Meta Information

BSc Information Technology – Yr 2

Semester 2

Dr. Brenda Mullally

1

Meta Information

50% Continuous assessment

50% Final exam – 3 questions do 2.

2 Lectures per week

2 Practicals per week

Meta Information

Continuous Assessment:

10%: lab work – keep a log of all the lab exercises you complete, this will be checked during and at the end of the semester.

40%: Individual project

Syllabus

Introduction

Web 2.0

Web information management

XML

Extensible style sheets (XSLT)

PHP

Syllabus

Introduction

Web 2.0

Web information management

XML

Extensible style sheets (XSLT)

Introduction

What is this module about?

The analysis, design, and implementation of the information that creates websites and other publications using XML.

First we will look at the access to, and the availability and modification of information on the web.

Introduction

Users of computers want applications that use multimedia graphics, images, animation, audio and video. They want applications that run on the Internet and the web and communicate with other applications.

Internet users want interactive websites where they can post comments, upload media, participate the creation of the website content.

Among all American adults ages 18 and older, the % in each group who go online in the following ways

		% use internet TOTAL	% use internet at home	% use home broadband	% use home dial-up	% use internet, but NOT from home
	All Americans ages 18+ (n=2,252)	85%	76%	70%	3%	9%
a	Men (n=1,029)	85	77	71	3	9
b	Women (n=1,223)	84	76	69	3	9
Race/ethnicity						
a	White, Non-Hispanic (n=1,571)	86[c]	79[bc]	74[bc]	3	6
b	Black, Non-Hispanic (n=252)	85[c]	70	64[c]	2	15[a]
c	Hispanic* (n=249)	76	63	53	6	13[a]
Age						
a	18-29 (n=404)	98[bcd]	87[bcd]	80[cd]	2	10[d]
b	30-49 (n=577)	92[cd]	83[cd]	78[cd]	2	10[d]
c	50-64 (n=641)	83[d]	75[d]	69[d]	4	8
d	65+ (n=570)	56	50	43	4	6
Education attainment						
a	No HS diploma (n=168)	59	45	37	4	14[d]
b	High school grad (n=630)	78[a]	65[a]	57[a]	4[cd]	13[cd]
c	Some College (n=588)	92[ab]	83[ab]	78[ab]	2	8[d]
d	College + (n=834)	96[abc]	93[abc]	89[abc]	2	3
Household income						
a	Less than $30,000/yr (n=580)	76	63	54	4[cd]	13[cd]
b	$30,000-$49,999 (n=374)	88[a]	76[a]	70[a]	4[c]	12[cd]
c	$50,000-$74,999 (n=298)	94[ab]	88[ab]	84[ab]	2	6
d	$75,000+ (n=582)	96[ab]	92[ab]	88[ab]	2	4
Urbanity						
a	Urban (n=763)	86[c]	76[c]	70[c]	2	9
b	Suburban (n=1,037)	86[c]	79[c]	73[c]	3	7
c	Rural (n=450)	80	70	62	4	10

8

www.pewinternet.org

Introduction

15 Most popular Websites

http://www.ebizmba.com/articles/most-popular-websites

Introduction

You have already learnt how to structure web sites using HTML, how to present web sites using CSS, how to include actions using JAVA script. These skills are useful for building applications that reside on the client side of a web based application.

This module will teach you about the management of data for web sites.

HTTP

Web pages

The main protocol used for communication between a browser and a Web server is HTTP. This protocol was designed to enable documents to be transferred but can be used with other types of data too. For Web documents the HTTP protocol works by sending commands over a TCP connection.

To understand how information gets passed from machine to machine we need to know how such systems can connect to each other. Generally, the main model used for the web is client-server.
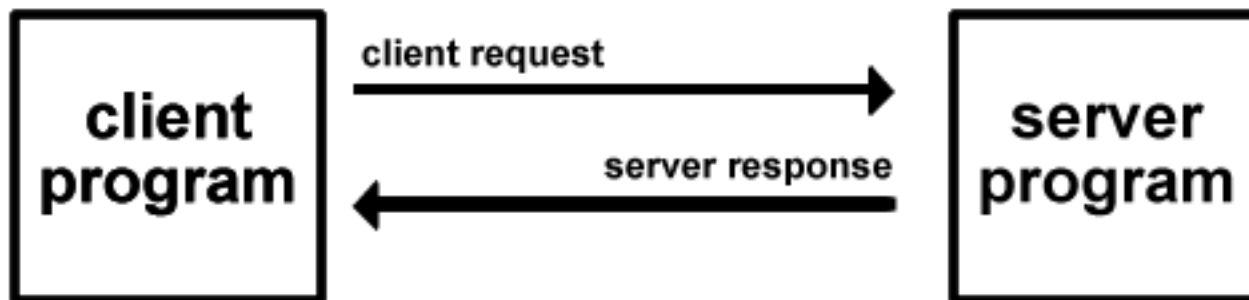
HTTP

Client/Server Computing:

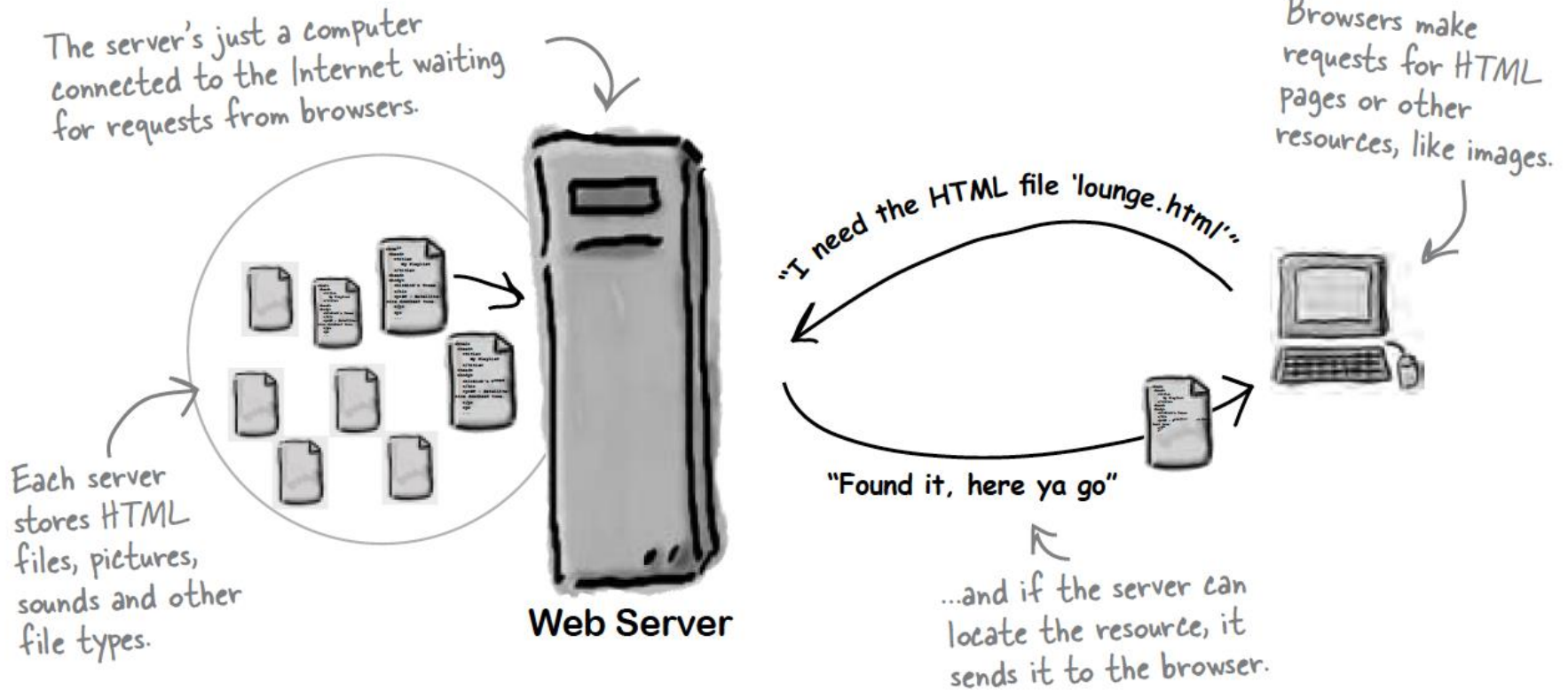The interaction between two programs when they communicate across a network.

A program at one site sends a request to a program at another site and awaits a response.

The requesting program is called a client; the program satisfying the request is called the server.
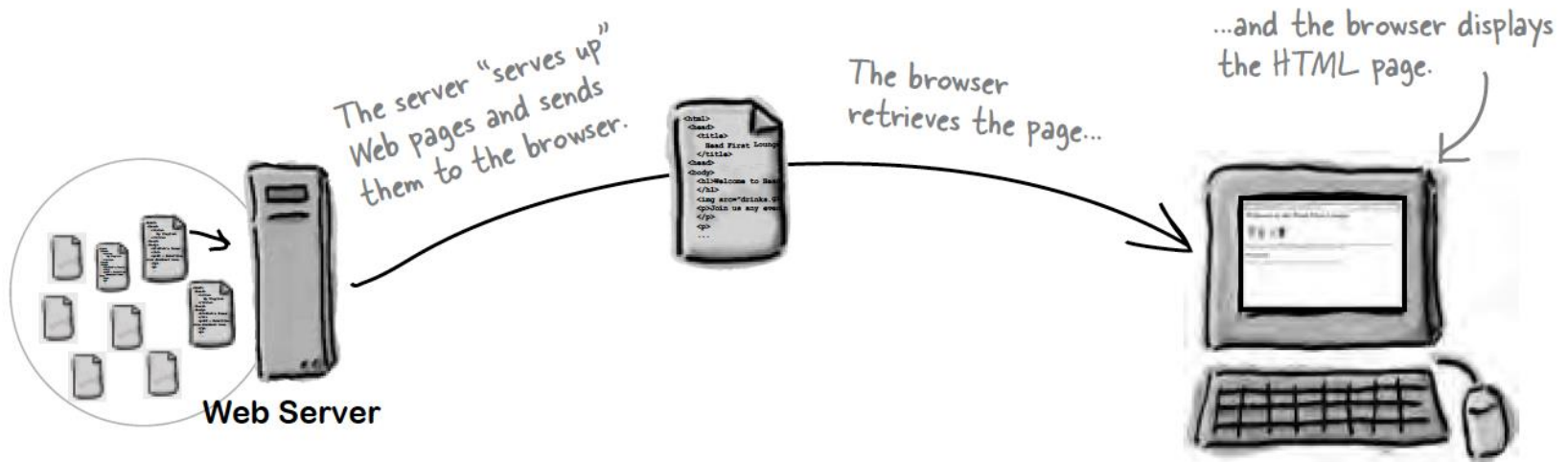
HTTP

Role of server

The server's just a computer connected to the Internet waiting for requests from browsers.

Browsers make requests for HTML pages or other resources, like images.

"I need the HTML file 'lounge.html'"

Each server stores HTML files, pictures, sounds and other file types.

"Found it, here ya go"

...and if the server can locate the resource, it sends it to the browser.

Web Server

HTTP

Role of Client

The server "serves up" Web pages and sends them to the browser.

The browser retrieves the page...

...and the browser displays the HTML page.

Web Server

HTTP

A typical HTTP session

Under HTTP there are 4 steps involved in communicating across the Web

Make the connection
Request the document
Respond to a request
Close the connection

1. Make a connection
The browser will open up a standard TCP connection to the server. Port 80 is the default port.

HTTP

2. Make the Request

Once the TCP connection is established, the browser will request a given document using the URL.

The command GET tells the server that the browser is trying to retrieve a document. The connection to the server should already have occurred so there is not need to specify the full DNS name. The browser can send a variety of other commands including **POST,** which sends Form data to the server, **PUT** which is used to transmit a data file to a server (probably a database server). The browser can also append an *Accept,* which will indicate the data types it can handle. The name of the application that makes the request may also be appended using the command *User-Agent*. So a typical request could look like:

GET / first.html HTTP 1.1
Accept: text/html
Accept: text/plain

HTTP

3. Server Response

When the server receives the HTTP request it locates the appropriate document and returns it. An HTTP response is required to have a particular form.

Different HTTP response status codes:
200 OK   This is the commonest code, it indicates that the message contains all the requested data.
 204 NO Content   The request was processed successfully but there was no data to return to browser
 400 Bad Request   The request from the client used invalid syntax and could not be processed
 401 Unauthorized some form of authorization information is needed before this resource can be accessed.
 404 Not Found   This is the commonest error response. It indicates that while the request was valid, the server could not find the document.

HTTP

4. Closing the connection

The client and server can both close the connection. This uses the standard TCP approach. If another document is required, the entire process must be repeated again.

Syllabus

Introduction

Web 2.0

Web information management

XML

Extensible style sheets (XSLT)

PHP

Web 2.0

In 1993 the web exploded in popularity and continued in growth throughout the 1990's. In 2003 there was a noticeable shift in how people and businesses were using the web and developing web-based applications.

The term Web 2.0 was coined to describe this new trend. Generally Web 2.0 companies use the web as a platform to create collaborative, community-based sites.

This trend has been assisted by the decreasing hardware costs and the increasing access to broadband connections. The abundance of digital media online would never have been possible without high-speed Internet access. Also the availability of open source software has resulted in cheaper customisable software options, this makes it easier to start new Web 2.0 companies and decreases the cost of failure.

Web 2.0

Web 1.0 was focused on relatively small number of companies and advertisers producing content for users to access, some called the web at the time the "brochure web".

Web 2.0 involves the user, not only is the content often created by users, but users help organise it, share it, remix it, critique it, update it etc.

Web 2.0 uses an architecture of participation, a design that encourages user interaction and community contributions.

Many Web 2.0 companies are built almost entirely on user-generated content and harnessing collective intelligence.

Social bookmarking sites allow users to recommend their favourite sites to others.

Social media sites enable the community to decide which news articles are the most significant.

Social networks have changed the way we interact and network.

Web 2.0

Search Engines:

In Web 2.0 "content is king" Finding content is key to the success of Web 2.0 business models. Search engines are the primary tools used to find information on the web.

Vertical search engines have the highest volume of traffic, horizontal search engines are becoming more popular but still lag behind.

Currently we use keywords to search, in the future natural language will be used.

People assisted search engines have also emerged such as www.mahalo.com and www.stumpleupon.com

Google has approximately 67% of the U.S search market, followed by Microsoft with 18%, Yahoo! with 11% Ask with 2.6% and AOL with 1.4%.

Web 2.0

Google is so popular that its name has been added to the Oxford English Dictionary – the verb "Google" means to find something on the Internet using the Google search engine.

Google uses its PageRank ™ algorithm which considers the number of links into a web page and the quality of the linking sites to determine the importance of the page. Google search also considers all of the content on the page, its fonts, its headers and the content of neighbouring pages. Sites with the highest PageRank will appear at the top of the search results.

Google also has some vertical search engines for images, news, videos, blogs etc.

Using Google WebServices you can build Google Maps and other Google services into your applications.

Google also now has a store where you can buy apps, music, books, there is also Google+ , Google Drive, Hangouts.

Web 2.0

User-generated Content:

Key to the success for many leading Web 2.0 websites such as Amazon, eBay and Monster is user generated content. eBay relies on the community to buy and sell items, Monster connects job seekers with employers and recruiters. User-generated content includes explicitly generated content such as articles, home videos and photos. It can also contain implicitly generated content.

What implicit content does Amazon hold?

Web 2.0

Collective Intelligence:

This concept harnesses the idea that collaboration can result in smart ideas, similar to a group brain storm session.
User-generated content is significant to Web 2.0 companies due to the innovative ways companies are harnessing collective intelligence. Amazon's personalised recommendations uses collective intelligence.

Web 2.0

- The Millennium Project connects futurists around the world to improve global foresight. It was founded in 1996 and is now an independent non-profit global participatory futures research think tank of futurists, scholars, business planners, and policy makers who work for international organizations, governments, corporations, NGOs, and universities.
- The Millennium Project manages a coherent and cumulative process that collects and assesses judgments from over 3,500 people since the beginning of the project selected by its 49 Nodes around the world

Web 2.0

- The work is distilled in its annual **"State of the Future"**, **"Futures Research Methodology"** series, special studies, and integrated into this **Global Futures System.**

- https://themp.org/

- How does eBay use collective intelligence?

Web 2.0

Wikis:

Websites that allow users to edit existing content and add new information, are prime examples of both user generated content and collective intelligence. The most popular wiki is Wikipedia, a community generated encyclopaedia with articles available in over 200 languages.

Wikipedia trusts its users to follow certain rules, such as not deleting accurate information and not adding biased information, while allowing community members to enforce the rules.

Moderation of content is costly so Web 2.0 companies rely on the community to do the policing. This is known as collaborative filtering, valuable material is promoted and offensive or inappropriate material is flagged.

Web 2.0

Blogging:

Blogs are websites consisting of entries listed in reverse chronological order. Interest in blogging has grown exponentially in recent years because of easy to use blogging software and increasingly economical Internet access.
Blogs can incorporate media and some are personal journal blogs, many include social networking features.
Blogs have become a significant news source, drawing traffic away from mainstream media, some call it participation journalism, blogging gives a voice to everyone with a computer and Internet access.

Web 2.0

Social Networking:

Social networking sites allow users to keep track of their existing interpersonal relationships and form new ones are experiencing extraordinary growth in Web 2.0. A large portion of the traffic on shopping sites comes from social networking websites.

The term network effects refers to the increased value of a network as its number of users grow. The more sellers and buyers on eBay the more valuable the site becomes to users. The more people on FaceBook the more friends for users to access.

http://visual.ly/social-media-market-share-2013

Web 2.0

Social Media:

Social media refers to any media shared online (e.g. Videos, music, photos, news etc). Consumer generated content has increased exponentially with the use of webcams and mobile phones.

YouTube – leading Internet video site. The entire site is based on user generated content.
Visual.ly – visual content
Spotify - music
Internet TV – NetFlix, Ustream
Digg – features news, videos an podcasts, all posted and rated by users.
Podcasting – digital audio file

Web 2.0

Tagging:

Tagging or labelling content, is part of the collaborative nature of Web 2.0. A tag is a user generated word or phrase that helps organise the web content and label it in a more human way.

Tag cloud – visual display of tags weighted by popularity. Many Web 2.0 sites include a graphical representation of popular tags, the size of tag text indicates its popularity, or a tag index is displayed.

Flickr – uses tagging as a way of organising personal photo collections. Users can search for photos by meaningful tags.

#tag in twitter

Stackoverflow.com uses tags

Web 2.0

Social Bookmarking:

Social bookmarking lets you share your Internet bookmarks through a website. Users can access these bookmarks from any computer and discover new sites by searching popular bookmarks and tags.

Delicious and stumbleupon are two examples.

Web 2.0

Rich Internet Applications (RIAs):

Rich Internet Applications are web applications that offer the responsiveness, rich features, and functionality of desktop applications. Early Internet applications supported only a basic HTML interface. They did not have the look and feel of a desktop application.
RIAs are as a result of more advanced technologies that allow greater responsiveness and advanced interfaces.

AJAX: Asynchronous JavaScript and XML allows for parts of a web page to be updated without having to reload the entire web page. This creates a more responsive interface allowing users to interact with the web page as the server processes requests. The technologies that make up AJAX are XHTML, CSS, JavaScript, the DOM and XML. AJAX forms a vital role in Web 2.0.

Web 2.0

Web Services, Mashups, Widgets and Gadgets:

Instead of re-inventing the wheel with every new project, developers can use existing companies' web services to create feature-rich applications.
Incorporating web services into new programs allows people to develop new applications quickly.
Application programming interfaces (API) are now offered through web services to encourage use of their services and data in the development of mashups, widgets and gadgets.
A mashup combines content or functionality from existing web services, websites and RSS feeds to serve a new purpose.

Web 2.0

Web Services, Mahups, Widgets and Gadgets:

Widgets, also referred to as gadgets, are mini applications designed to run either as standalone applications or as add-on features in web pages. Wordpress.com for example offers lots of widgets to add to your blog page. Widgipedia holds a catalog of widgets and gadgets for a variety of platforms. Amazon http://aws.amazon.com/ is a leading provider of web services, providing historical pricing data and e-commerce services. They offer storage services and queue services which allow businesses to pay for only the processing and storage space needed during any given period. This makes it possible for companies to save money while still being able to scale their storage and computing power to handle traffic surges.

Web 2.0

- For example, http://housingmaps.com is a mashup of Google maps and real-estate listings from Craigslist.

- www.trendsmap.com is a mashup of twitter and google maps.

- Mashups with maps are particularly popular, as are mashups using RSS feeds created using services such as Yahoo! Pipes.

- The mashup is reliant on one or more third parties, always check the terms of service when using web services.

Web 2.0

Location Based Services:

Location based services are applications that take your geographic location into consideration. While the term generally refers to services accessed on mobile devices using the Global Positioning System (GPS), it can also be used to describe web applications that take your location into account.

Search engines use localisation to provide you with geographically relevant content.

Google maps is one of the most popular mapping applications online. You can use Google maps to locate businesses in your area, get driving directions and live traffic information, create custom maps with images and more.

What other sites use location based services?

Web 2.0

RSS, Atom, JSON and VoIP:

Sites that offer RSS and Atom feeds can maintain an open connection with their readers. Users no longer have to regularly visit sites for updates – by subscribing to a site's feed, users receive updates as new information is posted to the site.

JavaScript Object Notation (JSON) was developed in 1999 as an alternative to XML, it is a text based data interchange format used to represent JavaScript objects as strings and transmit them over a network. JSON text is easy to produce and read.

Voice over Internet Protocol (VoIP) is the technology used to make free or inexpensive phone calls over the Internet.

Web 2.0

Web 2.0 Business Models:

The technologies and collaborative nature of Web 2.0 have opened up new business models. Some would not of been feasible ten years ago and now they are thriving.

Buying and selling domain names, blog, competitive intelligence, content network, discovery, file sharing, Internet radio, Internet TV, mashup, online advertising, online auction, online classifieds, online survey, open source, photo sharing, real estate, recommender system, reputation system, social networking, subscription site, travel site, vertical search engines, web analysis.