Glossary

part of [Software Engineering for Internet Applications](#) by [Eve Andersson](#), [Philip Greenspun](#), and [Andrew Grumet](#)

Abstract URL

> An abstract URL is one without a file extension, e.g., `http://foobar.com/contact-info` rather than `http://foobar.com/contact-info.html` or `http://foobar.com/contact-info.aspx`. If you publish only abstract URLs, you have the freedom to change your implementation technology without breaking users' bookmarks and links from other sites.

Acceptance Test

> A test performed by an end-user or system owner to verify that the delivered software functions correctly and meets requirements.

ACID Test

> A test proposed by IBM in the 1960s for transaction database management systems: Atomicity, Consistency, Isolation, Durability. An ACID-compliant database such as Oracle or SQL Server can guarantee that two updates will be done together (atomicity), that rules for integrity can be established and enforced (consistency), that concurrent users won't see each others' half-finished work (isolation), that information won't be lost even if a hard disk dies (durability). See [the "Basics" chapter](#) and *SQL for Web Nerds* at http://philip.greenspun.com/sql/ for more.

Application Server

> see ["Middleware"](#)

AOLserver

> Released in early 1995 as "NaviServer", AOLserver remains one of the most powerful Web server programs on the market and it is free and open-source. It is a multi-threaded server that provides a lot of support for connecting to relational database management systems. AOLserver is documented at www.aolserver.com.

Apache

> Vies with Microsoft Internet Information Server for the title of "world's most popular Web server". Apache was never very technically advanced, but it was the best of the free and open-source Web servers for a time and grew to dominance. More: http://philip.greenspun.com/panda/server.

API

> Application Programming Interface. An abstraction barrier between custom/extension code and a core, usually commercial, program. The goal of an API is to let you write programs that won't break when you upgrade the underlying system. The authors of the core program are saying, "Here are a bunch of hooks into our code. We guarantee and document that they will work a certain way. We reserve the right to change the core program, but we will endeavor to preserve the behavior of the API call. If we can't, then we'll tell you in the release notes that we broke an old API call."

ASP

> Active Server Pages, introduced by Microsoft in the mid-1990s. This is the standard programming system for Internet applications hosted on Windows servers. It is bundled with Internet Information Server (IIS) when you buy Windows. The fundamental idea is

that you write HTML pages with little embedded bits of Visual Basic, C# or other languages, that are interpreted by the server.

Audit Trail

A record of past activity. For instance, a log of all past values held by columns in a database row. Or a sequence of all cash register transactions over the last three months. Or a print-out of all customer service interactions related to a given order, regardless of whether communication takes place by telephone, email, or live chat with a rep.

Blog

An online journal, published frequently (often daily). Readers can post comments on each journal entry. Some blogs gain a wide readership, such as this one: http://blogs.law.harvard.edu/philg/. The term *blog* is a shortening of *weblog*.

Bozo Filter

An individual user request that the server filter out contributions from some particular other community member.

Cable Modem

A cable modem is an Internet connection provided by a cable TV operator, typically with at least 1.5 Mbits per second of download bandwidth (50-100 times faster than modems that work over analog telephone lines).

Cache

Computer systems typically incorporate capacious storage devices that are slow (e.g., disk drives) and smaller storage devices that are fast (e.g., memory chips, which are 100,000 times faster than disk). File systems and database management systems keep recently used information from the slow devices in a *cache* in the fast device.

CGI

Common Gateway Interface. This is a standard that lets programmers write Web scripts without depending on details of the Web server program being used. Thus, for example, an Internet service implemented in CGI could be moved from a site running AOLserver to a site running Apache. CGI scripts, which run as separately launched operating system processes, are typically very slow compared to scripts than run inside a Web server program.

Client/Server

In the 1960s, computers were so expensive that each company could have only one. "The computer" ran one program at a time, typically reading instructions and data from punch cards. This was *batch processing*. In the 1970s, that computer was able to run several programs simultaneously, responding to users at interactive terminals. This was *timesharing* (it would be nice if modesty prevented one of the authors from noting that this was developed by his lab at MIT circa 1960). In the 1980s, companies could afford lots of computers. The big computers were designated *servers* and would wait for requests to come in from a network of *client* computers. The client computer might sit on a user's desktop and produce an informative graph of the information retrieved from the server. The overall architecture was referred to as *client/server*. Because of the high cost of designing, developing, and maintaining the programs that run on the client machines, Corporate America is rapidly discarding this architecture in favor of *Intranet*: Client machines run a simple Web browser and servers do more of the work required to present the information.

Code Freeze

The point at which all coding stops, usually to allow software testing without the introduction of new bugs.

Collaborative Filtering

If you can persuade a group of people to rate movies on a 1-10 scale, for example, it becomes possible to identify people whose tastes are similar. Given a new movie that only a few people have seen and rated, a *collaborative filter* can identify others in the community who might like it. Some e-commerce sites provide this service, noting for example that "customers who bought the product you're looking at right now also tended to buy these other three things". Collaborative filtering is easy to program, but ultimately is a poor substitute for human reviewers and editors.

Community Site

A community site exists to support the interaction of an online community of users. These users typically come together because of a shared interest and are most vibrant when there is an educational dimension, i.e., when the more experienced users are helping the novices improve their skills.

Compression

When storing information in digital form, it is often possible to reduce the amount of space required by exploiting regular patterns in the data. For example, documents written in English frequently contain "the". A compression system might notice this fact and represent the complete word "the" (24 bits) with a shorter code. A picture containing your friend's face plus a lot of blue sky could be compressed if the upper region were described as "a lot of blue sky". All popular Web image, video, and sound formats incorporate compression.

Content Repository

Instead of having one SQL table for every different kind of content on a site, e.g., `articles`, `comments`, `news`, `questions`, `answers`, it is possible to define a single *content repository* table that is flexible enough to store all of these in one place. This approach to data modeling makes it simpler to perform queries such as "show me all the new stuff since yesterday" or "show me all the content contributed by User #37". With a content repository, it is also easier to program and enforce consistent site-wide policies regarding approval, editing, and administration of content.

Cookie

The Cookie protocol allows a Web application to conveniently maintain a "session" with a particular user. The Web server sends the client a "magic cookie" (piece of information) that the client is required to return on subsequent requests. The original specification is at http://home.netscape.com/newsref/std/cookie_spec.html.

Data Model

A data model is the structure in which a computer program stores persistent information. In a relational database, data models are built from tables. Within a table, information is stored in homogeneous columns, e.g., a column named `registration_date` would contain information only of type `date`. A data model is interesting because it shows what kinds of information a computer application can process. For example, if there is no place in the data model for the program to store the IP address from which content was posted, the publisher will never be able to automatically delete all content that came from the IP address of a spammer.

DNS

The Domain Name System translates human-readable hostnames, e.g., `www.google.com`, into machine-readable and network-routable IP addresses, e.g., `216.239.57.100`. DNS is a distributed application in that there is no single computer that holds translations for all possible hostnames. A domain registrar, e.g., www.register.com, records that the domain servers for the google.com domain are at particular IP addresses. A user's local *name server* will query the name servers for google.com to find the translation for the hostname `www.google.com`. Note that there is nothing magic about "www"; it is merely a conventional name for a computer that runs a Web server. The procedure for translating a hostname such as `froogle.google.com` is the same as that applied for `www`. *Round robin DNS* was an early load-balancing technique in which multiple computers at different IP addresses were configured to serve an application; browsers asking the DNS servers to translate the site's hostname would get different answers depending on when they asked, thus spreading out the users among the multiple computers hosting the application.

DTD

Document Type Definition. The specification of an XML document's schema, including its elements, attributes, and data structure. DTDs are used for validating that an XML document is well-formed. You can also share a DTD with your collaborators in order to agree upon the structure of XML documents that will be exchanged.

Dynamic Site

A dynamic site is one that is able to collect information from User A, serve it back to Users B and C immediately, and hide it from User D because the server knows that User D isn't interested in this kind of content. Dynamic sites are typically built on top of relational database management systems because these programs make it easy to organize content submitted by hundreds of concurrent users. An example of a simple dynamic site would be a classified ad system.

Emacs

World's most powerful text editor, written by Richard Stallman (RMS) in 1976 for the Incompatible Timesharing System (ITS) on the PDP-10s at MIT. Emacs has been subsequently ported to virtually every kind of computer hardware and operating system between 1976 and the present (including the Macintosh, Windows 95/NT, and every flavor of Unix). Good programmers tend to spend their entire working lives in Emacs, which is capable of functioning as a mail reader, USENET news reader, Web browser, shell, calendar, calculator, and Lisp evaluator. Emacs is infinitely customizable because users can write their own commands in Lisp. You can find out more about Emacs at ftp://publications.ai.mit.edu/ai-publications/pdf/AIM-519A.pdf (Stallman's 1979 MIT AI Lab report), at www.gnu.org (where you can download the source code for free), or by reading *Learning Emacs* (Cameron et al 1996; O'Reilly). If you want to program Emacs and then you'll want *Writing Gnu Emacs Extensions* (Bob Glickstein 1997; O'Reilly).

Filter

The best Web server APIs allow the programmer to say "run this little piece of code before [or after] serving files that match a particular URL pattern." Filters that run after a file is served are useful if you want to add extra logging to an application. Filters that run before a file is served or a script is run are useful for implementing a security policy in a consistent fashion, rather than relying on the authors of individual scripts to insert an authentication check.

Firewall

A computer that sits between a company's internal network of computers and the public Internet. The firewall's job is to make sure that internal users can get out to enjoy the benefits of the Internet while external crackers are unable to make connections to machines behind the firewall.

Flat-file

A flat-file database keeps information organized in a structured manner, typically in one big file. A desktop spreadsheet application is an example of a flat-file database management system. These are useful for Web publishers preparing content because a large body of information can be assembled and then distributed in a consistent format. Flat-file databases typically lack support for processing transactions (inserts and updates) from concurrent users. Thus, collaboration or e-commerce Web sites generally rely on a relational database management system as a back-end.

GIF

Graphical Interchange Format. Developed in 1987 by CompuServe, this is a way of storing compressed images with up to 256 colors. It became popular on the Web because it was the only format that could be displayed in-line by the first multi-platform Web browser (NCSA Mosaic). The JPEG image file format results in much better looking images with much smaller sized files.

HTML

Hyper Text Markup Language. Developed by Tim Berners-Lee, this specifies a format for the most popular kind of document distributed over the Web (via HTTP). Documented sketchily in this book, documented badly at http://www.w3.org, and documented well in *HTML: The Definitive Guide* (Musciano and Kennedy 2002; O'Reilly).

HTTP

Hyper Text Transfer Protocol. Developed by Tim Berners-Lee, this specifies how a Web browser asks for a document from a Web server. Question such as "how does a server tell the browser that a document has moved?" or "how does a browser ask the time that a document was last modified?" may be answered by reference to this protocol, which is documented badly at http://www.w3.org and documented well in various books such as *HTTP: The Definitive Guide* (David Gourley, Brian Totty; O'Reilly 2002).

IIS

Internet Information Server. A threaded Web server program that is included by Microsoft when you purchase the Windows operating system.

Java

Java is first a programming language, developed by Sun Microsystems around 1992, intended for use on the tiny computers inside cell phones and similar devices. Java is second an interpreter, the Java virtual machine, formerly compiled into popular Web browsers (back when Netscape Navigator was popular and before Sun sued Microsoft). Java is third a security system that purports to guarantee that a program downloaded from an untrusted source on the Internet can run safely inside the interpreter. Java is the only realistic way for a Web publisher to take advantage of the computing power available on a user's desktop. Java is generally a cumbersome language for server-side software development. For more background on the language, see the "Java" chapter from *Database Backed Web Sites* at http://philip.greenspun.com/wtr/dead-trees/53008.htm.

JPEG

Joint Photographic Experts Group. A bunch of people who sat down and designed a standard for image compression, conveniently titled "IS 10918-1 (ITU-T T.81)". This standard works particularly well for 24-bit color photographs. C-Cube Microsystems came up with the JFIF standard for encoding color images in a file. Such a file is what people commonly refer to as "a JPEG" and typically ends in ".jpg" or ".jpeg". The main problem with JFIF files is that they record only 8 bits per color, a vastly smaller range of intensities than is present in the natural world and significantly smaller than the 12- and 14-bits-per-color signals that come out of the best digital scanners and cameras. This defect and more are remedied in the JPEG 2000 standard. See www.jpeg.org for more about the standard.

LDAP

Lightweight Directory Access Protocol. A typical LDAP server is a simple network-accessible database where an organization stores information about its authorized users and what privileges each user has. Thus, rather than create a new employee account on 50 different computers, the new employee is entered into LDAP and granted rights to those 50 systems. If the employee leaves, revoking all privileges is as simple as removing one entry in the LDAP directory. LDAP is a bit confusing because original implementations were presented as alternatives to the Web and the relational database management system. Nowadays many LDAP servers are implemented using standard RDBMSes underneath, and they talk to the rest of the world via XML documents served over HTTP.

Linux

A free version of the Unix operating system, primarily composed of tools developed over a 15-year period by Richard Stallman and Project GNU. However, the final spectacular push was provided by Linus Torvalds who wrote a kernel (completed in 1994), organized a bunch of programmers Internet-wide, and managed releases.

Lisp

Lisp is the most powerful and also easiest-to-use programming language ever developed. Invented by John McCarthy at MIT in the late 1950s, Lisp is today used by the most sophisticated programmers pushing the limits of computers in mathematical physics, computer-aided engineering, and computer-aided genetics. Lisp is also used by thousands of people who don't think of themselves as programmers at all, only people who want to define shortcuts in AutoCAD or the Emacs text editor. The best introduction to Lisp is also the best introduction to computer science: *Structure and Interpretation of Computer Programs* (Abelson and Sussman 1996; MIT Press).

Log Analyzer

A program that reads a Web server's access log file (one line per request served) and produces a comprehensible report with summary statistics, e.g., "You served 234,812 requests yesterday to 2,039 different computers; the most popular file was /samoyed-faces.html".

Magnet Content

Material authored by a publisher in hopes of establishing an online community. In the long-run, a majority of the content in a successful community site will be user-authored.

Middleware

A vague term that, when used in the context of Internet applications, means "software sold to people who don't know how to program by people who don't know how to program." In theory, middleware sits between your relational database management

system and your application program and makes the whole system run more reliably, just like adding a bunch of extra moving parts to your car would make it more reliable.

MIME

Multi-Purpose Internet Mail Extensions. Developed in 1991 by Nathan Borenstein of Bellcore so that people could include images and other non-plain-text documents in e-mail messages. MIME is a critical standard for the World Wide Web because an HTTP server answering a request always includes the MIME type of the document served. For example, if a browser requests "foobar.jpg", the server will return a MIME type of "image/jpeg". The Web browser will decide, based on this type, whether or not to attempt to render the document. A JPEG image can be rendered by all modern Web browsers. If, for example, a Web browser sees a MIME type of "application/x-pilot" (for the .prc files that PalmPilots employ), the browser will invite the user to save the document to disk or select an appropriate application to launch for this kind of document.

Multi-modal

A multi-modal user interface allows you to interact with a piece of software in a variety of means simultaneously. For example, you may be able to communicate using a keyboard or stylus, or with your voice, or even with hand or face gestures. These are all "modes" of communication. The advent of GPRS makes simultaneous voice/keypad interaction possible on cellular telephones.

Operating System (OS)

A big complicated computer program that lets multiple simultaneously executing big complicated computer programs coexist peacefully on one physical computer. The operating system is also responsible for hiding the details of the computer hardware from the application programmers, e.g., letting a programmer say "I want to write ABC into a file named XYZ" without the programmer having to know how many disk drives the computer has or what company manufactured those drives. Examples of operating systems are Unix and Windows XP. Examples of things that try to be operating systems, but mostly fail to fulfill the "coexist peacefully" condition, are Windows 98 and the Macintosh OS.

Oracle

Oracle is the most popular relational database management system (RDBMS). It was developed by Larry Ellison's Oracle Corporation in the late 1970s.

Perl

Perl is a scripting language developed by Larry Wall in 1986 to make his Unix sysadmin job a little easier. It unifies a bunch of capabilities from disparate older Unix tools. Like Unix, Perl is perhaps best described as "ugly but fast and useful". Perl is free, has particularly powerful string processing operators, and quickly developed a large following and, therefore, a large library for CGI scripting. For more info, see www.perl.com or www.perl.org.

Historical Note: Lisp programmers forced to look at Perl code would usually say "if there were any justice in this world, the guys who wrote this would go to jail." In a rare case of Lisp programmers getting their wish, in 1995 Intel Corporation persuaded local authorities to send Randal Schwartz, author of *Learning Perl* (O'Reilly 2001), to the Big House for 90 days (plus 5 years of probation, 480 hours of community service, and $68,000 of "restitution" to Intel). Sadly, however, it seems that Schwartz's official crime was not corrupting young minds with Perl syntax and semantics. Most Unix sysadmins

periodically run a program called "crack" that tries to guess user passwords. When crack is successful, the sysadmins send out email saying "your password has been cracked; please change it to something harder to guess." Obviously they do not need the passwords since they have root access to all the boxes and can read any of the data contained on them. At a university, you get paid about $50,000/year for doing this. In Oregon if you do this for a multi-billion-dollar company that has recently donated $100,000 to the local law enforcement authorities, you've committed a crime. See http://www.lightlink.com/spacenka/fors/ for more on *State of Oregon v. Randal Schwartz*.

Persistence
> The continued existence of data. A *persistence mechanism* is something that provides long-term data storage, even when the application that created the data is no longer running. Examples include RDBMSes, XML documents, and flat-file databases.

RDBMS
> Relational Database Management System. A computer program that lets you store, index, and retrieve tables of data. The simplest way to look at an RDBMS is as a spreadsheet that multiple users can update. The most important thing that an RDBMS does is provide transactions.

Request Processor
> Portion of a Web server program that decides how to handle an incoming request. A well-designed request processor enables a publisher to expose only abstract URLs, e.g., "glossary" rather than "glossary.html". The job of the request processor is to dig around in the file system and find a document to deliver or a script to execute.

RMS
> Richard M. Stallman. In 1976, he developed Emacs, the world's best and most widely used text editor. He went on to develop gcc, the most widely used compiler for the C programming language, and won a $240,000 MacArthur fellowship in 1990. Stallman is the founder of the free software movement (see www.fsf.org), and Project GNU, which gave rise to Linux.

Robot
> In the technologically optimistic portion of the 20th century, robots were intelligent anthropomorphic machines that understood human speech, interpreted visual scenes, and manipulated objects in the real world. In the technologically realistic 21st century, robots are absurdly primitive programs that do things like "Go look up this book title at three different online bookstores and see who has the lowest price; fail completely if any one of the online bookstores has added a comma to their HTML page." Also known as *intelligent agents* (an intellectually vacuous term but useful for getting tenure if you're a university professor). Some simple but very useful examples of robots are the spiders or Web crawlers that fill the content database at public search engine sites such as AltaVista.

Scalable
> A marketing term used to sell defective software to executives at big companies. Internet applications are fundamentally concerned with processing updates from thousands of concurrent users. This is what database management systems were built for. Smart engineers build Web applications so that if the database is up and running, the Web site

will be up and running. Period. Adding more users to the site will inevitably require adding capacity to the database management system, no matter what other software is employed. The thoughtful engineer will realize that a provably *scalable* site is one that relies on no other software besides the database management system and the thinnest of software layers on top, such as Apache, AOLserver, or Microsoft IIS.

Semantic Tag

The most popular Web markup language is HTML, which provides for formatting tags, e.g., "this is a headline" or "this should be rendered in italics." This is useful for humans reading Web pages. What would be more useful for computer programs trying to read Web pages is a *semantic tag*, e.g., "the following numbers represent the price of the product in dollars", or "the following characters represent the date this document was initially authored". More: http://www.w3.org/RDF/.

SOAP

Simple Object Access Protocol. A way for a Web server to call a procedure on another, physically separate Web server, and get back a machine-readable result in a standardized XML format. Useful for building a Web page that combines dynamic information pulled from multiple foreign sites. Also useful for building a single Web form that can perform multiple actions at foreign sites on behalf of a user. See http://msdn.microsoft.com/soap/ and http://www.w3.org/TR/SOAP/.

SGML

Standard Generalized Markup Language, standardized in 1980. A language for marking up documents so that they could be parsed by computer programs. Each community of people that wishes to author and parse documents must agree on a Document Type Definition (DTD), which is itself a machine-parsable description of what tags a marked-up document must or may have. HTML is an example of an SGML DTD. XML is a simplified descendant of SGML.

Soft Launch

Placing a server on the public Internet but only telling a handful of people about it gives the developers a chance to see how real users interact with the system, fix bugs, and see how the servers handle a gradually increasing load. A *soft launch* like this is much safer than a Big Bang-style launch in which the server is made public just as a massive advertising campaign airs.

Spider

A spider or *Web crawler* is a program that exhaustively surfs all the links from a page and returns them to another program for processing. For example, all of the Internet search engine sites rely on spider robots to discover new Web sites and add them to their index. Another typical use of a spider is by a publisher against his or her own site. The spider program makes sure that all of the links function correctly and reports dead links.

SQL

Structured Query Language. Developed by IBM in the mid-1970s as a way to get information into and out of relational database management systems. A fundamental difference between SQL and standard programming languages is that SQL is declarative. You specify what kind of data you want from the database; the RDBMS is responsible for figuring out how to retrieve it. A full tutorial on SQL is available at http://philip.greenspun.com/sql/.

Static Site

A static Web site comprises content that does not change depending on the identity of the user, the time of day, or what other users might have contributed recently. A static Web site is typically built using static documents in HTML format with graphics in GIF format and images in JPEG format. Collectively, these are referred to as *static files*. Contrast with a dynamic site, in which content can be automatically collected from users, personalized for the viewer, or changed as a function of the time of day.

TCP/IP

Transmission Control Protocol and Internet Protocol. These are the standards that govern transmission of data among computer systems. They are the foundation of the Internet. IP is a way of saying "send these next 1000 bits from Computer A to Computer B". TCP is a way of saying "send this stream of data reliably between Computer A and Computer B" (it is built on top of IP). TCP/IP is a beautiful engineering achievement, documented beautifully in *TCP/IP Illustrated, Volume 1* (W. Richard Stevens 1994; Addison-Wesley).

Transaction

A *transaction* is a set of operations for which it is important that all succeed or all fail. On an e-commerce site, when a customer confirms a purchase, you'd like to send an order to the shipping department and simultaneously bill the customer's credit card. If the credit card can't be billed, you want to make sure that the order doesn't get shipped. If the shipping database can't accept the order, you want to make sure that the credit card doesn't get billed. RDBMSes such as Oracle provide significant support for implementing transactions.

UDDI

Universal Description, Discovery, and Integration. Like a worldwide Yellow Pages, this is an XML-based registry where companies can list the Web services they provide. More: uddi.org.

Unix

An operating system developed by Ken Thompson and Dennis Ritchie at Bell Laboratories in 1969, vaguely inspired by the advanced MULTICS system built by MIT. Unix really took off after 1979, when Bill Joy at UC Berkeley released a version for Digital's VAX minicomputer. Unix fragmented into a bewildering variety of mutually incompatible versions, thus enabling Microsoft Windows to take over most of the server market. The only surviving variants of Unix are Sun's Solaris and Linux.

URL

Uniform Resource Locator, also Uniform Resource Identifier (URI). A way of specifying the location of something on the Internet, e.g., "http://philip.greenspun.com/seia/glossary" is the URL for this glossary. The part before the colon specifies the protocol (HTTP). Legal alternatives include encrypted protocols such as HTTPS and legacy protocols such as FTP, news, gopher, etc. The part after the "//" is the server hostname ("philip.greenspun.com"). The part after the next "/" is the name of the file on the remote server. Also see "Abstract URL". More: http://www.w3.org/Addressing/.

USENET

A threaded discussion system that today connects millions of users from around the Internet into newsgroups such as rec.photo.equipment.35mm. The original system was

built in the late 1970s and ran on one of the wide-area computer networks later subsumed into the Internet.

Version Control System

A system for keeping track of multiple versions of a file, usually source code. Version control systems are most useful when many developers are working together on a project, to help prevent one developer from overwriting another developer's changes, and to make it easy to revert to a previous version of a file. An excellent open-source version control system is CVS, Concurrent Versions System: www.cvshome.org.

VoiceXML

A markup language used for the development of voice applications. Using only a traditional Web infrastructure, you can create applications that are accessible over the telephone. With VoiceXML, you can specify call flow, speech recognition, and text-to-speech. See the "Voice" chapter for more.

W3C

The World Wide Web Consortium. The W3C is a vendor-neutral industry consortium that promotes standards for the World Wide Web. Popular W3C standards include HTML, HTTP, URL, XML, SOAP, VoiceXML, and many more: www.w3.org.

WAP

Wireless Application Protocol. A set of standard communication protocols for wireless devices. See the "Mobile" chapter for more.

Web Service

These days, the term *Web service* typically refers to a modular application that can be invoked through the Internet. The consumers of Web services are other computer applications that communicate, usually over HTTP, using XML standards including SOAP, WSDL, and UDDI. Sometimes *Web service* will still be used in the older sense of the word, as a user-facing application like amazon.com or photo.net.

Weblog

See Blog.

Windows NT/2000/XP

A real operating system that can run the same programs with more or less the same user interface as the popular Windows 95/98 system. Windows NT was developed from scratch by a programming team at Microsoft that was mostly untainted by the people who brought misery to the world in the form of Windows 3.1/95. The latest versions of Windows work surprisingly well.

WML

Wireless Markup Language. An out-of-date markup language for the development of mobile browser applications. Replaced by XHTML-MP.

Workflow

The management of steps in a business processes. A workflow specifies what tasks need to be done, in what order (sometimes linearly, sometimes in parallel), and who has permission to perform each task. Most tasks are performed by humans but they can also be automated processes.

WSDL

Web Services Description Language. A way for a Web server to answer, in a machine-readable form, the question "what services do you provide?" with said services ultimately to be provided by SOAP. See http://www.w3.org/TR/wsdl.

WYSIWYG

What You See Is What You Get. A WYSIWYG word processor, for example, lets a user view an on-screen document as it will appear on the printed page, e.g., with text in italics appearing on-screen in italics. This approach to software was pioneered by Xerox Palo Alto Research Center in the 1970s and widely copied since then, notably by the Apple Macintosh. WYSIWYG is extremely effective for structurally simple documents that are printed once and never worked on again. WYSIWYG is extremely ineffective for the production of complex documents and documents that must be maintained and kept up-to-date over many years. Thus Quark Xpress and Adobe FrameMaker facilitated a tremendous boom in desktop publishing, while Microsoft FrontPage and similar WYSIWYG tools for Web page construction have probably hindered development of interesting Web applications.

XHTML

The next generation of HTML, compliant with XML standards. Although it is very similar to the current HTML, it follows a stricter set of rules, thus allowing for better automatic code validation. This structure also makes it possible to embed other XML-based languages such as MathML (for equations) and SMIL (for multimedia) inside of XHTML pages. More: www.wdvl.com/Authoring/Languages/XML/XHTML/

XHTML-MP

XHTML Mobile Profile. A strict subset of XHTML, used as a markup language for wireless application development. See the "Mobile" chapter for more.

XML

Extensible Markup Language, a simplified version of SGML with enhanced features for defining hyperlinks. As with SGML, it solves the trivial problem of defining a syntax for exchanging structured information but doesn't do any of the hard work of getting users to agree on semantic structure.

Return to Table of Contents

eve@eveandersson.com, philg@mit.edu, aegrumet@mit.edu