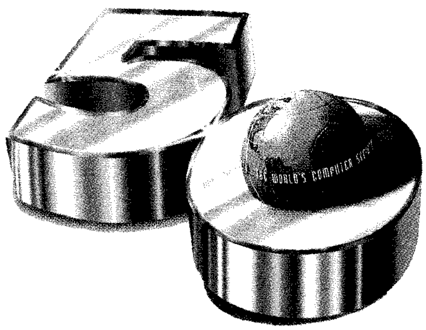


WWW: Past, Present, and Future

Tim Berners-Lee
World Wide Web Consortium, Massachusetts Institute of Technology

The World Wide Web is simply defined as the universe of global network-accessible information. It is an abstract space within which people can interact, and it is chiefly populated by interlinked pages of text, images, and animations, with occasional sounds, videos, and three-dimensional worlds.

The Web marks the end of an era of frustrating and debilitating incompatibility between computer systems. It has created an explosion of accessibility, with many potential social and economical impacts.

The Web was designed to be a space within which people could work on a project. This was a powerful concept, in that

- people who build a hypertext document of their shared understanding can refer to it at all times;
- people who join a project team can have access to a history of the team's activities, decisions, and so on;
- the work of people who leave a team can be captured for future reference; and
- a team's operations, if placed on the Web, can be machine-analyzed in a way that could not be done otherwise.

The Web was originally supposed to be a personal information system and a tool for groups of all sizes, from a team of two to the entire world.

People have rapidly developed new features for the Web, because of its tremendous commercial potential. This has made the maintenance of global Web interoperability a continuous task. This has also created a number of areas into which research must continue.

BEFORE THE WEB

The hypertext concept can be traced back to such work as Vannevar Bush's famous 1945 article "As We May Think."[1] In the article, Bush proposed the Memex machine, which would use binary coding, photocells, and instant photography to let people make and follow microfilm cross-references.

The concept continued in the mid-1960s as Douglas Englebart developed the NLS,[2] which used digital computers and provided hypertext e-mail and documentation sharing. Meanwhile, in 1965, Ted Nelson coined the word hypertext.[3]

Nonetheless, in 1980, the world still suffered from incompatible networks, incompatible disk formats, incompatible data formats, and incompatible character-encoding schemes. This made any attempt to transfer information between systems daunting and impractical. This was frustrating because people were increasingly using computers to handle information, a large amount of important information was already stored in computers, and many of the computers were networked. However, the systems in use, including those that were proprietary and those that physicists wrote for their own use, were incompatible.

The Web marks the end of frustrating incompatibility between computers. Maintaining the interoperability is a continuous, global task.

When I worked at CERN, the European particle physics laboratory, I designed the Enquire program for my own use. The system, which formed

0018-9162/96/$5.00 © 1996 IEEE

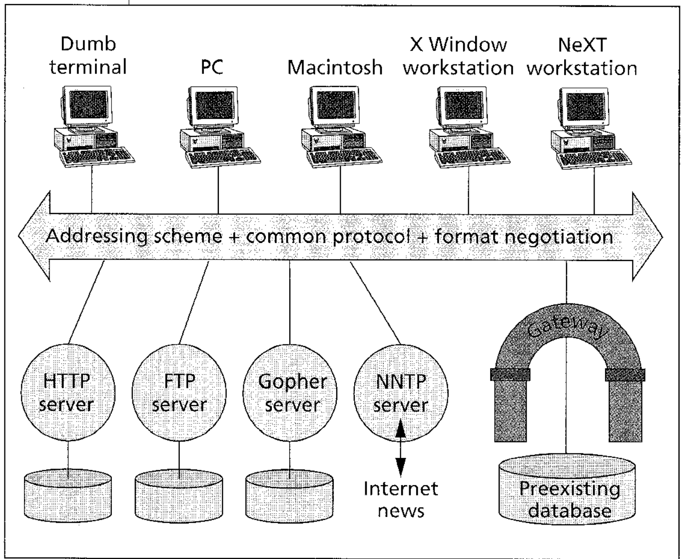


Figure 1. Diagram of Web architecture from early discussions on the Web that took place in January 1990. The Web's architecture lets users on any of a variety of clients read and, where appropriate, manipulate information on any provider's server. (The fourth server from the left is the Network News Transfer Protocol server). This interoperability requires common standards represented in the figure by the arrow: the URI addressing scheme, such protocols as HTTP, and HTTP's ability to negotiate the transfer of a mutually understood format, such as HTML.

the conceptual basis for Web development, allowed random links but was not usable across a wide-area network.

DESIGN GOALS

The Web's major goal was to be a shared information space through which people and machines could communicate. This space was to be inclusive, rather than exclusive. It would serve as a public and private information system, and would include high-value, carefully checked and designed material, as well as obscure ideas that make sense to only a few people.

I wanted the interaction between person and hypertext to be so intuitive that the machine-readable information space would give an accurate representation of the state of people's interactions and work patterns. Machine analysis would then become a powerful management tool, helping us solve the type of problems that beset the management of large organizations.

The Web's design was based on a few criteria:

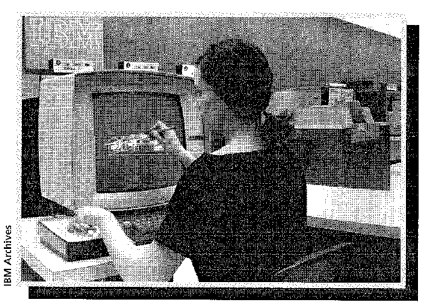
- An information system must be able to record random associations between arbitrary objects, unlike most database systems.
- If two sets of users start to use the system independently, making a link from one system to another should be an incremental effort, not requiring unscalable operations such as the merging of link databases.
- Attempts to limit users to particular languages or operating systems are doomed to failure.
- Information must be available on all platforms.
- Attempts to force users to deal with information in the same way computers deal with information are doomed to failure.
- To insure accuracy, information must be easy to enter and correct.

BASIC ARCHITECTURAL PRINCIPLES

The Web's architecture, shown in Figure 1, was proposed in 1989. It was designed to meet the above criteria and follow established principles of software design adapted to the network environment.

Independence of specifications

Flexibility was clearly a key goal. Every specification that was needed to ensure interoperability constrain the Web's implementation. Therefore, there should be as few specifications as possible (the principle of minimal constraint), and the necessary specifications should be made independently (the principles of modularity and information hiding). This would let you replace parts of the design while preserving the basic architecture.



IBM Archives

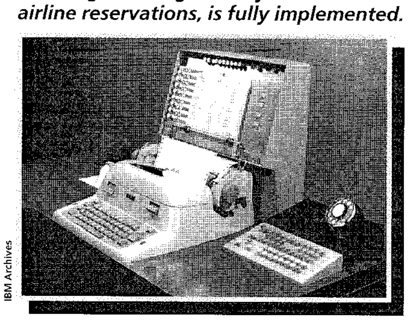
1964 IBM announces the System/360 "third-generation" line of computers.

1964 Basic (Beginner's All-Purpose Symbolic Instruction Code) is developed at Dartmouth by John Kemeny and Thomas Kurtz. It spawns many variations.

```
10 print "Hello World!"
20 goto 10
```

http://www.latec.edu/~acm/HelloWorld.shtml

1964 IBM's seven-year-long Sabre project, allowing travel agents anywhere to make airline reservations, is fully implemented.



IBM Archives

1964

If done properly, you could mix older and newer specifications. Thus, the old FTP protocol could be mixed with the new HTTP protocol in the address space, and conventional text documents could be mixed with new hypertext documents.

The principle of minimal constraint was a major factor in the Web's adoption. People frequently had to make minor and incremental changes to adopt the Web, first as a parallel technology to existing systems and then as the principal technology. The ability of many people to evolve from the past to the present within general architectural principles inspires hope that future evolution will be equally smooth and incremental.

Universal resource identifiers

Typically, hypertext systems were built around a database of links. Because they were centralized, they became unmanageable when scaled up. However, they guaranteed that links would be consistent and they removed links to documents that had been deleted. The removal of this guarantee was the principal compromise made in the Web architecture. Letting people make a reference without consulting the reference's destination permitted the scalability that the Web subsequently exploited.

The power of a link in the Web is that it can point to any document (or any other type of resource) in the universe of information. This requires a global space of identifiers. These universal resource identifiers (URIs) are the primary elements of Web architecture. The well-known structure starts with a prefix such as http:, which indicates the space into which the rest of the string points.

The URI space is universal in that any new space that has an identifying, naming, or addressing syntax can be mapped into a printable syntax, can be given a prefix, and can then become part of URI space. The properties of any URI depend on the properties of the space into which it points. Depending on these properties, some spaces are known as name spaces, and some are known as address spaces. However, a space's properties do not depend only on technical factors, such as its definition and syntax, and the protocols it uses to dereference names. The properties also depend on social factors, such as agreements about how identifiers can be allocated and reallocated.

Fortunately, the Web architecture does not depend on whether a URI is a name or an address, although the phrase URL (uniform resource locator) was coined by the Internet Engineering Task Force to indicate that most URIs are more like addresses than names. We await the definition of more powerful name spaces. This is not a trivial matter, and it is a problem that is as much social and administrative as it is technical.

HTTP

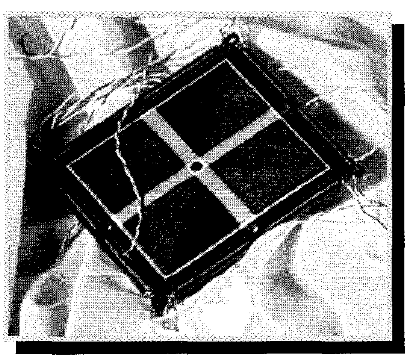
In 1989, FTP existed as a standard for accessing remote data. However, FTP was insufficiently rich in features and too slow to be optimal for the Web. So, I designed the hypertext transfer protocol, a new protocol designed to operate with the speed necessary for traversing hypertext links. HTTP URIs are resolved into the addressed document by splitting them into two parts. The first part is applied to the domain name service to discover a suitable server, and the second half is an opaque string that is handed to that server.

HTTP lets a client specify language and data format preferences. This lets a server select a suitable specific object when you request a generic URI. This feature, known as format negotiation, is a key element of independence between HTTP and HTML specifications. Format negotiation is implemented in various HTTP servers, but clients tend to use it infrequently, partly because of the time needed to transmit the preferences and partly because generic URIs have historically been the exception.

We await the definition of more powerful name spaces. This is not a trivial matter, and it is a problem that is as much social and administrative as it is technical.

HTML

For the interchange of hypertext, I defined the Hypertext Markup Language as a data format to be transmitted on the network. HTML was chosen because it resembled some SGML-based systems and thus would be adopted more quickly by both the documentation community, which preferred the SGML syntax, and the hyper-

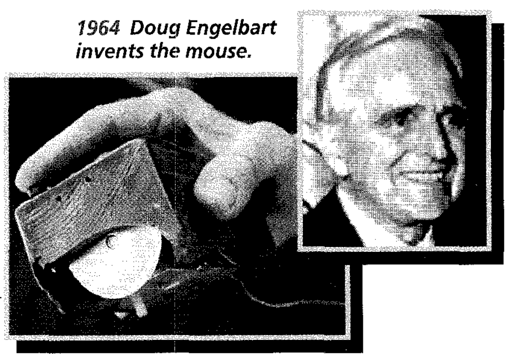


University of Virginia

***1964** With a speed of 9 megaflops, Control Data Corp.'s CDC 6600, designed by Seymour Cray, claims the title of first commercially successful supercomputer.*

***1964** IBM develops a computer-aided design system.*

***1964** Doug Engelbart invents the mouse.*



Bootstrap Institute

1964

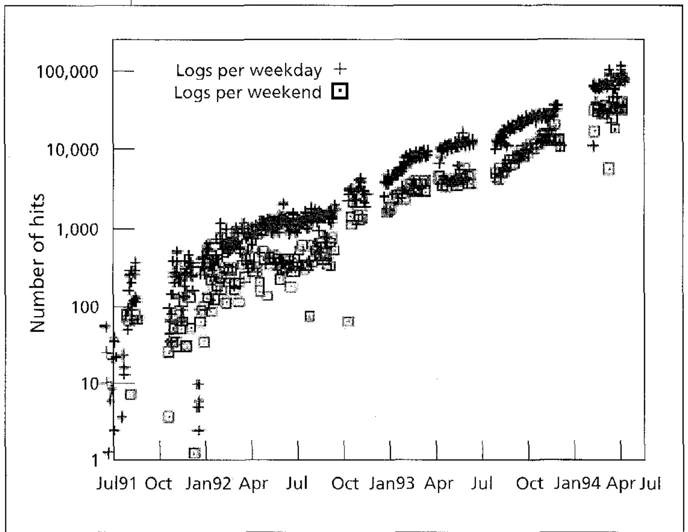


Figure 2. Web client growth from July 1991 to July 1994. A sign of the growth in early Web usage was the increase in the number of hits on the first Web server, info.cern.ch. The number of Web clients between July 1991 and July 1994 grew by a factor of 10 each year, as shown in the graph. Missing points represent lost data.

text community, in which only SGML-based standards had been proposed.

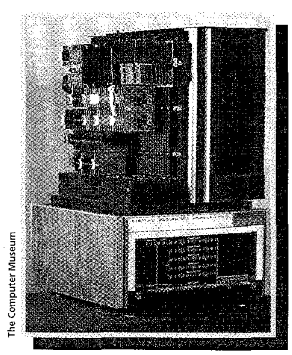
The adoption of SGML encouraged these communities to accept the Web more easily, but SGML turned out to have a very complex and poorly defined syntax. The attempt to achieve full SGML compatibility and HTML's ease of use bedeviled experts for a long time.

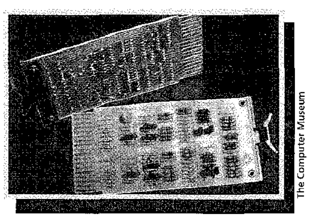
EARLY HISTORY

The Web's road from conception to adoption certainly had its curves. It was impossible to convince many people to use the original system, which had a small audience and content only about itself.

The key steps in the Web's early history were as follows:

- I wrote the initial World Wide Web prototype from October to December 1990 using the NeXTStep development environment. This program allowed the simple addition of new links and new documents, because it was a WYSIWYG editor that browsed at the same time. However, the limited deployment of NeXTStep limited the number of people who could try the prototype and experience its power. The initial Web of hypertext was written with links to sound and graphic files and was published by a simple HTTP server.
- To encourage global acceptance, Nicola Pellow in 1991 wrote a line-mode browser, a portable hypertext browser that lets you retrieve Web information on any platform. At the time, this was all that many people saw of the Web.
- To seed the Web with data, Bernd Pollermann and I wrote a second server, which provided a gateway into a legacy phone book database on a CERN mainframe. This was the first useful Web application, so many people saw the Web as a phone book program with a strange user interface. However, this got the line-mode browser onto a few desks. This gateway server was followed by a number of others, providing enough servers to make a Web client a useful tool for the physicists who drove the Web's development.
- The Internet community at large was encouraged to port the Web program to other platforms. In 1992, Erwise, Midas, and ViolaWWW for X Windows and Cello for Windows were among the resulting clients. Unfortunately, they were only browsers, although ViolaWWW was based on an interpreted mobile code language (Viola) and was comparable in some respects to HotJava.
- The University of Minnesota's Internet Gopher[4] was seen for a long time as the preferred information system, because it avoided HTML's complexities. However, that changed because potential users were concerned about the university's licensing fee.
- In 1993, Marc Andreessen of the National Center for Supercomputer Applications, having seen ViolaWWW, wrote Mosaic, a Web client for X Windows. Mosaic was easy to install, was later allowed in-line images, and



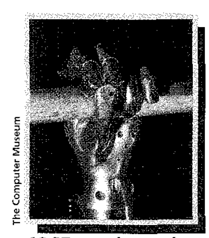


1965 J.A. Robinson develops unification, the underpinning of logic programming and important to many of today's programming technologies.

1965 DEC debuts the first minicomputer, the PDP-8, which used transistor circuitry modules.

1965 Project MAC, a large collaborative time-sharing project, leads to the Multics operating system.

1965 Maurice Wilkes proposes the use of a cache memory on the basis of an idea by Gordon Scarott.



1965 At the University of Belgrade, Rajko Tomovic makes one of the earliest attempts to develop an artificial limb with a sense of touch.

1965

eventually became very popular.

- In 1994, Navisoft created a browser/editor that was reminiscent of the original Web program, which could also browse and edit in the same mode.

The first Web server offered suggestions on ways to find and run both clients and servers. The server had a page on etiquette, which included such conventions as using the e-mail address "Webmaster" (taken from the term postmaster) for a server administrator, and using "www" as the host name of an HTTP server.

An early metric of growing Web usage was the increase in the number of hits on the first Web server, info.cern.ch. (That server, which originally ran on the same machine as the first client, has been replaced by www.w3.org.) From July 1991 to July 1994, the number of Web clients grew exponentially, by a factor of 10 per year, as Figure 2 shows

CURRENT SITUATION

At this point, the general public became aware of the Web. HTML, which was intended to be the warp and woof of a hypertext tapestry crammed with rich and varied data types, became surprisingly widespread. The Web started to drive, rather than rely on, the extent of computer availability and Internet connectivity. URL syntaxes such as http: became as well-known to many people as the 800 area code for toll-free telephone numbers.

Incompatibilities and tensions

The common standards of URIs, HTTP, and HTML have permitted the Web's growth and have allowed companies and universities to devote resources to the Web's exploitation and extension. This has resulted in many new data types and protocols.

For example, the ability of HTTP to handle arbitrary data formats has allowed easy expansion of the Web. So, for example, it has been easy to introduce the Virtual Reality Markup Language, the 3D scene description language, and the Java byte-code format for the transfer of mobile program code. It has been more difficult for servers to know what has been supported by clients, because the format negotiation system has not been widely deployed in clients.

This has led to such problems as the deplorable engineering practice of checking, in the server, the browser's manufacturer and version against a table kept by the server. This practice makes it difficult to introduce new clients, and, of course, it is difficult to maintain. As a result, new clients sometimes pretend to be well-known clients in order to extract sufficiently rich data from servers. This has been accompanied by an insufficiency in the MIME types used to describe data—text/html refers to many HTML levels, image/png refers to any Portable Network Graphics format graphic, and Java files are shipped without any visible indication of the runtime support they will require.

WWW Consortium: Toward compatibility

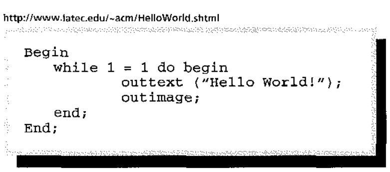
There was concern that fragmentation of Web standards would destroy the universe of information upon which so many technical and commercial developments were being built. This led to the formation of the World Wide Web Consortium in 1994. The consortium has about 150 members, including all major Web technology developers and many other organizations that increasingly use or depend on the Web.

The consortium is based at the Massachusetts Institute of Technology in Cambridge, Massachusetts, and at the Institut Nationale pour la Récherche en Informatique et Automatique (National Institute for Computer and Automation Research) in France. It provides a vendor-neutral forum where competing companies can work together to develop common specifications for the common good.

Web protocol developments are sometimes driven by the infrastructure's technical needs, by particular applications, by the connection between the Web and society, or by a combination of these factors.

Web protocol developments are sometimes driven by the infrastructure's technical needs, by particular applications, by the connection between the Web and society, or by a combination of these factors.

An example occurred in response to concerns by parents, schools, and government agencies that young children would gain access to indecent, violent, or otherwise inappropriate material on the Internet. To avoid govern-



1967 Ole-Johan Dahl and Kristen Nygaard at the Norwegian Computing Centre complete a general-purpose version of the language Simula, the first object-oriented language.

1967 Fairchild introduces its 3800 8-bit ALU chip.

1967 At Texas Instruments, Jack Kilby, Jerry Merryman, and James Van Tassel invent a four-function handheld calculator.

1967 Donald Knuth writes about algorithms and data structures as entities separate from the programs they are used in.

1968 A conference sponsored by the NATO Science Committee addresses the "software crisis" and introduces the term "software engineering."

1968 Edsger Dijkstra writes about the harmful effects of the goto statement, and interest in structured programming burgeons.

1967 — 1968

ment restriction or censorship of the Internet, the World Wide Web Consortium formed the Platform for Internet Content Selection to introduce protocol elements and data formats to the Web architecture.

PICS has given parents a way to refer to their choice of independent Internet content-rating services. This lets parents, rather than government agencies, define what is indecent for their children. Like the Internet and the Web, PICS is a decentralized solution.

PICS has established a specification for labels that can be read by filter software. The labels are sets of attribute-value pairs. Each label carries a URL that, when dereferenced, provides machine-readable and human-readable explanations of the semantics' attributes and possible values. Each URL points to a document written by a rating service. Parents can use these documents to obtain content-related material about various Web sites.

Some promising new ideas involve analysis not only of Web content but of people's interaction with that content. This could let the providers of material determine its quality and relevance.

You may obtain PICS labels on a CD-ROM or from a server. (PICS labels may be digitally signed so users can verify their authenticity.) A third party may also provide you with labels that describe information from a second party.

This technology may be used in other ways. PICS labels may be used for other types of content ratings (such as whether material is appropriate for adult or scholarly use). Meanwhile, the label-querying protocol may be used as an annotation-retrieval protocol, so, once deployed, it could let label servers present annotations as well as normal PICS labels.

Security and electronic commerce

The Web already permits the exchange of information, so the interchange of money is a logical next step. There are cryptographic and other schemes for a variety of payment systems. There are also various approaches toward using cryptography to ensure the confidentiality, integrity, and authenticity of Web transactions.

In the face of so many different schemes and approaches, the W3C is working on ways to encourage interoperability and competition. For example, the W3C is working on a protocol for the selection of payment mechanisms and on the Digital Signature Initiative for technology that would let people "sign" applets, public documents, and so on, for purposes of verifying authenticity.

Machine interaction

Material on the Web has principally been machine-analyzed through textual indexing by search engines. Search engines have proven remarkably useful, as they can search large indexes quickly and find obscure documents. However, their searches frequently produce a lot of junk because they generally take into account only the words documents contain and have little or no concept of document usefulness or quality.

Some promising new ideas involve analysis not only of Web content but of people's interaction with that content. This could let the providers of material determine its quality and relevance. Some of these sophisticated tools have been described as agents (a term usually used for mobile programs) because they act on behalf of the user.

Mobile agents are not widely deployed. Mobile code lets users create interesting human data interfaces (such as Java applets) and also bootstrap into new distributed applications. Mobile code could have a great impact on client and server software architecture. However, progress will be limited as long as users don't trust mobile programs (or fixed Web-searching programs, for that matter) to act on their behalf.

FUTURE DIRECTIONS

Future development of the Web will be based on

- improving the infrastructure, to provide a more functional, robust, efficient, and available service;
- enhancing the Web as a means of communication and interaction between people; and
- letting the Web contain rich data in a form understandable by machines, thus letting machines more effectively interact with the Web.

***1968** The first computers to incorporate integrated circuits—the B2500 and B3500—are introduced by Burroughs.*

***1968** A Federal Information Processing Standard encourages use of the six-digit data format (YYMMDD) for information interchange, sowing the seeds of the "Year 2000 Crisis."*


Intel Corporation

***1968** Robert Noyce, Andy Grove, and Gordon Moore establish Intel, which is incorporated on July 18.*

***1968** The Seymour Cray-designed CDC 7600 supercomputer achieves 40-megaflops performance.*

***1968** The Rand Corp. presents a decentralized communications network concept to ARPA.*

1968

Infrastructure

In the short term, protocol designers are increasing the efficiency of HTTP communications, particularly for those who use telephone modems.

There are also long-term concerns of a different nature. When the Web was designed, anyone could start an Internet server without having to register it with a central authority and without regard to the number of HTTP servers in existence. This enabled the Web to accommodate its increasing popularity.

Today, this is not enough. There are so many clients that the load on many servers sometimes becomes excessive. We need servers that can operate independently of the number of clients.

Moreover, for the Web to be a useful mirror of real life, the infrastructure must be able to cope when the number of people reading any document changes rapidly and dramatically.

Meanwhile, the business community is relying on the Web so much that server or network outages are not considered acceptable. We thus need a fault-tolerant Web architecture. This need can be addressed by the automatic (and sometimes preemptive) replication of data. We need an adaptive system that could configure in a way that would best use the available resources to optimize service quality. This complex issue includes the problems of

- categorizing documents and users so they can be handled in groups,
- anticipating the high usage of document groups by user groups,
- deciding on the optimal placement of copies of data for rapid access,
- developing an algorithm that will take a URL and find the cheapest or nearest copy of the corresponding data, and
- resolving these problems even though different parts of the Web infrastructure are funded by different groups with different priorities and policies.

Communication medium

Research into improving the Web as a communications medium has primarily concentrated on the data formats for various displayable document types: continued HTML extensions, the new Portable Network Graphics specification, VRML, and so on. There will always be new formats, so perhaps a more powerful and more consistent set of formats will eventually displace HTML.

Eventually, the Web must change in other ways to reach its potential as a medium for human communications.

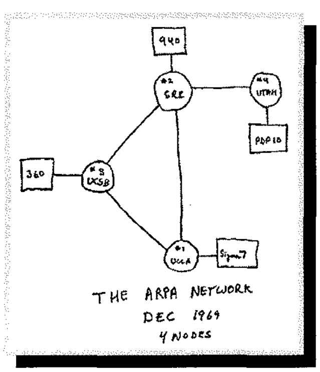
INTRANET. The most famous aspect of the Web is the corporate site that targets the general consumer population. However, the power of using the Web within an organization is increasingly being appreciated, in the form of private intranets.

Intranets must be able to control access to the information they contain. Once this has been done, the use of an intranet server increases because the participants belong to the same organization and share a level of trust. This encourages information-sharing at a more spontaneous and direct level than that found on the Internet.

When the Web was designed, anyone could start an Internet server without having to register it. Today, this is not good enough.

COLLABORATIVE USE. Web protocols could better encourage collaborative use by

- having better editors that would permit direct interaction with Web data;
- notifying interested parties when information changes;
- integrating audio and video Internet conferencing technologies;
- using hypertext links that represent in a visible and analyzable way the semantics of such human processes as argument, peer review, and work-flow management;
- having third-party annotation servers;
- enabling systems to determine whether users are members of the group to which access is supposed to be limited; and
- representing links as first class objects with version control, authorship, and ownership.



The Computer Museum

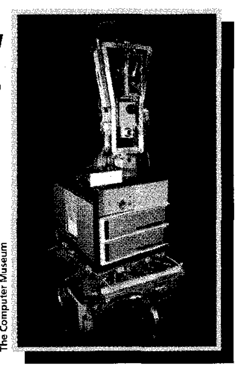
1969 Bell Labs withdraws from Project MAC, which developed Multic, and begins to develop Unix.

1969 The RS-232-C standard is introduced to facilitate data exchange between computers and peripherals.

1969 The US Department of Defense commissions Arpanet for research networking, and the first four nodes become operational at UCLA, UC Santa Barbara, SRI, and the University of Utah.

1970 Winston Royce publishes "Managing the Development of Large Software Systems," which outlines the waterfall development method.

1970 Shakey, developed at SRI International, is the first robot to use artificial intelligence to navigate.



The Computer Museum

1969 — 1970

PERSONAL INFORMATION SYSTEM. We should be able to easily use the Web as a personal information system. However, using the Web will not be easy until global and personal data are handled consistently. The computer interface that typically uses a desktop metaphor should thus be integrated with hypertext. There are not many major differences between desktop and Web clients. For example, file systems have links (aliases or shortcuts), like Web documents. Meanwhile, useful information-management objects, such as folders and nested lists, will have to be transferable in standard ways to exist on the Web.

The file name's importance will decrease until the ubiquitous file name dialog box disappears. The important aspects of information can best be stated in a title (of an article, for example), rather than a file name, and in various types of links, such as the enclosure of a file within a folder, the appearance of an e-mail address in a message's destination field, and the relationship of a document to its author. These semantically rich assertions make sense. If the user specifies essential information, such as the availability and reliability levels required for access to a document, then the system can manage disk space in a way that will maximize service quality.

For machine reasoning over a global domain to be effective, machines must be able to verify the authenticity of assertions found on the Web.

The result will hopefully be a consistent and intuitive universe of information.

Machine interaction

With the growth of commercial Web applications, machine analysis could let computers facilitate business by serving as agents with the power to act financially.

To do this, potentially useful data on the Web would have to be available in a machine-readable form with defined semantics, as in the Electronic Data Interchange.[5] In EDI, digital versions are produced of such forms as sale offers, bills of sale, title deeds, and invoices. Each form's semantics are defined by a human-readable specification document.

Alternatively, we could define general-purpose languages in which assertions could be made. Within these assertions, axiomatic concepts could be defined in human-readable documents. In this case, the language's power to combine concepts from different areas could allow development of a more powerful information space on which we could base machine-reasoning systems. Such knowledge representation (KR) languages have not had a significant impact on computer applications, but the same was true of hypertext before the Web gave it global scope.

There is a relationship between developments in the machine processing of global data and those in cryptographic security. For machine reasoning over a global domain to be effective, machines must be able to verify the authenticity of assertions found on the Web. This requires a global security infrastructure that permits documents with verifiable "signatures." Similarly, a global security infrastructure must be able to manipulate fairly complex assertions. It is perhaps this chicken-and-egg interdependence that, along with government restrictions on cryptography use, has delayed the deployment of either system.

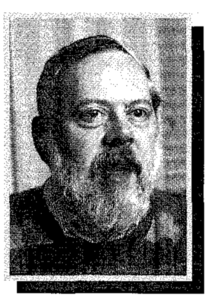
The PICS system, with its machine-readable labels, may be a first step in this direction.

ETHICAL AND SOCIAL CONCERNS

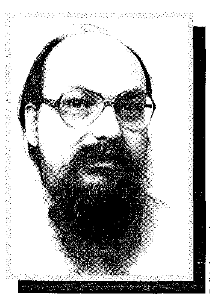
Engineers who design Web protocols have increasingly found that their work is not just academic or technical, but also raises ethical and social issues that they should address. For example, the PICS initiative showed that network protocols can affect the type of society we build within the information space.

Now, privacy is an issue. Two people talking in the middle of a wheat field have privacy. However, do they have the right to a really private conversation over the Internet? Strong cryptography could provide such privacy. However, governments have limited its use because of national security concerns.

Meanwhile, our intellectual property concepts, central to our culture, are not expressed in a way that maps onto the abstract information space. Intellectual property concerns the way authors' creations can eventually benefit readers. Independently of this, the system must be able to copy data for reasons of efficiency and reliability. The con-



Bell Laboratories



Bell Laboratories

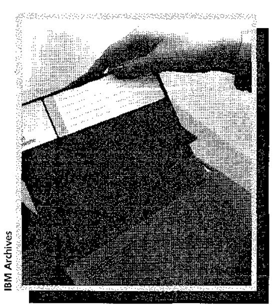
1970 Unix is developed at Bell Labs by Dennis Ritchie and Kenneth Thomson.

1970 The Computer Group News becomes Computer, a monthly magazine for all Computer Society members.

1970 RCA's MOS (metal-oxide semiconductor) technology promises cheaper and smaller ICs.

1970 Xerox establishes the Palo Alto Research Center at Stanford University for computer research.

1970 E.F. Codd describes the relational model.



IBM Archives

1970 The floppy disk and the daisywheel printer make their debut.

1970

cept of copyright as it relates to the making of copies makes little sense on the Web.

Of course, once a system has made copies of supposedly private material, there is a risk the government could legally seize those copies. Indeed, privacy on the Web can be compromised in many ways. For example, the providers of Web content want to determine the demographic makeup of the people who browse their material, but this could compromise those people's privacy.

In the long term, there are questions as to what will happen to the world's cultures when geography becomes weakened as a diversifying force. Will this lead to a monolithic world culture that is like the culture of the US, or will it foster even more disparate interest groups than exist today? Will the Web help create a true democracy by informing the voting public of the realities behind government decisions, or will it encourage ghettos of bigotry where emotional intensity rather than truth gains readership? It is for us to decide, and in doing so, we must assess the simple engineering decisions that will affect the outcome.

THE WEB, LIKE THE INTERNET, is designed to let users communicate with other users without seeing the machinery that makes the system work. If the laws of our countries can respect this, the system can continue operating that way. If not, engineers will have to learn how to design systems so that user-to-user functionality is guaranteed regardless of what happens in between.

It is more appropriate for laws to affect what people read and write, which is what really affects society, than what electrons do in wires. The latter type of laws would stunt technological development.

Meanwhile, cryptography is doing for confidentiality what TCP did for reliable delivery (providing it even when the underlying network did not do so). Additional protocols may do this for information ownership, payment, and other facets of interaction that currently cannot take place on the Web.

For the information space to be a powerful place in which to solve the problems of future generations, its integrity—including its independence of hardware, packet route, operating system, and application software brand—is essential. The laws of our countries will have to work hand-in-hand with network protocol specifications to ensure that the Web's properties are consistent, reliable, and fair. ∎

References

1. V. Bush, "As We May Think," *Atlantic Monthly*, July 1945.
2. D. Englebart, "Boosting Our Collective IQ," *Selected Readings*, Bootstrap Institute/BLT Press, Fremont, Calif., 1995.
3. T. Nelson, *Literary Machines 90.1*, Mindful Press, Sausalito, Calif., 1990.
4. F. Anklesaria et al., "The Internet Gopher Protocol (A Distributed Document Search and Retrieval Protocol)," Request for Comments No. 1436, Univ. of Minnesota Network Working Group, http://ds.internic.net/rfc/rfc1436.txt.
5. *EDI Reference Desk*, Data Interchange Standards Assoc., Alexandria, Va., 1996; http://polaris.disa.org/edi/edihome.htp.

Tim Berners-Lee *is the director of the World Wide Web Consortium and a principal research scientist at the Massachusetts Institute of Technology's Laboratory for Computer Science. His principal research interest is achieving human potential through computing and networking technology. He has published a number of journal and magazine articles about the World Wide Web and its development. Berners-Lee received a BA in physics from Oxford University.*

Contact Berners-Lee at the Laboratory for Computer Science, Massachusetts Institute of Technology, 545 Technology Square, Cambridge MA 02139; http://www.w3.org/People/Berners-Lee.