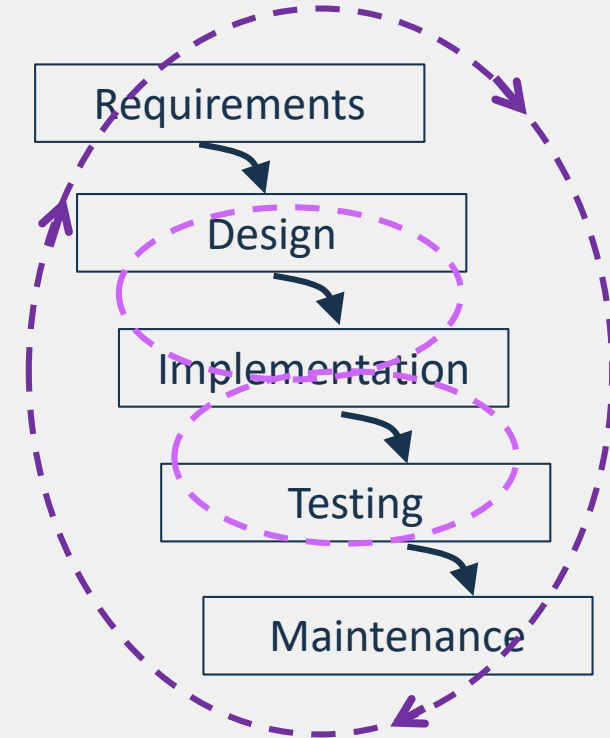
More likely there will be iteration of stages

Review, design, build, repeat (Marcotte, 2014)

Marcotte, E., Responsive Web Design, 2nd edition, A Book Apart, 2014

In Summary

- Today we have looked at:

  - Introduction to the Internet and WWW

  - Introduction to network protocols

  - Introduction to website development process

- In the next lecture we will look at

  - Structure of a web page

  - HTML

APPENDIX

A history of browsers - timeline

- A nice timeline
  - http://evolutionofweb.appspot.com/

HISTORY OF BROWSERS

Lynx

Mosaic

Netscape and Internet Explorer
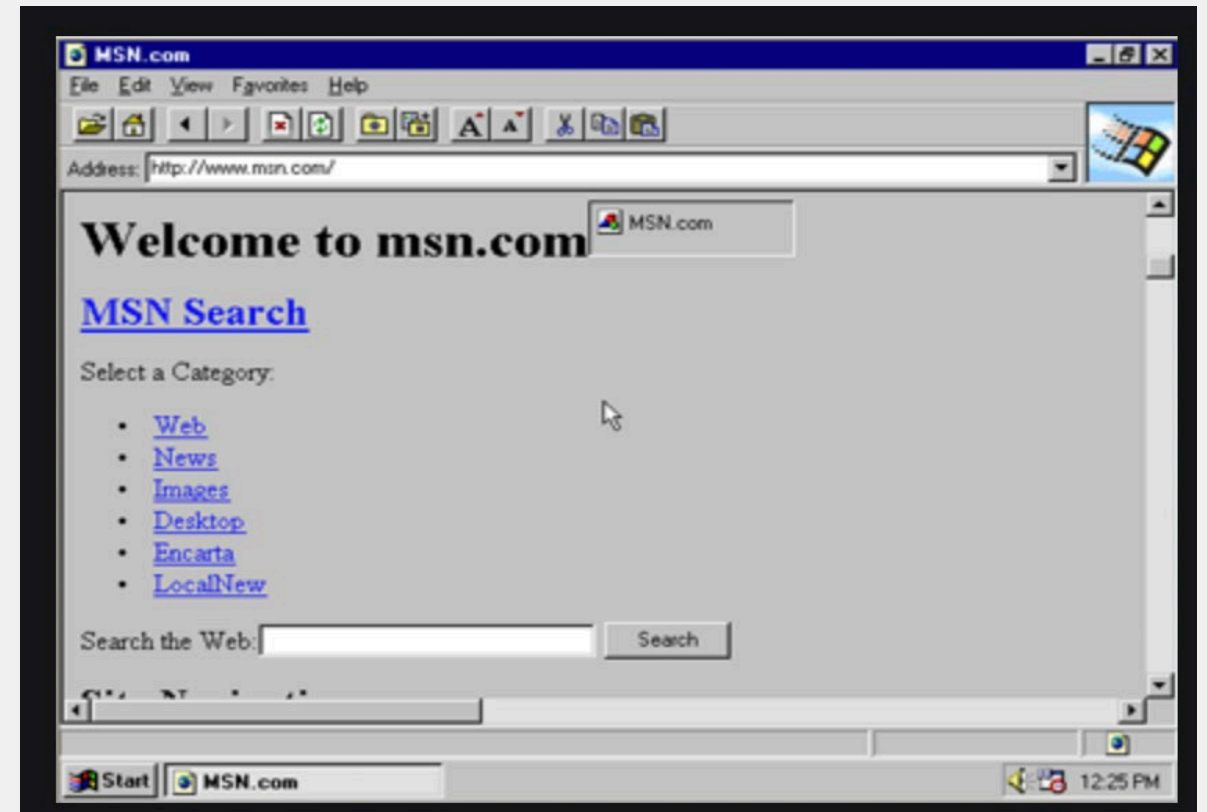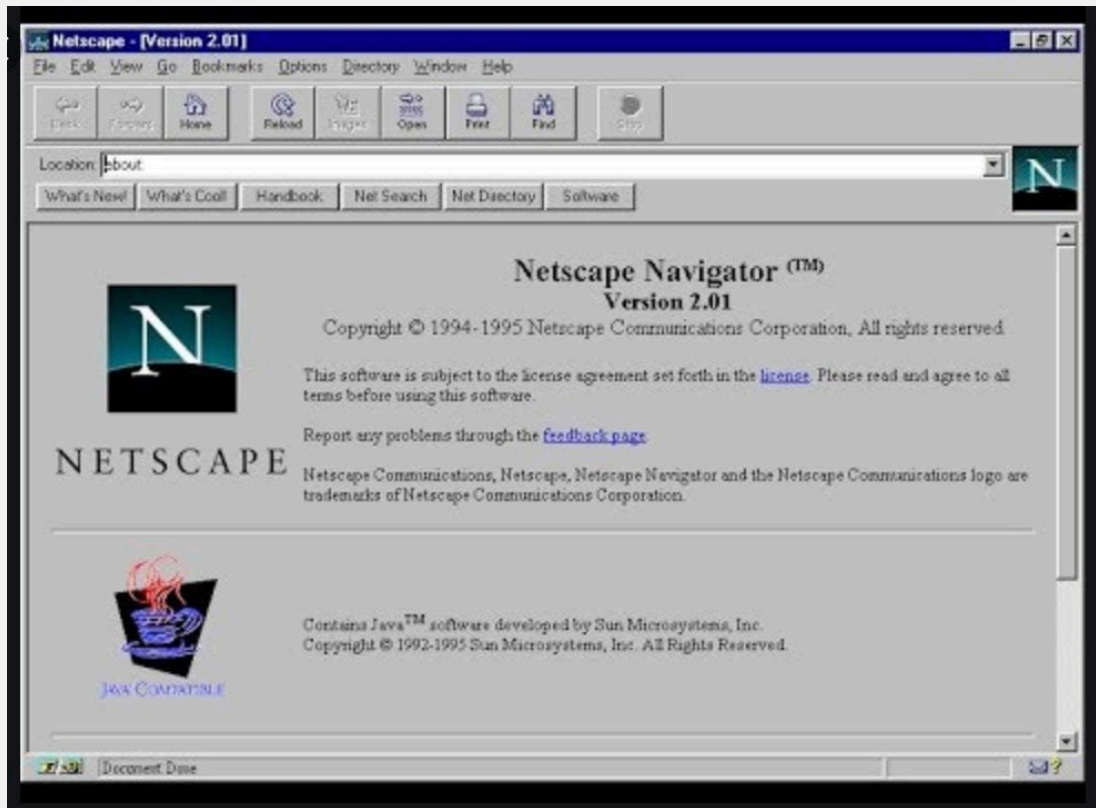
Today's browsers

A history of devices

Technology Cycles Have Tended to Last Ten Years

Mainframe Computing 1960s
Mini Computing 1970s
Personal Computing 1980s
Desktop Internet Computing 1990s
Mobile Internet Computing 2000s
Wearable / Everywhere Computing 2014+
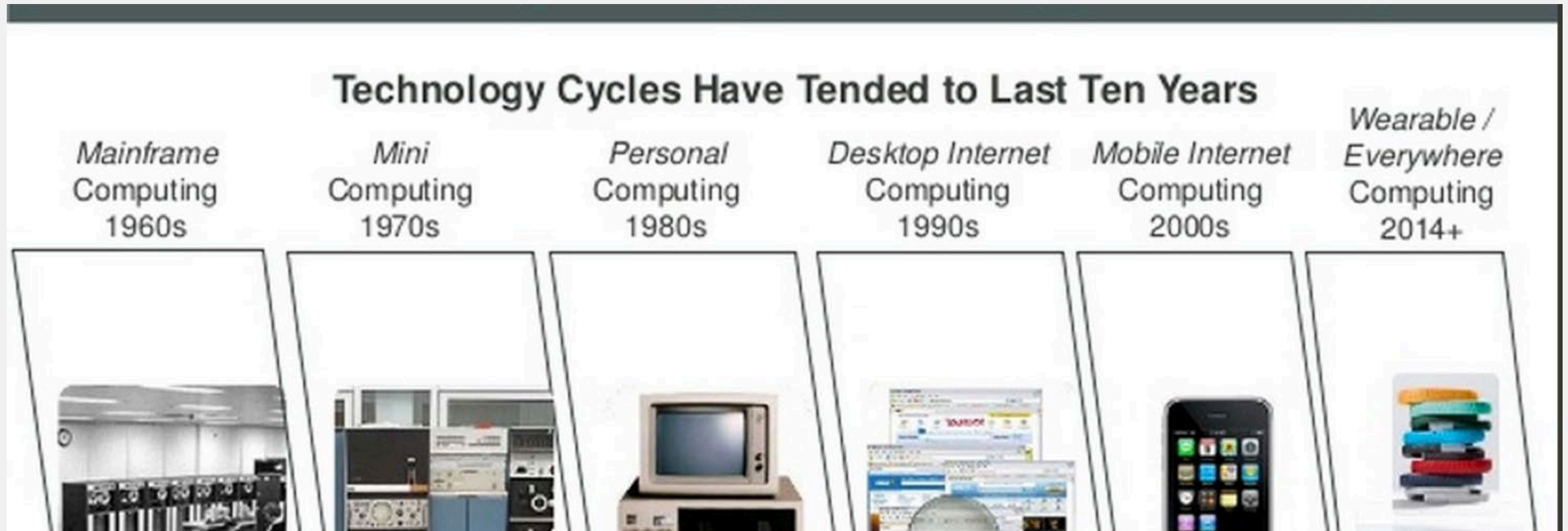
Mobile devices

HTTP PROTOCOL

HTTP response status codes

- Grouped in 5 classes

  - Informational responses (100–199)

  - Successful responses (200–299)

  - Redirects (300–399)

  - Client errors (400–499)

  - Server errors (500–599)

HTTP response status codes - examples

- <u>200 OK</u> - The request has succeeded.

- <u>301 Moved Permanently</u> - The URL of the requested resource has been changed permanently.

- <u>400 Bad Request</u> – wrong syntax

- <u>401 Unauthorized – unauthenticated.</u>

- <u>404 Not Found - The server can not find the requested resource.</u>

HTTP response status codes – examples 2

- 500 Internal Server Error - The server has encountered a situation it doesn't know how to handle.

- 502 Bad Gateway - the server got an invalid response.

- 503 Service Unavailable - The server is not ready to handle the request.

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