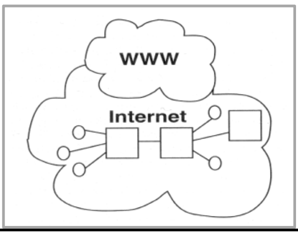
World Wide Web

The internet and the World Wide Web

- The concept of Internet and World Wide Web must be distinguished:
  - The Internet is the interconnection of networks managed by private or public bodies.
  - The World Wide Web (WWW) is an information space (i.e. a set of hw and sw entities that can be univocally addressed and a set of tools for their management) that can be accessed though the network.
- The WWW World Wide Web (WWW) is an internet service. Together with e-mail it is the most known and used service of the Internet. It offers a digital space for document publishing, software distribution and user-developed services.
- The WWW was created by Tim Berners-Lee when was a researcher at CERN in Geneve. Conventionally its starting date is assumed to be August 6th 1991, when he put the first website on-line on the Internet. World Wide Web is based on standards that are maintained by the World Wide Web Consortium (W3C).



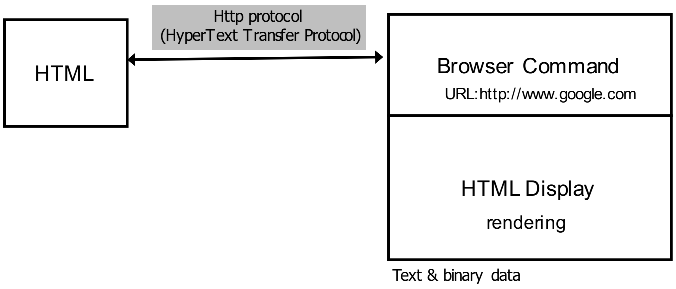
W3C WORLD WIDE WEB consortium

Main features

- There are several characteristics of the World Wide Web that have contributed to its success:
  - architecture based on public domain standards
  - client-server based architecture
  - capability of managing different media
- It is based on three main standards:
  - HTTP *Protocol* (*HyperText Transfer Protocol*) to communicate between the client and the server
  - *Addressing based on Uniform Resource Identifier* (URI) to refer to any addressable entity such as documents (text, images, sounds, etc.), programs.
  - HTML *Language* (*HyperText Markup Language*) to define web pages

HTTP Protocol (Hyper Text Transfer Protocol)

- HTTP Protocol (Hyper Text Transfer Protocol) allows the communication between a client (f.e. a browser) and a web server. It permits to transfer and manage data that are formatted according to the HTML language in a way that is independent from the system used.
- HTTP is an *ASCII* (8 bit) *protocol*, i.e. any HTTP message (i.e. a client request or control data by the server) is a string of ASCII characters. Data that are provided by the server are not necessarily ASCII data (f.e. they can be binary data of images, video…)


Http protocol
(HyperText Transfer Protocol)

HTML

Browser Command

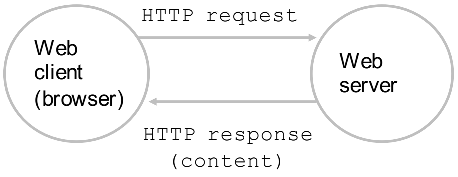
URL:http://www.google.com

HTML Display

rendering

Text & binary data

- Any HTTP interaction between the client and the server follows the schema:
  - client opens a transport connection between the client and the server using *TCP*
  - client sends a request with the URL of the resource requested
  - server sends a reply including the data from the requested URL
  - server closes the transport connection
- Because of the Connect, Request, Response, Disconnect nature of HTTP it is said to be a *stateless* protocol i.e. from one web page to the next there is nothing in the protocol that allows a web program to maintain program "state". Every interaction is independent from the others.



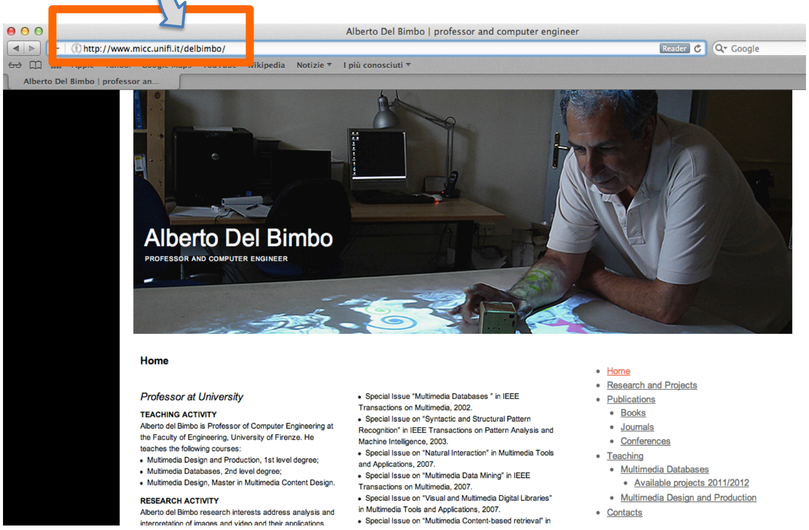
URL (Uniform Resource Locators)

- HTTP uses URI (Universal Resource Identifier) to refer to each entity. URI can be specified either by location (URL) or by name (URN),
- URL (Uniform Resource Locator) univocally defines the location of a resource in the network. Any URL is composed of three parts:
  - access method (specifies the way in which we want to access the resource)
  - *host* (specifies where the resource is located)
  - *identity* (specifies the resource identifier)
- Example: aaa://bbb: ccc/ddd/eee

Access method:
aaa = protocol (http, ftp, https, etc). As a default http is assumed.
Host:
bbb = host name (no default). Can be defined either as a domain address or as IP addres
ccc = port TCP that is used for transmission. If not specified the default port is used for the protocol that has been selected: 80 for http, 21 for ftp,443 for https.
Identity:
ddd = pathname. For http identifies a path from the *root* defined in the server.
eee = filename. If not specified, it corresponds to an index file that is configured in the server. As a default it is indicated as *index.html* or *index.htm*.

URL (Uniform Resource Locator) aaa://bbb: ccc/ddd/eee

http://www.micc.unifi.it [: default] /delbimbo/



Alberto Del Bimbo | professor and computer engineer

http://www.micc.unifi.it/delbimbo/

Alberto Del Bimbo

PROFESSOR AND COMPUTER ENGINEER

Home

Professor at University

TEACHING ACTIVITY

Alberto del Bimbo is Professor of Computer Engineering at the Faculty of Engineering, University of Firenze. He teaches the following courses:

- Multimedia Design and Production, 1st level degree;
- Multimedia Databases, 2nd level degree;
- Multimedia Design, Master in Multimedia Content Design.

RESEARCH ACTIVITY

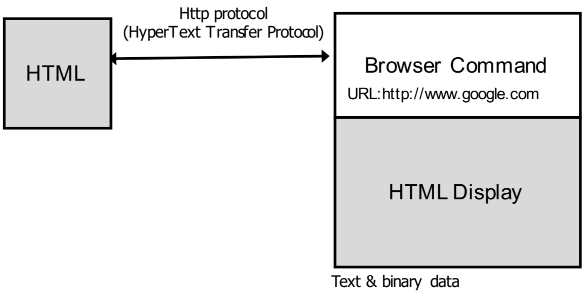
Alberto del Bimbo research interests address analysis and interpretation of images and video and their applications.

- Special Issue "Multimedia Databases " in IEEE Transactions on Multimedia, 2002.
- Special Issue on "Syntactic and Structural Pattern Recognition" in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003.
- Special Issue on "Natural Interaction" in Multimedia Tools and Applications, 2007.
- Special Issue on "Multimedia Data Mining" in IEEE Transactions on Multimedia, 2007.
- Special Issue on "Visual and Multimedia Digital Libraries" in Multimedia Tools and Applications, 2007.
- Special Issue on "Multimedia Content-based retrieval" in

- Home
- Research and Projects
- Publications
  - Books
  - Journals
  - Conferences
- Teaching
  - Multimedia Databases
    - Available projects 2011/2012
  - Multimedia Design and Production
- Contacts

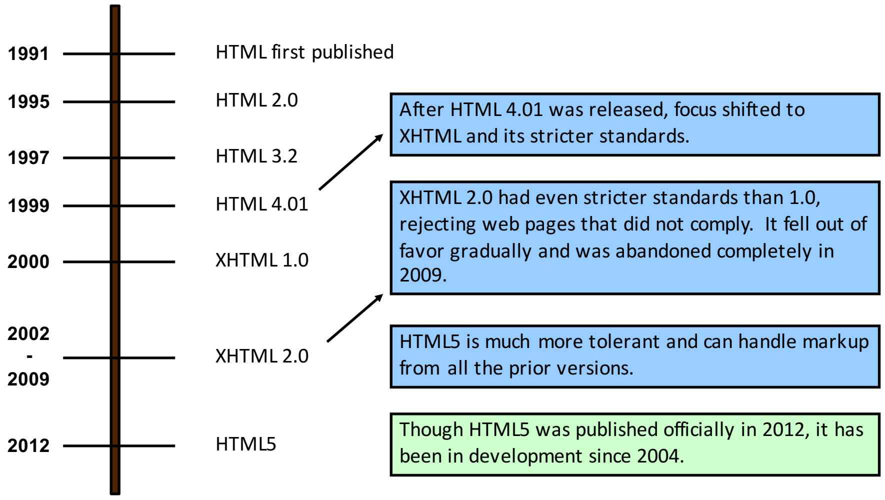
HTML (Hyper Text Markup Language)

- HTML (Hyper Text Markup Language) is a markup language that is used to define the web document format. The term markup identifies a sequence of characters and symbols (tags) that are inserted in a document in order to indicate to a browser program how the content must be displayed or the logical structure of the document. Formatting commands are explicitely inserted in the document text. An HTML document is an ASCII file



History of HTML

HTML evolution is under the control of the W3C consortium



1991 HTML first published

1995 HTML 2.0

1997 HTML 3.2

1999 HTML 4.01

2000 XHTML 1.0

2002 - 2009 XHTML 2.0

2012 HTML5

After HTML 4.01 was released, focus shifted to XHTML and its stricter standards.

XHTML 2.0 had even stricter standards than 1.0, rejecting web pages that did not comply. It fell out of favor gradually and was abandoned completely in 2009.

HTML5 is much more tolerant and can handle markup from all the prior versions.

Though HTML5 was published officially in 2012, it has been in development since 2004.

- HTML defines the way in which a web page (also referred to as HTML page) should appear. This is obtained through appropriate tags included in the text.
- As the client receives an HTML page the following operations are performed:
  - tags are interpreted
  - the page is formatted according to the tags and is adapted to the client constrains (screen resolution, window size…)
  - the page is displayed
- HTML tags are of one of two types:
  - tags for text formatting
  - tags for other purposes (user interaction…)

A few HTML tags

```
<head>  <title>Title of document goes here</title>
</head>

<body>   Visible text goes here… </body>

Basic Tags
<h1>Largest Heading</h1>
<h2> . . . </h2>
<h3> . . . </h3>
<h4> . . . </h4>
<h5> . . . </h5>
<h6>Smallest Heading</h6>

<p>This is a paragraph.</p>
<br> (line break)
<!-- This is a comment -->

<b>Bold text</b>
<em>Emphasized text</em>
<i>Italic text</i>
<small>Smaller text</small>
<strong>Important text</strong>
```

……….

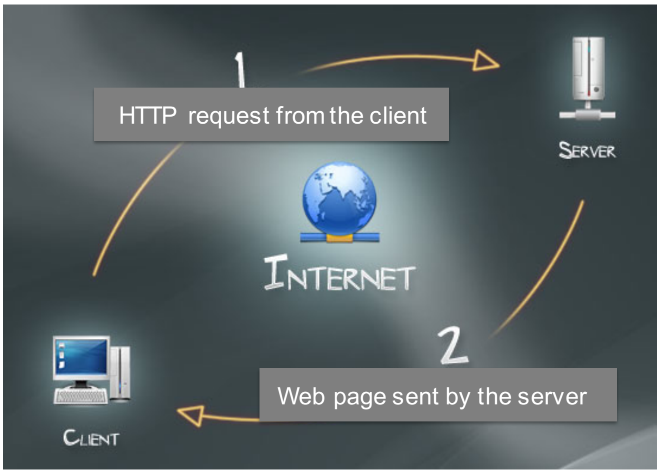
XML, HTML, HTML5

- XML is a standard metalanguage: "*a common syntax for expressing structure in data*".
  It is used to define new markup languages and permits to create your own structured and personalized documents.
- XML defines a separation between data definition and data presentation in order to ease document exchange between applications. According to this with XML it is possible to validate a data structure that has been defined by the user: the user agent assumes that XML data are structured according to the standard specifications.

- XML is an official specification of the World Wide Web Consortium (W3C) and is the common ground of all technical specifications that have been released by the W3C.

- HTML + XML = XHTML
  XHTML is very similar to HTML, but adopts the XML syntax to formally describe the parts that compose a document. Surpassed by HTML5

Client – Server dialogue



Server (Server Agents)

- A server is a process that executes on a computer system. Typically it:
  - listens for requests by a client agent
  - operates in order to satisfy any requests received

- A server should be able to satisfy multiple requests and at the same time continue to listen for new requests. This is performed in one of two ways:
  - *Cloned server*: as a new request arrives at the server, the server creates a server replica that is dedicated to satisfy the new request. The server then returns to listen for new requests.
  - *Multithreaded server* : only one copy of the server exists that is designed so to generate multiple threads.

Client (User Agents)

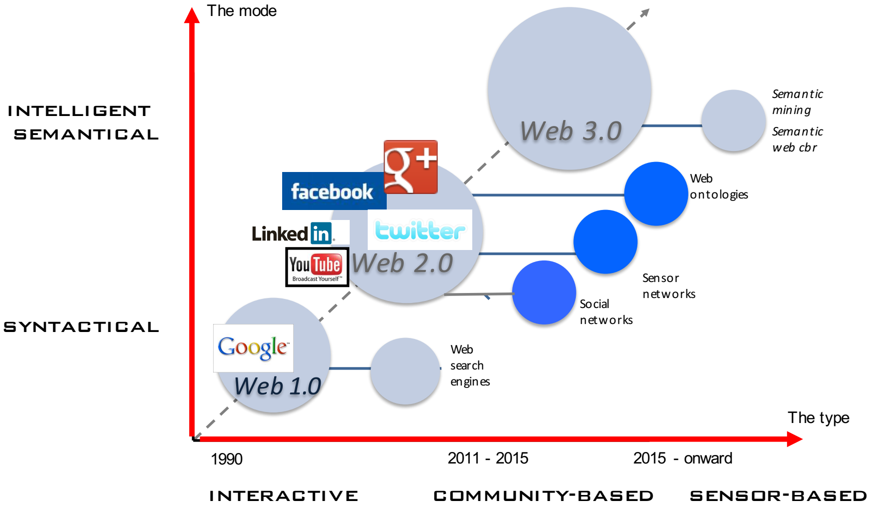
- Clients or user agents permit to the final user access and navigation on the web. They are commonly referred to as *Browsers*: *Microsoft Explorer, Firefox, Safari, Google Chrome…..*

- Among the Browser's duties:
  - Send requests for data to the server
  - Receive data from the server
  - Support visualization of the web page requested
  - Permit operations on the data received

- To extend their functions, browsers use plug-ins, i.e. specialized libraries of executable code that are loaded in memory on request. Among the new features provided by plug-ins are fe. search-engines, virus scanners, or the ability to manage new video formats.

- Well-known browser plug-ins include QuickTime Player and the Java plug-in, which can launch a user-activated Java applet on a web page

THE WEB

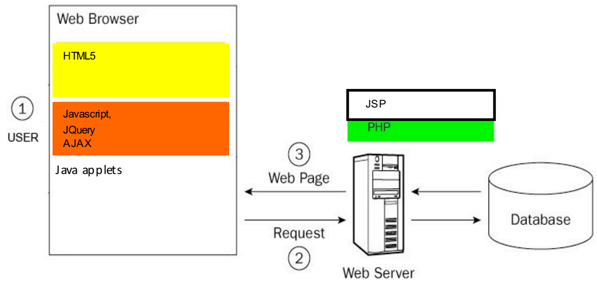


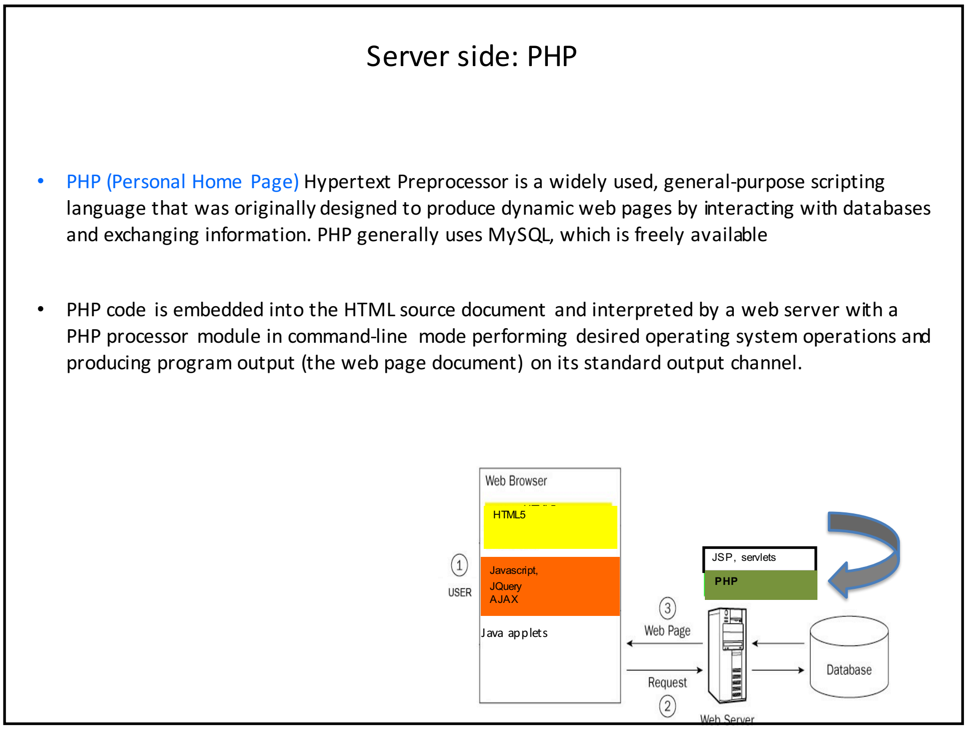
Web 2.0

- Web 1.0 was conceived only for web page browsing. With Web 2.0 Browser new functions were added to the sole presentation of web pages that permit a more complete user interactivity, such as access to remote repositories…
- Web pages in this framework are referred to as dynamic web pages as opposite to the static web pages of the Web 1.0

Web 2.0 main technologies

- A Client-Server application in the Web 2.0 framework can exploit functions that are executed either at the Server or the Client side:
  - Server Side
    - PHP scripting language
    - JSP Java Server Pages
  - Client Side:
    - HTML5
    - JavaScript, JQuery
    - Ajax
    - Java applets



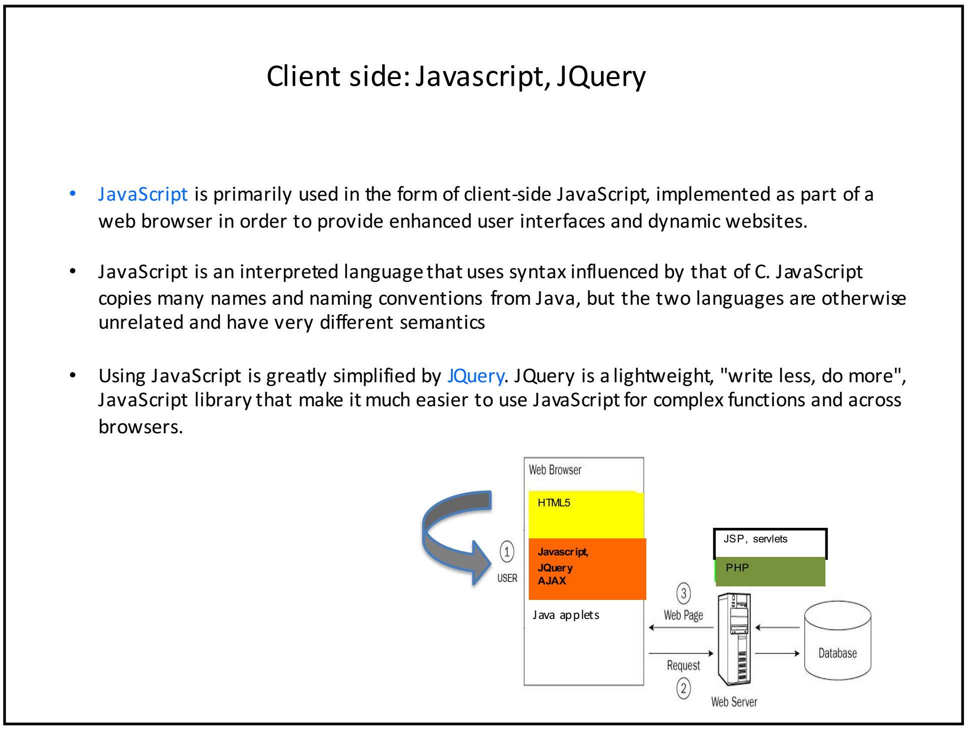


Server side: PHP

- PHP (Personal Home Page) Hypertext Preprocessor is a widely used, general-purpose scripting language that was originally designed to produce dynamic web pages by interacting with databases and exchanging information. PHP generally uses MySQL, which is freely available

- PHP code is embedded into the HTML source document and interpreted by a web server with a PHP processor module in command-line mode performing desired operating system operations and producing program output (the web page document) on its standard output channel.

Web Browser
HTML5
Javascript,
JQuery
AJAX
Java applets
1
USER
JSP, servlets
PHP
3
Web Page
Request
2
Database
Web Server



Client side: Javascript, JQuery

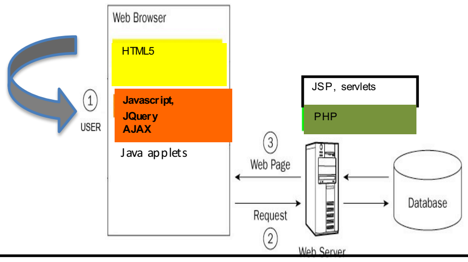
- JavaScript is primarily used in the form of client-side JavaScript, implemented as part of a web browser in order to provide enhanced user interfaces and dynamic websites.

- JavaScript is an interpreted language that uses syntax influenced by that of C. JavaScript copies many names and naming conventions from Java, but the two languages are otherwise unrelated and have very different semantics

- Using JavaScript is greatly simplified by JQuery. JQuery is a lightweight, "write less, do more", JavaScript library that make it much easier to use JavaScript for complex functions and across browsers.

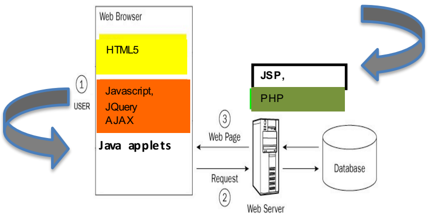
Client side : Ajax

- Ajax (Asynchronous JavaScript and XML), is a group of interrelated web development techniques used to create interactive web applications.
- With Ajax, web applications can retrieve data from the server asynchronously in the background without interfering with the display and behavior of the existing page. In other words the page is not reloaded. Typically functions that are requested are written in Javascript language.
- The use of Ajax has led to an increase in interactive animation on web pages. Data is retrieved using the XML HttpRequest object or through the use of Remote Scripting in browsers that do not support it. Ajax is a multi-platform technique that can be used on many operating systems and web browsers, with many open source implementations.



JSP, Java applets

- Server side: JSP (Java Server Pages) allows Java code to be interleaved with static web markup content, with the resulting page being compiled and executed on the server to deliver an HTML document. It is an alternative solution to PHP for dynamic web pages.
    - JSP uses predefined tags inserted in the HTML code and call for predefined functions in Java code. The compiled pages must be executed within a Java Virtual Machine (JVM) that integrates with the host operating system.
- Client side: Java applets are used to provide interactive features to web applications that cannot be provided by HTML alone. In response to the user action an applet can change the provided graphic content.
    - Java applets can run in a Web browser using a Java Virtual Machine.

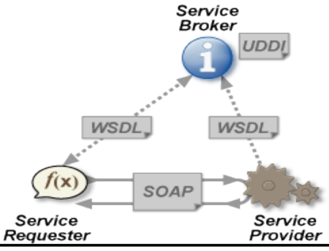


Interoperability: web services

- HTTP, originally designed for human-to-machine communication, is also utilized for machine-to-machine communication, more specifically for transferring machine readable file formats

- A Web service is defined by the W3C as "*a software system designed to support interoperable machine-to-machine interaction over a network*". It is a software function provided at a network address over the Web with the service *always on* as in the concept of utility computing.

- According to W3C, web service standardizes:
  - A directory called UDDI (Universal Description, Discovery and Integration), that defines which software system should be contacted for which type of data.
  - the interface to the service, described with a machine-readable format (the Web Services Description Language WSDL);
  - the way in which data are sent, through XML messages that are included in an envelope (the SOAP) and transmitted with the http protocol



Web 2.0 distinguishing applications

- Web 2.0 main distinguishing applications are:
  - *Traditional*: RSS, Blogs, Wikis, Multimedia sharing, Folksonomy, Audio blogging and podcasting, …
  - *Social Networking* : Facilitates meeting people, finding like-minds and sharing content
  - *Collaborative*: Collaborative reference works (like Wikipedia) using wiki-like software tools ("*everyone is an expert on something*").
  - *Aggregation and mash-up services*: Gather information from diverse sources across the Web and publish in one place. Pull together data from different sources to create a new service. Tracking and filtering content….
  - *Replicate office-style software in the browser*: Web-based desktop application/document tools. Replicate desktop applications. Based on technological developments.

From Web 1.0 to Web 2.0

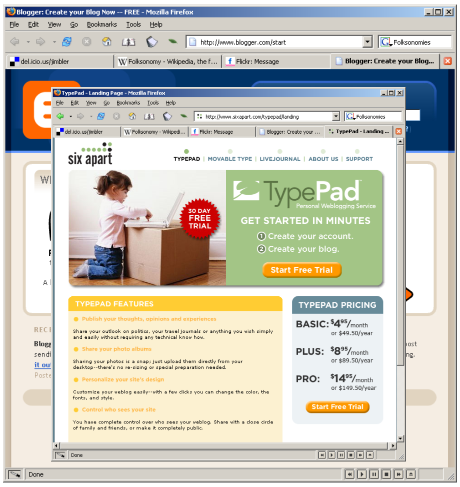
Web 1.0 (1993-2003) Pretty much HTML pages viewed through a browser		Web 2.0 (2003- beyond) Web pages, plus a lot of other "content" shared over the web, with more interactivity; more like an application than a "page"
"Read"	**Mode**	"Write" & Contribute
"Page"	**Primary Unit of content**	"Post / record"
"static"	**State**	"dynamic"
Web browser	**Viewed through…**	Browsers, RSS Readers, anything
"Client Server"	**Architecture**	"Web Services"
Web Coders	**Content Created by…**	Everyone
"geeks"	**Domain of…**	"mass amatuerization"

RSS

- RSS (Really Simple Syndication) is a family of web feed formats used to publish frequently updated works such as news headlines, blog entries, audio, and video—in a standardized format.
- An RSS document called a "*feed*" includes summarized text, plus metadata such as publishing dates and authorship. A standardized XML file format allows the information to be published once and viewed by many different programs.
- RSS feeds can be read using software called an "*RSS reader*" which can be web-based, desktop-based, mobile device or any computerized Internet-connected device.
  - The user subscribes to a feed by entering the feed's URI into the reader or by clicking an RSS icon in a browser that initiates the subscription process.
  - The RSS reader checks the user's subscribed feeds regularly for new work, downloads any updates that it finds, and provides a user interface to monitor and read the feeds.
- Web feeds benefit publishers by letting them syndicate content automatically. They benefit readers who want to subscribe to timely updates from favored websites or to aggregate feeds from many sites into one place.

Blog

- A blog (*web log*) is a discussion or informational site published on the World Wide Web and consisting of discrete entries ("*posts*") typically displayed in reverse chronological order.


Blogger: Create your Blog Now -- FREE - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.blogger.com/start
Folksonomies
del.icio.us/jimbler
Folksonomy - Wikipedia, the f...
Flickr: Message
Blogger: Create your Blog...
TypePad - Landing Page - Mozilla Firefox
http://www.sixapart.com/typepad/landing
TypePad - Landing ...
six apart
TYPEPAD | MOVABLE TYPE | LIVEJOURNAL | ABOUT US | SUPPORT
30 DAY FREE TRIAL
TypePad
Personal Weblogging Service
Start Free Trial
Done

GET STARTED IN MINUTES

① Create your account.
② Create your blog.

TYPEPAD FEATURES

- Publish your thoughts, opinions and experiences

Share your outlook on politics, your travel journals or anything you wish simply and easily without requiring any technical know how.

- Share your photo albums

Sharing your photos is a snap; just upload them directly from your desktop--there's no re-sizing or special preparation needed.

- Personalize your site's design

Customize your weblog easily--with a few clicks you can change the color, the fonts, and style.

- Control who sees your site

You have complete control over who sees your weblog. Share with a close circle of family and friends, or make it completely public.

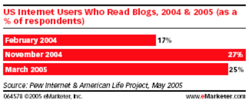
TYPEPAD PRICING

BASIC: \$4⁹⁵/month or \$49.50/year

PLUS: \$8⁹⁵/month or \$89.50/year

PRO: \$14⁹⁵/month or \$149.50/year

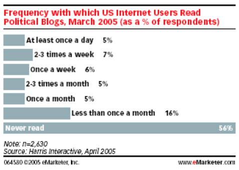
US Internet Users Who Read Blogs, 2004 & 2005 (as a % of respondents)



Source: Pew Internet & American Life Project, May 2005

064578 ©2005 eMarketer, Inc. www.eMarketer.com

Frequency with which US Internet Users Read Political Blogs, March 2005 (as a % of respondents)



Note: n=2,630
Source: Harris Interactive, April 2005

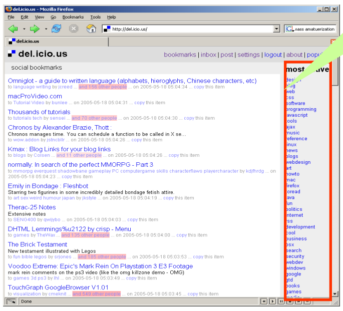
064580 ©2005 eMarketer, Inc. www.eMarketer.com

Folksonomy

- A "folksonomy" is a spontaneous, collaborative work to categorize links by a community of users. Users take control of organize the content together.

Tags: Descriptive words applied by users to links. Tags are searchable

My Tags: Words I've used to describe links in a way that makes sense to me



Omniglot - a guide to written language (alphabets, hieroglyphs, Chinese characters, etc)
to language writing by jcreed ... and 156 other people ... on 2005-05-18 05:04:34 ... copy this item

macProVideo.com
to Tutorial Video by bunlee ... on 2005-05-18 05:04:31 ... copy this item

Thousands of tutorials
to tutorials tech by sensei ... and 70 other people ... on 2005-05-18 05:04:30 ... copy this item

Chronos by Alexander Brazie, Thott :
Chronos manages time. You can schedule a function to be called in X se...
to wow.addon by jstncbllr ... on 2005-05-18 05:04:26 ... copy this item

Kmax : Blog Links for your blog links
to blogs by Colsen ... and 11 other people ... on 2005-05-18 05:04:26 ... copy this item

normally: In search of the perfect MMORPG - Part 3
to mmorpg everquest shadowbane gameplay PC computergame skills characterflaws playercharacter by kdjfhrdg ... on 2005-05-18 05:04:23 ... copy this item

Emily in Bondage : Fleshbot
Starring two figurines in some incredibly detailed bondage fetish attire.
to art sex weird humour japan by jkstyle ... on 2005-05-18 05:04:19 ... copy this item

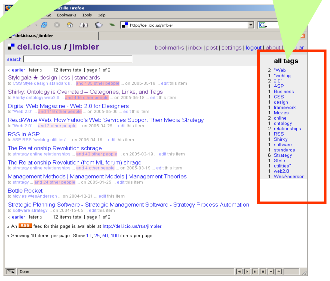
Therac-25 Notes
Extensive notes
to SENG400 by qwijybo ... on 2005-05-18 05:04:03 ... copy this item

DHTML Lemmings%u2122 by crisp - Menu
to games by TheWax ... and 135 other people ... on 2005-05-18 05:04:00 ... copy this item

The Brick Testament
New testament illustrated with Legos
to fun bible legos by srjones ... and 185 other people ... on 2005-05-18 05:03:53 ... copy this item

Voodoo Extreme: Epic's Mark Rein On Playstation 3 E3 Footage
mark rein comments on the ps3 video (like the omg killzone demo - OMG)
to games 3d ps3 by lhl ... on 2005-05-18 05:03:49 ... copy this item

TouchGraph GoogleBrowser V1.01
to visualization by cmeknit ... and 549 other people ... on 2005-05-18 05:03:45 ... copy this item



« earlier | later » 12 items total | page 1 of 2

Stylegala ★ design | css | standards
to CSS Style design standards ... and 135 other people ... on 2005-05-18 ... edit this item

Shirky: Ontology is Overrated -- Categories, Links, and Tags
to Shirky ontology web2.0 ... and 405 other people ... on 2005-05-18 ... edit this item

Digital Web Magazine - Web 2.0 for Designers
to "Web 2.0" ... and 119 other people ... on 2005-05-06 ... edit this item

Read/Write Web: How Yahoo's Web Services Support Their Media Strategy
to "Web 2.0" ... and 3 other people ... on 2005-04-29 ... edit this item

RSS in ASP
to ASP RSS "weblog utilities" ... on 2005-04-16 ... edit this item

The Relationship Revolution schrage
to strategy online relationships ... and 43 other people ... on 2005-03-19 ... edit this item

The Relationship Revolution (from ML forum) shrage
to strategy online relationships ... and 4 other people ... on 2005-03-19 ... edit this item

Management Methods | Management Models | Management Theories
to strategy ... and 24 other people ... on 2005-01-25 ... edit this item

Bottle Rocket
to Movies WesAnderson ... on 2004-12-21 ... edit this item

Strategic Planning Software - Strategic Management Software - Strategy Process Automation
to software strategy ... on 2004-12-05 ... edit this item

« earlier | later » 12 items total | page 1 of 2

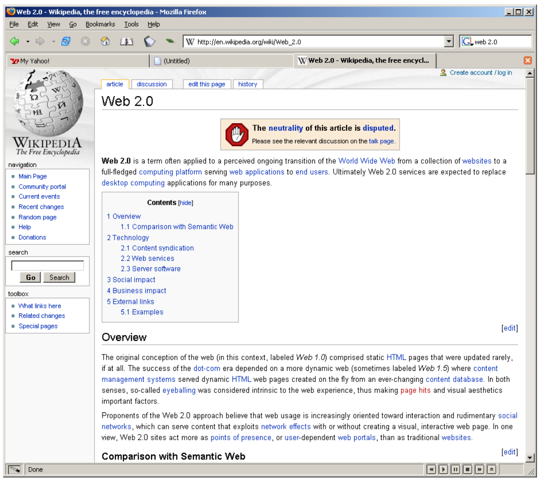
» An RSS feed for this page is available at http://del.icio.us/rss/jimbler.

» Showing 10 items per page. Show 10, 25, 50, 100 items per page.

Del.icio.us is an example of Folksonomy to organize bookmarks

Wiki

- A wiki is a website which allows its users to add, modify, or delete its content via a web browser usually using a simplified markup language or a rich-text editor. Wikis are powered by wiki software. Most are created collaboratively.



Web 2.0

The neutrality of this article is disputed.
Please see the relevant discussion on the talk page.

Web 2.0 is a term often applied to a perceived ongoing transition of the World Wide Web from a collection of websites to a full-fledged computing platform serving web applications to end users. Ultimately Web 2.0 services are expected to replace desktop computing applications for many purposes.

Contents [hide]
1 Overview
1.1 Comparison with Semantic Web
2 Technology
2.1 Content syndication
2.2 Web services
2.3 Server software
3 Social impact
4 Business impact
5 External links
5.1 Examples

[edit]
Overview

The original conception of the web (in this context, labeled *Web 1.0*) comprised static HTML pages that were updated rarely, if at all. The success of the dot-com era depended on a more dynamic web (sometimes labeled *Web 1.5*) where content management systems served dynamic HTML web pages created on the fly from an ever-changing content database. In both senses, so-called eyeballing was considered intrinsic to the web experience, thus making page hits and visual aesthetics important factors.

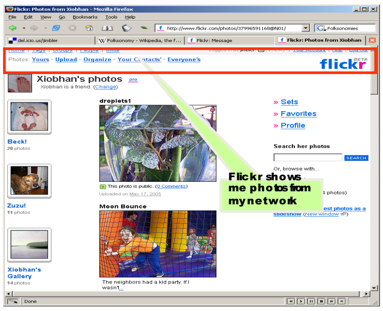
Proponents of the Web 2.0 approach believe that web usage is increasingly oriented toward interaction and rudimentary social networks, which can serve content that exploits network effects with or without creating a visual, interactive web page. In one view, Web 2.0 sites act more as points of presence, or user-dependent web portals, than as traditional websites.

[edit]
Comparison with Semantic Web

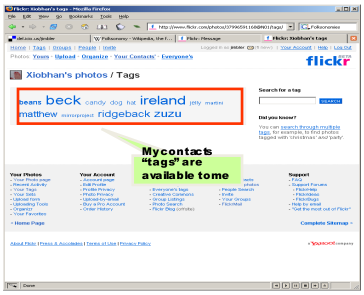
Wikipedia is a Collaborative Dictionary

Social networks

- A social networking service is a web-based platform to build social relations among people who, for example, share interests, activities, backgrounds...It consists of a representation of each user, his/her social links, and a variety of additional services, such as e-mail and instant messaging among the others.


Flickr: Photos from Xiobhan - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.flickr.com/photos/37996591168@N01/
Folksonomies
del.icio.us/jimbler
Folksonomy - Wikipedia, the f...
Flickr: Message
Flickr: Photos from Xiobhan
Photos: Yours · Upload · Organize · Your Contacts' · Everyone's
flickr BETA
Xiobhan's photos pro
Xiobhan is a friend. (Change)
droplets1
This photo is public. (0 Comments)
Uploaded on May 17, 2005
Moon Bounce
The neighbors had a kid party. If I wasn't...
Beck!
20 photos
Zuzu!
11 photos
Xiobhan's Gallery
14 photos
» Sets
» Favorites
» Profile
Search her photos
SEARCH
Or, browse with...
Done

Flickr shows me photos from my network



My contacts "tags" are available to me

Flickr combines a social network with user generated content. Users can work together to collaborate on photo projects and use each others' tags to find new photos. Flickr also has an API for web services to integrate photo collections with blogs and other apps.

Web 3.0

- Web 3.0: “the sensor-web, in which the architecture of participation will be an automatic byproduct of the devices we carry around with us.”

- Not providing semantics in the links is one of the main navigational problems of the World Wide Web: It is not until one opens the destination page of a link that one finds out that its content is not of interest.

- Semantic Concept Extraction and Ontology Building is Web 3.0