CSIT 70 Web 2.0 – The Web's Edge

Instructor Teresa Pelkie

History and Development of the World Wide Web – Class # 1

Overview of Web 2.0

The term "Web 2.0" was coined in late 2004 by O'Reilly Media's Dale Dougherty. He stated that far from having crashed, the Web was becoming filled with new types of applications and websites, which was the beginning of a new second phase of architecture and development for the World Wide Web. Some of the characteristics are:

- Mass publishing with content controlled by the end-user
- New ways of communicating using the Web
- Tag based classification and searching
- New delivery methods of information
- Sharing and re-use of information
- The Web as a platform of communities used to connect with others
- Conversation, collaboration, contribution, and participation
- The transition of websites from isolated silos of information to sources of content and functionality

History of the World Wide Web:

The Web has revolutionized our lives and work more than any other technological development in recent history.

ARPANET - 1969

President Dwight D. Eisenhower created the Advanced Research Projects Agency **(**ARPA**)** in 1958 which created a computer network called ARPANET (around 1969). ARPANET established the protocols, most notably TCP/IP, which are still used on the Internet today. In the late 70's other networks - USENET**,** BITNET**,** CSNET and NSFNET **–** joined ARPANET. They called the connection between multiple networks inter-networking, or the Internet for short. The Internet consisted of protocols such as SMPT, FTP, Gopher, and TELNET.

The World Wide Web Begins - 1990

During 1989-1990, Tim Berners-Lee, at the European Organization for Nuclear Research (CERN), Switzerland, develop a system for linking and navigating large amounts of information. He developed the HyperText Transfer Protocol (HTTP) and the HyperText Markup Language (HTML).HTTP is the protocol that a Web browser uses to send a request to the server, and the server used to send a reply back to the browser. HTML is the language in which Web pages are written.

Most early Internet users were government and military employees, graduate students and computer scientists. Using the World Wide Web, the Internet became much more accessible. Colleges and universities began to connect to the Internet, and businesses soon followed. The Web browsers were text based browsers, and communication was still primitive.

Mosaic Web Browser – 1993

The turning point for the World Wide Web was the introduction of the Mosaic Web browser in 1993, a graphical browser developed by a team at the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign (UIUC), led by Marc Andreessen. Funding for Mosaic came from the High-Performance Computing and Communications Initiative, a funding program initiated by then-Senator Al Gore's High Performance Computing and Communication Act of 1991 also known as the Gore Bill. The Mosaic Web browser could display text, links, images, and had most of the features that modern browsers have today.

Netscape Web Browser – 1994 - 1999
In mid 1994, Silicon Graphics founder Jim Clark collaborated with Marc Andreessen, and ultimately founded Netscape Communications. The Netscape Navigator browser was released, which introduced many innovative features over the years which enhanced the Web browsing experience.

Other Browsers
Microsoft introduced the Internet Explorer Browser in1995. Eventually Internet Explorer and Mozilla Firefox (2004) became the predominant browsers.

The Early Web

- Read only
- Needed technical skill to create pages
- Slow, hosting expensive
- Fixed size, colors vary

Today

- An interactive application

HTML and HTTP – the Client/Server Principle

The Web browser (client) and the Web server communicate via the HTTP protocol.

1. The browser sends a **request** (ie for a Web file) to the server. The request is packaged in a TCP/IP package.
2. After the browser sends the request the connection is closed. The browser sits and waits for the reply.
3. This is known as "**synchronous**" communication – meaning that the browsers sits and waits for the reply before it can continue.
4. The server receives the request, processes it, and sends back a **reply** (ie for a Web file) to the browser.
5. After the reply is sent, the connection is closed.

Web's Effect on Business and Communication:

- No longer do people or companies need to be in close proximity to communicate or do business.
- There are no more boundaries, we can communicate, work together and collaborate no matter where we are.
- The Netscape IPO (stock) release in 1995 signified that a company can make money through the Internet – this was the start of the "dot-com bubble"

Main Goal of the Web – the Quest for Information (why Berners-Lee developed the Web)

- Linking – most revolutionary aspect of the Internet
- Hyperlinks that connect to a resource – a Universal Resource Locator (URL) = address of a resource One of the
- Became a critical aspect of information location and retrieval
- Navigation – linking within a group of pages (website) – going back and forth

How We Find Information

1.Search Engines:
Search engines go back to 1993. They generally show results based on how often the search terms occurred in the page, or how strongly associated the search terms were within each resulting page. They consist of:

1. Crawler / spider / bot – finds pages and copies them to the search engine
2. Indexer – extracts terms, organizes the terms and URL into the database
3. Runtime system – answers the search query, displays the results using a ranking algorithm
4. Popular Search Engines"
    - Excite (1993)
    - WebCrawler, Lycos, Infoseek – Netscape's default search tool, AltaVista (1994)
    - Inktomi / HotBot (1995)
    - Ask Jeeves, Google (1997 Larry Page and Sergey Brin)

2.Directories and Portals:
A directory classifies information into a hierarchical organization. A portal is an entry point to a number of topical sections. They are created by humans, rather than computer programs. Some examples are:

- Virtual Library –set up by Tim Berners-Lee (1990)
- EINet Galaxy, Yahoo (1994)
- LookSmart (1995)
- Open Directory Project / DMOZ (1998)

Google's PageRank

Google's algorithm uses a patented system called **PageRank** (named after co-founder Larry Page) to help rank web pages that match a given search string. The PageRank algorithm is based on the pages linking to them. The PageRank derives from human-generated links, and is thought to correlate well with human concepts of importance.. Previous keyword-based methods of ranking search results, used by many search engines that were once more popular than Google, would rank pages by how often the search terms occurred in the page, or how strongly associated the search terms were within each resulting page. In addition to PageRank, Google also uses other secret criteria for determining the ranking of pages on result lists.

Google not only indexes and caches web pages but also takes "snapshots" of other file types, which include PDF, Word documents, Excel spreadsheets, Flash SWF, plain text files, online videos such as YouTube and much more.

Commercialization of the Web

Security / Encryption:
1994 included an important security protocol called Secure Socket Layer (SSL) that encrypted messages on both the sending and receiving side of an online transaction. SSL ensured that personal information like names, addresses and credit card numbers could be encrypted as they passed over the Internet**.**

Online Payments:
1994- 1995 First Virtual and CyberCash Internet payment systems emerged for online payments. PayPal was founded in 1998 and was acquired by eBay in 2002.

Digital Certificates:
1995, a company called Verisign began developing digital IDs, or certificates, that verified the identity of online businesses and to certify that a Web site's e-commerce servers were properly encrypted and secure.

Software:
E-commerce shopping carts

E-commerce Begins:

- 1995 - Amazon.com (Jeff Beezos) set the standard for a customer-oriented e-commerce Web site. Users could search available titles by keyword, author or subject. They could browse books by category and even get personalized recommendations. People can now buy from companies – business to customer.
- 1995 - Pierre Omidyar, a software programmer, started coding a simple Web site he called AuctionWeb, which later became E-bay. People can now buy from other people – customer to customer.

Advertising:

- DoubleClick (banner ads) - customer choose the ads
- Google ADSense – ads are generated by user query
- Google AdWords – ads are generated by user query

Hardware Developments

- Cheeper, smaller, faster, more powerful
- VoIP, Broadband and wireless technology for fast, easy access

Software Developments

- New languages to complement HTML – XML, CSS, scripting
- Free tools, open source software, Web services