

World Wide Web

From Wikipedia, the free encyclopedia

*"The Web" and "WWW" redirect here. For other uses, see Web and WWW (disambiguation).
For the world's first browser, see WorldWideWeb.*

The **World Wide Web** (or the "**Web**") is a system of interlinked, hypertext documents accessed via the Internet. With a Web browser, a user views Web pages that may contain text, images, and other multimedia and navigates between them using hyperlinks. The Web was created around 1990 by the Englishman Sir Tim Berners-Lee and the Belgian Robert Cailliau working at CERN in Geneva, Switzerland. Since then, Berners-Lee has played an active role in guiding the development of Web standards (such as the markup languages in which Web pages are composed), and in recent years has advocated his vision of a Semantic Web.



Contents

- 1 How the Web works
- 2 Caching
- 3 History
- 4 Standards
- 5 Java and JavaScript
- 6 Sociological implications
- 7 Publishing Web pages
- 8 Statistics
- 9 Speed issues
- 10 Link rot and Web archival
- 11 Academic conferences
- 12 WWW prefix in Web addresses
- 13 Pronunciation of "www"
- 14 See also
- 15 References
- 16 External links

How the Web works

Viewing a Web page or other resource on the World Wide Web normally begins either by typing the URL of the page into a Web browser, or by following a hypertext link to that page or resource. The first step, behind the scenes, is for the server-name part of the URL to be resolved into an IP address by the global, distributed Internet database known as the Domain name system or DNS. The browser then establishes a TCP connection with the server at that IP address.

The next step is for an HTTP request to be sent to the Web server, requesting the resource. In the case of a typical Web page, the HTML text is first requested and parsed by the browser, which then makes additional requests for graphics and any other files that form a part of the page in quick succession. When considering web site popularity statistics, these additional file requests give rise to the difference between one single 'page view' and an associated number of server 'hits'.

The Web browser then renders the page as described by the HTML, CSS and other files received,

incorporating the images and other resources as necessary. This produces the on-screen page that the viewer sees.

Most Web pages will themselves contain hyperlinks to other related pages and perhaps to downloads, source documents, definitions and other Web resources.

Such a collection of useful, related resources, interconnected via hypertext links, is what has been dubbed a 'web' of information. Making it available on the Internet created what Tim Berners-Lee first called the **WorldWideWeb** (note the name's use of CamelCase, subsequently discarded) in 1990.^[1]

Caching

If the user returns to a page fairly soon, it is likely that the data will not be retrieved from the source Web server, as above, again. By default, browsers cache all web resources on the local hard drive. An HTTP request will be sent by the browser that asks for the data *only if it has been updated since the last download*. If it has not, the cached version will be reused in the rendering step.

This is particularly valuable in reducing the amount of Web traffic on the Internet. The decision about expiration is made independently for each resource (image, stylesheet, JavaScript file etc., as well as for the HTML itself). Thus even on sites with highly dynamic content, many of the basic resources are only supplied once per session or less. It is worth it for any Web site designer to collect all the CSS and JavaScript into a few site-wide files so that they can be downloaded into users' caches and reduce page download times and demands on the server.

There are other components of the Internet that can cache Web content. The most common in practice are often built into corporate and academic firewalls where they cache web resources requested by one user for the benefit of all. Some search engines such as Google or Yahoo! also store cached content from Web sites.

Apart from the facilities built into Web servers that can ascertain when physical files have been updated, it is possible for designers of dynamically generated web pages to control the HTTP headers sent back to requesting users, so that pages are not cached when they should not be — for example Internet banking and news pages.

This helps with understanding the difference between the HTTP 'GET' and 'POST' verbs - data requested with a GET may be cached, if other conditions are met, whereas data obtained after POSTing information to the server usually will not.

History

See also: History of the Internet - CERN, Where the Web Was "WWW" born
(<http://info.cern.ch/>)

The underlying ideas of the Web can be traced as far back as 1980, when, at CERN in Switzerland, the Englishman Tim Berners-Lee built ENQUIRE (referring to *Enquire Within Upon Everything*, a book he recalled from his youth). While it was rather different from the Web in use today, it contained many of the same core ideas (and even some of the ideas of Berners-Lee's next project after the WWW, the Semantic Web).

In March 1989, Tim Berners-Lee wrote a proposal^[2], which referenced ENQUIRE and described a

more elaborate information management system. With help from Robert Cailliau, he published a more formal proposal for the World Wide Web^[3] on November 12, 1990.

A NeXTcube was used by Berners-Lee as the world's first web server and also to write the first web browser, WorldWideWeb in 1990. By Christmas 1990, Berners-Lee had built all the tools necessary for a working Web:^[4] the first Web browser (which was a Web editor as well), the first Web server and the first Web pages^[5] which described the project itself.



This NeXTcube used by Berners-Lee at CERN became the first Web server.

On August 6, 1991, he posted a short summary of the World Wide Web project on the `alt.hypertext` newsgroup^[6]. This date also marked the debut of the Web as a publicly available service on the Internet.

The crucial underlying concept of hypertext originated with older projects from the 1960s, such as Ted Nelson's Project Xanadu and Douglas Engelbart's oN-Line System (NLS). Both Nelson and Engelbart were in turn inspired by Vannevar Bush's microfilm-based "memex," which was described in the 1945 essay "As We May Think".

Berners-Lee's breakthrough was to marry hypertext to the Internet. In his book *Weaving The Web*, he explains that he had repeatedly suggested that a marriage between the two technologies was possible to members of *both* technical communities, but when no one took up his invitation, he finally tackled the project himself. In the process, he developed a system of globally unique identifiers for resources on the Web and elsewhere: the Uniform Resource Identifier.

The World Wide Web had a number of differences from other hypertext systems that were then available:

- The WWW required only unidirectional links rather than bidirectional ones. This made it possible for someone to link to another resource without action by the owner of that resource. It also significantly reduced the difficulty of implementing Web servers and browsers (in comparison to earlier systems), but in turn presented the chronic problem of link rot.
- Unlike predecessors such as HyperCard, the World Wide Web was non-proprietary, making it possible to develop servers and clients independently and to add extensions without licensing restrictions.

On April 30, 1993, CERN announced^[7] that the World Wide Web would be free to anyone, with no fees due. Coming two months after the announcement that gopher was no longer free to use, this produced a rapid shift away from gopher and towards the Web. An early popular Web browser was ViolaWWW which was based upon HyperCard.

Scholars generally agree, however, that the turning point for the World Wide Web began with the introduction^[8] of the Mosaic web browser^[9] in 1993, a graphical browser developed by a team at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign (NCSA-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*.^[10] Prior to the release of Mosaic, graphics were not commonly mixed with text in Web pages and its popularity was less than older protocols in use over the Internet, such as Gopher protocol and Wide area information server. Mosaic's graphical user interface allowed the Web to become by far the most popular Internet protocol.

Standards

Many formal standards and other technical specifications define the operation of different aspects of the World Wide Web, the Internet, and computer information exchange. Many of the documents are the work of the World Wide Web Consortium (W3C), headed by Berners-Lee, but some are produced by the Internet Engineering Task Force (IETF) and other organizations.

Usually, when Web standards are discussed, the following publications are seen as foundational:

- Recommendations for markup languages, especially HTML and XHTML, from the W3C. These define the structure and interpretation of hypertext documents.
- Recommendations for stylesheets, especially CSS, from the W3C.
- Standards for ECMAScript, a.k.a. JavaScript, from Ecma International.
- Recommendations for the Document Object Model, from W3C.

Additional publications provide definitions of other essential technologies for the World Wide Web, including, but not limited to, the following:

- *Uniform Resource Identifier (URI)*, which is a universal system for referencing resources on the Internet, such as hypertext documents and images. URIs, often called URLs, are defined by the IETF's RFC 3986 (<http://tools.ietf.org/html/rfc3986>) / STD 66: *Uniform Resource Identifier (URI): Generic Syntax*, as well as its predecessors and numerous URI scheme-defining RFCs;
- *HyperText Transfer Protocol (HTTP)*, especially as defined by RFC 2616 (<http://tools.ietf.org/html/rfc2616>): *HTTP/1.1* and RFC 2617 (<http://tools.ietf.org/html/rfc2617>): *HTTP Authentication*, which specify how the browser and server communicate with each other.

Java and JavaScript

A significant advance in Web technology was Sun Microsystems' Java platform. It enables Web pages to embed small programs (called applets) directly into the view. These applets run on the end-user's computer, providing a richer user interface than simple web pages. Java client-side applets never gained the popularity that Sun had hoped for, for a variety of reasons including lack of integration with other content (applets were confined to small boxes within the rendered page) and the fact that many computers at the time were supplied to end users without a suitably installed JVM, and so required a download by the user before applets would appear. Adobe Flash now performs many of the functions that were originally envisioned for Java applets including the playing of video content, animation and some rich UI features. Java itself has become more widely used as a platform and language for server-side and other programming.

JavaScript, on the other hand, is a scripting language that was initially developed for use within Web pages. The standardized version is ECMAScript. While its name is similar to Java, JavaScript was developed by Netscape and it has almost nothing to do with Java, apart from that, like Java, its syntax is derived from the C programming language. In conjunction with a Web page's Document Object Model, JavaScript has become a much more powerful technology than its creators originally envisioned. The manipulation of a page's Document Object Model after the page is delivered to the client has been called Dynamic HTML (DHTML), to emphasize a shift away from *static* HTML displays.

In its simplest form, all the optional information and actions available on a JavaScripted Web page

will have been downloaded when the page was first delivered. Ajax ("Asynchronous JavaScript And XML") is a JavaScript-based technology that may have a significant effect on the development of the World Wide Web. Ajax provides a method whereby large or small parts *within* a Web page may be updated, using new information obtained over the network in response to user actions. This allows the page to be much more responsive, interactive and interesting, without the user having to wait for whole-page reloads. Ajax is seen as an important aspect of what is being called Web 2.0. Examples of Ajax techniques currently in use can be seen in Gmail, Google Maps etc.

Sociological implications



This article or section may contain original research or unverified claims.

Please help Wikipedia by adding references. See the talk page for details.

The Web, as it stands today, has allowed global interpersonal exchange on a scale unprecedented in human history. People separated by vast distances, or even large amounts of time, can use the Web to exchange—or even mutually develop—their most intimate and extensive thoughts, or alternately their most casual attitudes and spirits. Emotional experiences, political ideas, cultural customs, musical idioms, business advice, artwork, photographs, literature, can all be shared and disseminated digitally with less individual investment than ever before in human history. Although the existence and use of the Web relies upon material technology, which comes with its own disadvantages, its information does not use physical resources in the way that libraries or the printing press have. Therefore, propagation of information via the Web (via the Internet, in turn) is not constrained by movement of physical volumes, or by manual or material copying of information. By virtue of being digital, the information of the Web can be searched more easily and efficiently than any library or physical volume, and vastly more quickly than a person could retrieve information about the world by way of physical travel or by way of mail, telephone, telegraph, or any other communicative medium.

The Web is the most far-reaching and extensive medium of personal exchange to appear on Earth. It has probably allowed many of its users to interact with many more groups of people, dispersed around the planet in time and space, than is possible when limited by physical contact or even when limited by every other existing medium of communication combined.

Because the Web is global in scale, some have suggested that it will nurture mutual understanding on a global scale. By definition or by necessity, the Web has such a massive potential for social exchange, it has the potential to nurture empathy and symbiosis, but it also has the potential to incite belligerence on a global scale, or even to empower demagogues and repressive regimes in ways that were historically impossible to achieve previously.

Publishing Web pages

The Web is available to individuals outside mass media. In order to "publish" a Web page, one does not have to go through a publisher or other media institution, and potential readers could be found in all corners of the globe.

Unlike books and documents, hypertext does not need to have a linear order from beginning to end. It is not necessarily broken down into the hierarchy of chapters, sections, subsections, etc.

Many different kinds of information are now available on the Web, and for those who wish to know other societies, their cultures and peoples, it has become easier. When traveling in a foreign country

or a remote town, one might be able to find some information about the place on the Web, especially if the place is in one of the developed countries. Local newspapers, government publications, and other materials are easier to access, and therefore the variety of information obtainable with the same effort may be said to have increased, for the users of the Internet.

Although some Web sites are available in multiple languages, many are in the local language only. Additionally, not all software supports all special characters, and RTL languages. These factors would challenge the notion that the World Wide Web will bring a unity to the world.

The increased opportunity to publish materials is certainly observable in the countless personal pages, as well as pages by families, small shops, etc., facilitated by the emergence of free Web hosting services.

Statistics

According to a 2001 study,^[11] there were more than 550 million documents on the Web, mostly in the "invisible Web". A 2002 survey of 2,024 million Web pages^[12] determined that by far the most Web content was in English: 56.4%; next were pages in German (7.7%), French (5.6%) and Japanese (4.9%). A more recent study which used web searches in 75 different languages to sample the Web determined that there were over 11.5 billion web pages in the publicly indexable Web as of the end of January 2005.^[13]

Speed issues

Frustration over congestion issues in the Internet infrastructure and the high latency that results in slow browsing has led to an alternative name for the World Wide Web: the *World Wide Wait*. Speeding up the Internet is an ongoing discussion over the use of peering and QoS technologies. Other solutions to reduce the World Wide Wait can be found on W3C (<http://www.w3.org/Protocols/NL-PerfNote.html>).

Standard guidelines for ideal Web response times are (Nielsen 1999, page 42):

- 0.1 second (one tenth of a second). Ideal response time. The user doesn't sense any interruption.
- 1 second. Highest acceptable response time. Download times above 1 second interrupt the user experience.
- 10 seconds. Unacceptable response time. The user experience is interrupted and the user is likely to leave the site or system.

These numbers are useful for planning server capacity.

Link rot and Web archival

Over time, many Web resources pointed to by hyperlinks disappear, relocate, or are replaced with different content. This phenomenon is referred to in some circles as "link rot" and the hyperlinks affected by it are often called "dead links".

The ephemeral nature of the Web has prompted many efforts to archive the Web. The Internet Archive is one of the most well-known efforts; they have been archiving the Web since 1996.

Academic conferences

The major academic event covering the WWW is the World Wide Web series of conferences, promoted by IW3C2 (<http://www.iw3c2.org/>). There is a list (<http://www.iw3c2.org/conferences/>) with links to all conferences in the series.

WWW prefix in Web addresses

"www" is commonly found at the beginning of Web addresses because of the long-standing practice of naming Internet hosts (servers) according to the services they provide. So for example, the host name for a Web server is often "www"; for an FTP server, "ftp"; and for a USENET news server, "news" or "nntp" (after the news protocol NNTP). These host names appear as DNS subdomain names, as in "www.example.com".

This use of such prefixes is not required by any technical standard; indeed, the first Web server was at "nxoc01.cern.ch"^[14] and even today many Web sites exist without a "www" prefix. The "www" prefix has no meaning in the way the main website is shown. The "www" prefix is simply one choice for a Web site's subdomain name.

Some Web browsers will automatically try adding "www." to the beginning, and possibly ".com" to the end, of typed URLs if no host is found without them. Internet Explorer, Mozilla Firefox, Safari and Opera will also prefix "http://www." and append ".com" to the address bar contents if the Control and Enter keys are pressed simultaneously. For example, entering "example" in the address bar and then pressing either just Enter or Control+Enter will usually resolve to "http://www.example.com", depending on the exact browser version and its settings.

Pronunciation of "www"

In English, WWW (pronounced "double you double you double you") is the longest possible three-letter acronym (TLA) to pronounce, requiring nine syllables. The late Douglas Adams once quipped:

The World Wide Web is the only thing I know of whose shortened form takes three times longer to say than its long form.

– Douglas Adams, *The Independent on Sunday*, 1999

In practice it is sometimes shortened, in English usage, to "triple double-you", run together as "DubaDubaDub-u.", or even just "Dub-Dub-Dub". In other languages, "www" may be pronounced like "veh-veh-veh". The early "w³" abbreviation is now defunct.

In Chinese, the World Wide Web is commonly translated to *wàn wéi wǎng* (万维网), which satisfies "www" and literally means "ten-thousand dimensional net".

See also

- Deep web
- First image on the Web

- Search engine
- Streaming media
- Web 2.0
- Web directory
- Web operating system
- Web services
- Web science
- Webology
- Website
- Website architecture

References

- Fielding, R.; Gettys, J.; Mogul, J.; Frystyk, H.; Masinter, L.; Leach, P.; Berners-Lee, T. (June 1999). *"Hypertext Transfer Protocol — HTTP/1.1 (ftp://ftp.isi.edu/in-notes/rfc2616.txt)"*. Request For Comments 2616. Information Sciences Institute.
 - Berners-Lee, Tim; Bray, Tim; Connolly, Dan; Cotton, Paul; Fielding, Roy; Jeckle, Mario; Lilley, Chris; Mendelsohn, Noah; Orchard, David; Walsh, Norman; Williams, Stuart (December 15, 2004). *"Architecture of the World Wide Web, Volume One (http://www.w3.org/TR/webarch/)"*. Version 20041215. W3C.
 - Polo, Luciano (2003). World Wide Web Technology Architecture: A Conceptual Analysis (http://newdevices.com/publicaciones/www/). *New Devices*. Retrieved on July 31, 2005.
1. ^ "WorldWideWeb: Proposal for a HyperText Project (http://www.w3.org/Proposal.html)", Tim Berners-Lee & Robert Cailliau, 1990
 2. ^ *Information Management: A Proposal (http://www.w3.org/History/1989/proposal.html)*
 3. ^ proposal for the World Wide Web (http://www.w3.org/Proposal)
 4. ^ http://www.w3.org/People/Berners-Lee/WorldWideWeb
 5. ^ first Web pages (http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html)
 6. ^ short summary of the World Wide Web project (http://groups.google.com/groups?selm=6487%40cernvax.cern.ch)
 7. ^ http://intranet.cern.ch/Chronological/Announcements/CERNAnnouncements/2003/04-30TenYearsWWW/Welcome.html
 8. ^ http://www.livinginternet.com/w/wi_mosaic.htm
 9. ^ http://www.totic.org/nscp/demodoc/demo.html
 10. ^ http://www.cs.washington.edu/homes/lazowska/faculty.lecture/innovation/gore.html
 11. ^ http://www.brightplanet.com/technology/deepweb.asp
 12. ^ http://www.netz-tipp.de/languages.html
 13. ^ http://www.cs.uiowa.edu/~assignori/web-size/
 14. ^ http://www.w3.org/People/Berners-Lee/FAQ.html

External links

- World Wide Web Consortium (http://www.w3.org/)
- Official WWW conference, 2007 (http://www2007.org/)
 - Official WWW conference, 2006 (http://www2006.org/)
- Open Directory — Computers: Internet: Web Design and Development (http://dmoz.org/Computers/Internet/Web_Design_and_Development/)
- Early archive of the first web site (http://www.w3.org/History/19921103-hypertext/hypertext/WWW/)
- Internet Statistics: Growth and Usage of the Web and the Internet (http://www.mit.edu/people/mkgray/net/)
- A comprehensive site (http://www.watchessegment.com/)
- Living Internet (http://www.livinginternet.com/w/w.htm) A comprehensive history of the Internet, including the World Wide Web.
- World Wide Web Size (http://www.worldwidewebsite.com/) Daily estimated size of the

World Wide Web.

- Relics of the Internet (<http://relics.monkhq.com/>) Daily Internet Historical Finds. Dive into the past of the Net.

Retrieved from "http://en.wikipedia.org/wiki/World_Wide_Web"

Categories: Articles that may contain original research | Articles with unsourced statements since March 2007 | All articles with unsourced statements | World Wide Web | English inventions | Internet history | Hypertext

-
- This page was last modified 00:24, 3 July 2007.
 - All text is available under the terms of the GNU Free Documentation License. (See **Copyrights** for details.) Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a US-registered 501(c)(3) tax-deductible nonprofit charity.