# Introduction Web Development I

16th September

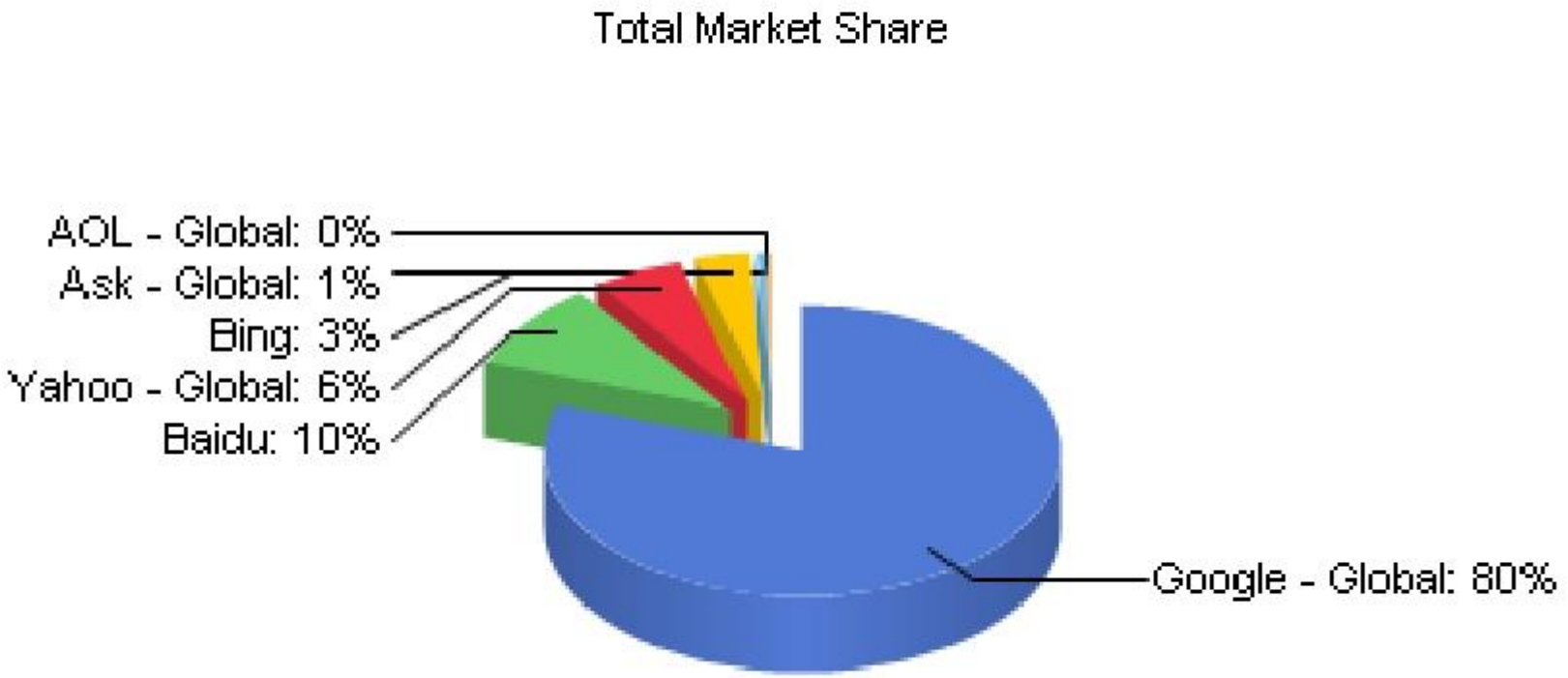
# Introduction to Web Development

- Search engines
- Discussion boards, bulletin boards, other online collaboration or knowledge-sharing resources
- Future directions to WWW

# Search Engines

- Directories
- Search engines
  - Meta-search engines
  - Search techniques
    - Word and Exact-Phrase Searches
    - Boolean Search
    - Other searches
- Synchronous/Asynchronous communication
- Browsers

# Popular Search Engines & Search Indexes (2011)




Source: http://marketshare.hitslink.com/search-engine-market-share.aspx?qprid=4

# Search Index (aka Search Directory)

- Popular Search Indexes:
  - Yahoo! http://www.yahoo.com
  - Open Directory http://www.dmoz.org
- A search directory maintains a hierarchical category of topics and places web site listings into these categories.
- Search either by typing in a search term OR "drilling" down into the hierarchy for relevant sites.
- Each site is reviewed by an editor before accepted into a directory

# Search Directory

# Search Engine Components

- Popular Search Engines:
  - Google http://www.google.com
  - Baidu http://www.baidu.com
- Search engines use the following components:
  - Robot or “spider”
  - Database (also used by search indexes)
  - Search form (also used by search indexes)

# Search Engine Robot

- Also called a **spider** or **bot**
  - A computer program that follows hyperlinks and "walks" the Web -- accessing and documenting Web pages.
  - Categorizes the pages and stores information in a database.
- May access the following components of Web pages:
  - title,
  - meta tag keywords & meta tag description
  - text in headings
  - other text on the page
  - hyperlinks
- See http://www.robotstxt.org

# Search Engine Database

- Database:
  - A collection of information organized so that its contents can easily be accessed, managed, and updated.
- Database Management Systems (DBMSs)
  - Configure and manage database
    - Oracle, MySQL, Microsoft SQL Server, IBM DB2
- Search Engine Database
  - Contains information about web pages

# Search Engine Search Form

- The search form is the graphical user interface that allows a user to request a word or phrase to search for. It is usually just a text box and a submit button.
- The visitor to the search engine types words (called keywords) related to their search into the text box.
- When the form is submitted, the data typed into the text box is sent to a server-side script that searches the database using the keywords you have entered.
- The search results (also called a result set) is a list that contains information such as the URLs for web pages that meet your criteria.

# Search Engine Results Page (SERP)

- A list of items that describe Web pages matching the search terms.
  - Each item contains a link to a page along with additional information that might include the page title, a brief description, the first few lines of text, the size of the page, and so on.
- The order the Web page items are displayed in the SERP may depend on:
  - paid advertisements
  - alphabetical order
  - link popularity
- Each search engine has their own policy for ordering the search results.

# Effective Search Techniques

- Be specific
- Don‟t use common words
  - e.g. “internet”, “program”...
- Learn to adjust your query
  - “run”, “runs” and “running”
- Use synonyms
  - use “jog” OR “run”
- Use quoted phrases for exact matches for phrases
- Use uppercase where appropriate
  - most tools are case sensitive (upper case means only upper case)

# Effective Search Techniques

- Understand the search tools you are using
  - look for a Tips link
- Use Boolean Queries where possible
  - AND = all words present
  - OR = one or more words
  - NOT = word not present
  - can use parentheses to refine search
- + and - (alternative to boolean)
  - + = word must be present
  - - = work must not be present

# Synchronous Communication

- Synchronous communication
  - Text-only based web chatting
    - Bulletin Board system
    - Internet Relay Chat
  - Instant Messaging
  - Webcasting

Is a podcast synchronous communication?

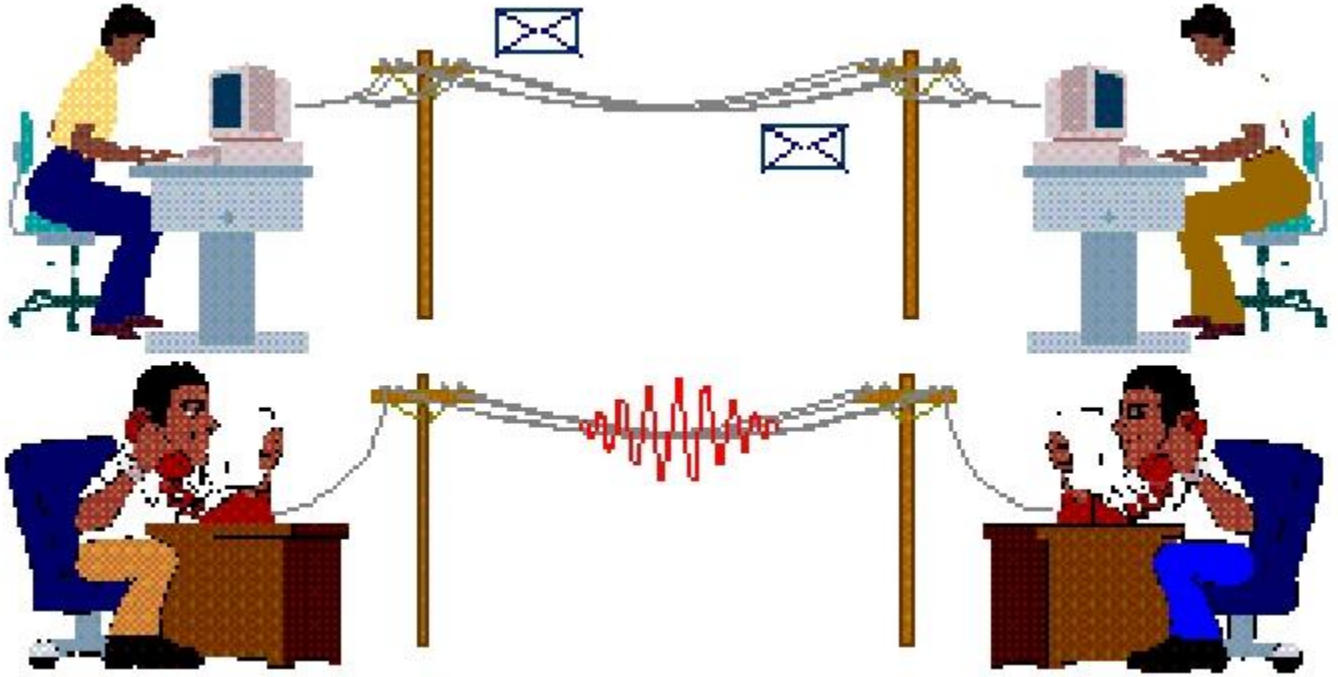
# Synchronous Communication

- Text-Only based web chatting
  - Bulletin Board system
    - Computer application dedicated to exchanging messages and files on a network
    - Free to use
    - User needs an account to log in
  - Internet Relay Chat (IRC)
    - Based on BBS technology
    - Enables forums for discussion and real-time chatting
    - Very popular amongst UNIX users
    - Example: http://www.efnet.org

# Asynchronous Communication

- Asynchronous communication
  - Mailing lists
  - Discussion groups/Web forums
  - Emailing

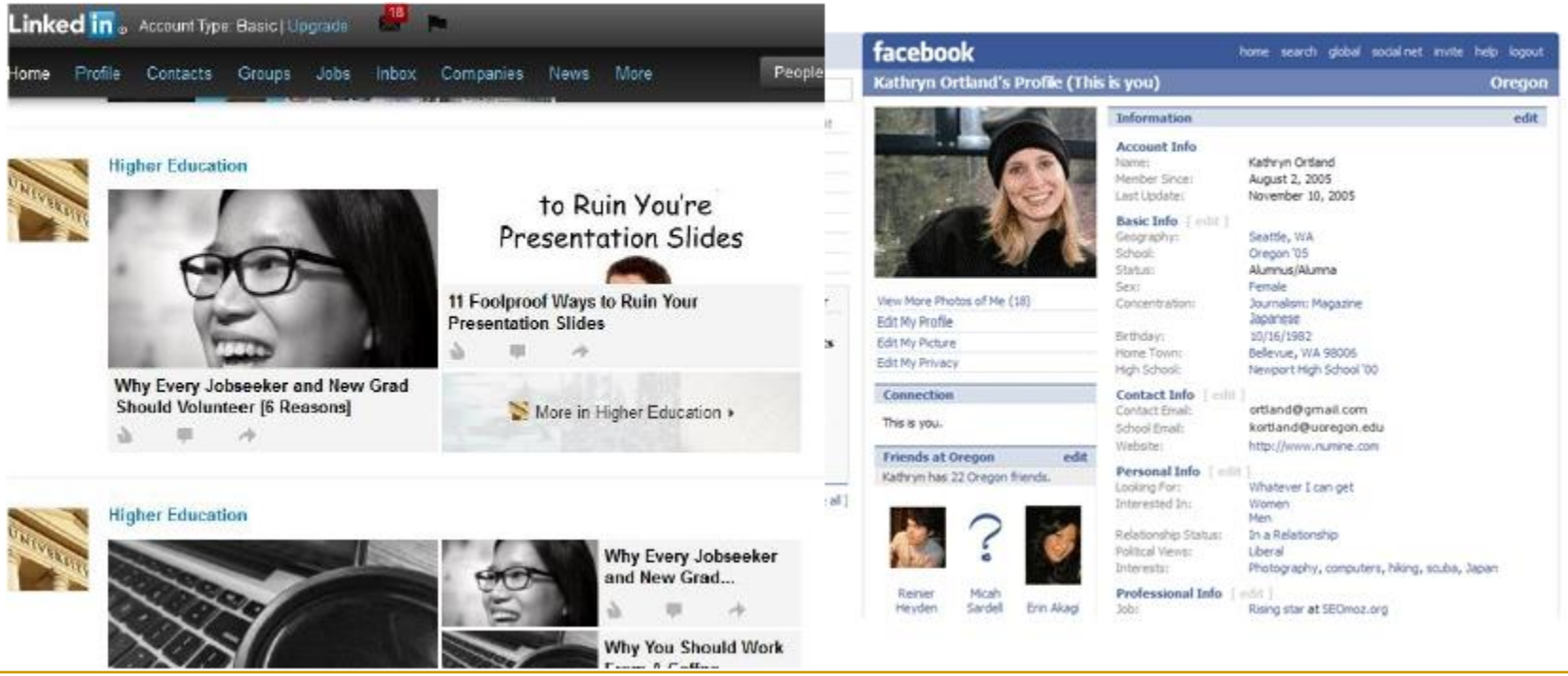
# Asynchronous versus synchronous communication



# Communication

Using whatever SNS site that you use, identify

1 form of synchronous and one form of asynchronous communication



# Browsers

- Microsoft‟s **Internet Explorer**
- Mozilla **Firefox**
- Google **Chrome**

# Web Browser Capabilities

- Browsing URLs
- History Lists
- Printing, Saving, Sending
- Bookmarking Favorites
- Viewing Document Source
- Image Formatting Support
- Caching
- Toolbar Buttons
- Security
- Native Multimedia Support …

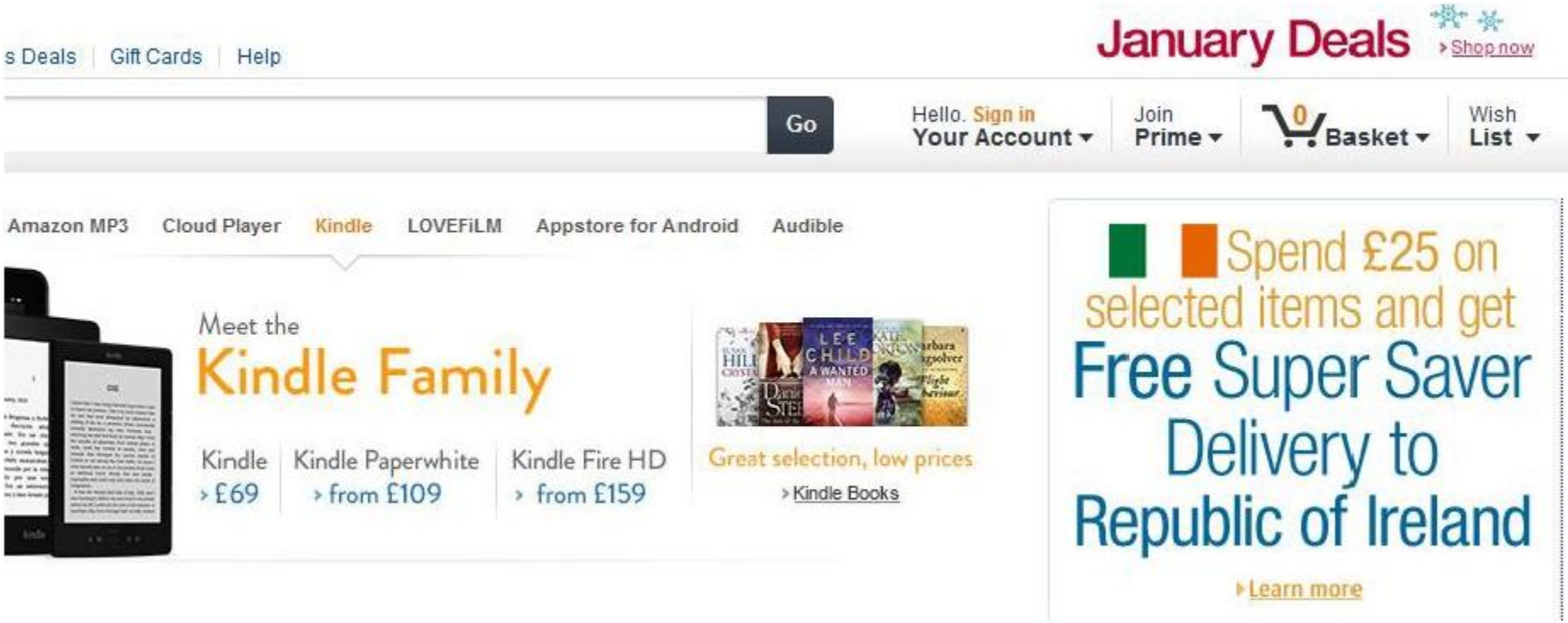
# Future Internet & Web Trends

- Continued importance of E-Commerce
- Wireless Web access
- IPV6
- Web Services
- Blogs
- Twitters
- RSS
- Wikis
- Web 2.0
- HTML5

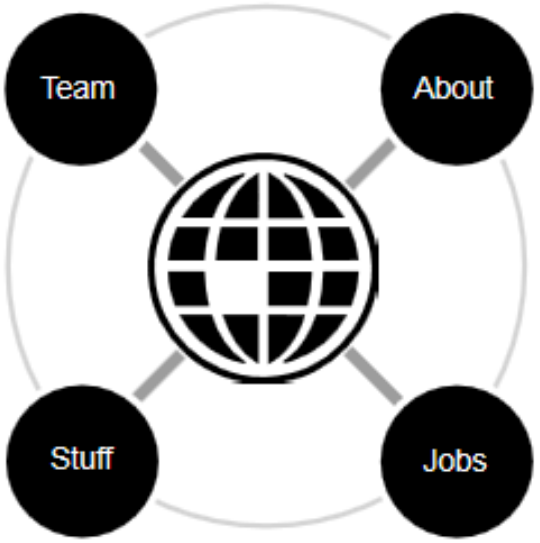
# Future Internet & Web Trends

- Security
  - “Death of the Open Web”?
- The World into the Web
  - Browser-based Everything
  - The Internet as a Collective Consciousness
- The Web into the World
  - The “Web of Things”
  - Interactive Landscapes
  - The “Innerweb”

# Web-sites analysed at home work



# Web-sites analysed at home work




## Stupid Fun Club

Welcome to the Stupid Fun Club

# Web-sites analysed at home work

# Web-sites analysed at home work

# Web-sites analysed at home work