

# SCIENTIFIC AMERICAN

Permanent Address: <http://www.scientificamerican.com/article.cfm?id=long-live-the-web>



## Long Live the Web: A Call for Continued Open Standards and Neutrality

The Web is critical not merely to the digital revolution but to our continued prosperity—and even our liberty. Like democracy itself, it needs defending

By Tim Berners-Lee | Monday, November 22, 2010 | 20 comments

The world wide web went live, on my physical desktop in Geneva, Switzerland, in December 1990. It consisted of one Web site and one browser, which happened to be on the same computer. The simple setup demonstrated a profound concept: that any person could share information with anyone else, anywhere. In this spirit, the Web spread quickly from the grassroots up. Today, at its 20th anniversary, the Web is thoroughly integrated into our daily lives. We take it for granted, expecting it to “be there” at any instant, like electricity.

The Web evolved into a powerful, ubiquitous tool because it was built on egalitarian principles and because thousands of individuals, universities and companies have worked, both independently and together as part of the World Wide Web Consortium, to expand its capabilities based on those principles.

The Web as we know it, however, is being threatened in different ways. Some of its most successful inhabitants have begun to chip away at its principles. Large social-networking sites are walling off information posted by their users from the rest of the Web. Wireless Internet providers are being tempted to slow traffic to sites with which they have not made deals. Governments—totalitarian and democratic alike—are monitoring people’s online habits, endangering important human rights.

If we, the Web’s users, allow these and other trends to proceed unchecked, the Web could be broken into fragmented islands. We could lose the freedom to connect with whichever Web sites we want. The ill effects could extend to smartphones and pads, which are also portals to the extensive information that the Web provides.

Why should you care? Because the Web is yours. It is a public resource on which you, your business, your community and your government depend. The Web is also vital to democracy, a communications channel that makes possible a continuous worldwide conversation. The Web is now more critical to free speech than any other medium. It brings principles established in the U.S. Constitution, the British Magna Carta and other important documents into the network age: freedom from being snooped on, filtered, censored and disconnected.

Yet people seem to think the Web is some sort of piece of nature, and if it starts to wither, well, that’s just one of those unfortunate things we can’t help. Not so. We create the Web, by designing computer protocols and software; this process is completely under our control. We choose what properties we want it to have and not have. It is by no means finished (and it’s certainly not dead). If we want to track what government is doing, see what companies are doing, understand the true state of the planet, find a cure for Alzheimer’s disease, not to mention easily share our photos with our friends, we the public, the scientific community and the press must make sure the Web’s principles remain intact—not just to preserve what we have gained but to benefit from the great advances that are still to come.

### Universality Is the Foundation

Several principles are key to assuring that the Web becomes ever more valuable. The primary design principle underlying the Web’s usefulness and growth is universality. When you make a link, you can link to anything. That means people must be able to put anything on the Web, no matter what computer they have, software they use or human language they speak and regardless of whether they have a wired or wireless Internet connection. The Web should be usable by people with disabilities. It must work with any form of information, be it a document or a point of data, and information of any quality—from a silly tweet to a scholarly paper. And it should be accessible from any kind of hardware that can connect to the Internet: stationary or mobile, small screen or large.

These characteristics can seem obvious, self-maintaining or just unimportant, but they are why the next blockbuster Web site or the new homepage for your kid’s local soccer team will just appear on the Web without any difficulty. Universality is a big demand, for any system.

Decentralization is another important design feature. You do not have to get approval from any central authority to add a page or make a link. All you have to do is use three simple, standard protocols: write a page in the HTML (hypertext markup language) format, name it with the URI naming convention, and serve it up on the Internet using HTTP (hypertext transfer protocol). Decentralization has made widespread innovation possible and will continue to do so in the future.

The URI is the key to universality. (I originally called the naming scheme URI, for universal resource identifier; it has come to be known as URL, for uniform resource locator.) The URI allows you to follow any link, regardless of the content it leads to or who publishes that content. Links turn the Web's content into something of greater value: an interconnected information space.

Several threats to the Web's universality have arisen recently. Cable television companies that sell Internet connectivity are considering whether to limit their Internet users to downloading only the company's mix of entertainment. Social-networking sites present a different kind of problem. Facebook, LinkedIn, Friendster and others typically provide value by capturing information as you enter it: your birthday, your e-mail address, your likes, and links indicating who is friends with whom and who is in which photograph. The sites assemble these bits of data into brilliant databases and reuse the information to provide value-added service—but only within their sites. Once you enter your data into one of these services, you cannot easily use them on another site. Each site is a silo, walled off from the others. Yes, your site's pages are on the Web, but your data are not. You can access a Web page about a list of people you have created in one site, but you cannot send that list, or items from it, to another site.

The isolation occurs because each piece of information does not have a URI. Connections among data exist only within a site. So the more you enter, the more you become locked in. Your social-networking site becomes a central platform—a closed silo of content, and one that does not give you full control over your information in it. The more this kind of architecture gains widespread use, the more the Web becomes fragmented, and the less we enjoy a single, universal information space.

A related danger is that one social-networking site—or one search engine or one browser—gets so big that it becomes a monopoly, which tends to limit innovation. As has been the case since the Web began, continued grassroots innovation may be the best check and balance against any one company or government that tries to undermine universality. GnuSocial and Diaspora are projects on the Web that allow anyone to create their own social network from their own server, connecting to anyone on any other site. The Status.net project, which runs sites such as identi.ca, allows you to operate your own Twitter-like network without the Twitter-like centralization.

### **Open Standards Drive Innovation**

Allowing any site to link to any other site is necessary but not sufficient for a robust Web. The basic Web technologies that individuals and companies need to develop powerful services must be available for free, with no royalties. Amazon.com, for example, grew into a huge online bookstore, then music store, then store for all kinds of goods because it had open, free access to the technical standards on which the Web operates. Amazon, like any other Web user, could use HTML, URI and HTTP without asking anyone's permission and without having to pay. It could also use improvements to those standards developed by the World Wide Web Consortium, allowing customers to fill out a virtual order form, pay online, rate the goods they had purchased, and so on.

By “open standards” I mean standards that can have any committed expert involved in the design, that have been widely reviewed as acceptable, that are available for free on the Web, and that are royalty-free (no need to pay) for developers and users. Open, royalty-free standards that are easy to use create the diverse richness of Web sites, from the big names such as Amazon, Craigslist and Wikipedia to obscure blogs written by adult hobbyists and to homegrown videos posted by teenagers.

Openness also means you can build your own Web site or company without anyone's approval. When the Web began, I did not have to obtain permission or pay royalties to use the Internet's own open standards, such as the well-known transmission control protocol (TCP) and Internet protocol (IP). Similarly, the Web Consortium's royalty-free patent policy says that the companies, universities and individuals who contribute to the development of a standard must agree they will not charge royalties to anyone who may use the standard.

Open, royalty-free standards do not mean that a company or individual cannot devise a blog or photo-sharing program and charge you to use it. They can. And you might want to pay for it if you think it is “better” than others. The point is that open standards allow for many options, free and not.

Indeed, many companies spend money to develop extraordinary applications precisely because they are confident the applications will work for anyone, regardless of the computer hardware, operating system or Internet service provider (ISP) they are using—all made possible by the Web's open standards. The same confidence encourages scientists to spend thousands of hours devising incredible databases that can share information about proteins, say, in hopes of curing disease. The confidence encourages governments such as those of the U.S. and the U.K. to put more and more data online so citizens can inspect them, making government increasingly transparent. Open standards also foster serendipitous creation: someone may use them in ways no one imagined. We discover that on

the Web every day.

In contrast, not using open standards creates closed worlds. Apple's iTunes system, for example, identifies songs and videos using URIs that are open. But instead of "http:" the addresses begin with "itunes:," which is proprietary. You can access an "itunes:" link only using Apple's proprietary iTunes program. You can't make a link to any information in the iTunes world—a song or information about a band. You can't send that link to someone else to see. You are no longer on the Web. The iTunes world is centralized and walled off. You are trapped in a single store, rather than being on the open marketplace. For all the store's wonderful features, its evolution is limited to what one company thinks up.

Other companies are also creating closed worlds. The tendency for magazines, for example, to produce smartphone "apps" rather than Web apps is disturbing, because that material is off the Web. You can't bookmark it or e-mail a link to a page within it. You can't tweet it. It is better to build a Web app that will also run on smartphone browsers, and the techniques for doing so are getting better all the time.

Some people may think that closed worlds are just fine. The worlds are easy to use and may seem to give those people what they want. But as we saw in the 1990s with the America Online dial-up information system that gave you a restricted subset of the Web, these closed, "walled gardens," no matter how pleasing, can never compete in diversity, richness and innovation with the mad, throbbing Web market outside their gates. If a walled garden has too tight a hold on a market, however, it can delay that outside growth.

### **Keep the Web separate from the Internet**

Keeping the web universal and keeping its standards open help people invent new services. But a third principle—the separation of layers—partitions the design of the Web from that of the Internet.

This separation is fundamental. The Web is an application that runs on the Internet, which is an electronic network that transmits packets of information among millions of computers according to a few open protocols. An analogy is that the Web is like a household appliance that runs on the electricity network. A refrigerator or printer can function as long as it uses a few standard protocols—in the U.S., things like operating at 120 volts and 60 hertz. Similarly, any application—among them the Web, e-mail or instant messaging—can run on the Internet as long as it uses a few standard Internet protocols, such as TCP and IP.

Manufacturers can improve refrigerators and printers without altering how electricity functions, and utility companies can improve the electrical network without altering how appliances function. The two layers of technology work together but can advance independently. The same is true for the Web and the Internet. The separation of layers is crucial for innovation. In 1990 the Web rolled out over the Internet without any changes to the Internet itself, as have all improvements since. And in that time, Internet connections have sped up from 300 bits per second to 300 million bits per second (Mbps) without the Web having to be redesigned to take advantage of the upgrades.

### **Electronic Human Rights**

Although internet and web designs are separate, a Web user is also an Internet user and therefore relies on an Internet that is free from interference. In the early Web days it was too technically difficult for a company or country to manipulate the Internet to interfere with an individual Web user. Technology for interference has become more powerful, however. In 2007 BitTorrent, a company whose "peer-to-peer" network protocol allows people to share music, video and other files directly over the Internet, complained to the Federal Communications Commission that the ISP giant Comcast was blocking or slowing traffic to subscribers who were using the BitTorrent application. The FCC told Comcast to stop the practice, but in April 2010 a federal court ruled the FCC could not require Comcast to do so. A good ISP will often manage traffic so that when bandwidth is short, less crucial traffic is dropped, in a transparent way, so users are aware of it. An important line exists between that action and using the same power to discriminate.

This distinction highlights the principle of net neutrality. Net neutrality maintains that if I have paid for an Internet connection at a certain quality, say, 300 Mbps, and you have paid for that quality, then our communications should take place at that quality. Protecting this concept would prevent a big ISP from sending you video from a media company it may own at 300 Mbps but sending video from a competing media company at a slower rate. That amounts to commercial discrimination. Other complications could arise. What if your ISP made it easier for you to connect to a particular online shoe store and harder to reach others? That would be powerful control. What if the ISP made it difficult for you to go to Web sites about certain political parties, or religions, or sites about evolution?

Unfortunately, in August, Google and Verizon for some reason suggested that net neutrality should not apply to mobile phone-based connections. Many people in rural areas from Utah to Uganda have access to the Internet only via mobile phones; exempting wireless from net neutrality would leave these users open to discrimination of service. It is also bizarre to imagine that my fundamental right to access the information source of my choice should apply when I am on my WiFi-connected computer at home but not when I use my cell phone.

A neutral communications medium is the basis of a fair, competitive market economy, of democracy, and of science. Debate has risen again in the past year about whether government legislation is needed to protect net neutrality. It is. Although the Internet and Web generally thrive on lack of regulation, some basic values have to be legally preserved.

### No Snooping

Other threats to the web result from meddling with the Internet, including snooping. In 2008 one company, Phorm, devised a way for an ISP to peek inside the packets of information it was sending. The ISP could determine every URI that any customer was browsing. The ISP could then create a profile of the sites the user went to in order to produce targeted advertising.

Accessing the information within an Internet packet is equivalent to wiretapping a phone or opening postal mail. The URIs that people use reveal a good deal about them. A company that bought URI profiles of job applicants could use them to discriminate in hiring people with certain political views, for example. Life insurance companies could discriminate against people who have looked up cardiac symptoms on the Web. Predators could use the profiles to stalk individuals. We would all use the Web very differently if we knew that our clicks can be monitored and the data shared with third parties.

Free speech should be protected, too. The Web should be like a white sheet of paper: ready to be written on, with no control over what is written. Earlier this year Google accused the Chinese government of hacking into its databases to retrieve the e-mails of dissidents. The alleged break-ins occurred after Google resisted the government's demand that the company censor certain documents on its Chinese-language search engine.

Totalitarian governments aren't the only ones violating the network rights of their citizens. In France a law created in 2009, named Hadopi, allowed a new agency by the same name to disconnect a household from the Internet for a year if someone in the household was *alleged* by a media company to have ripped off music or video. After much opposition, in October the Constitutional Council of France required a judge to review a case before access was revoked, but if approved, the household could be disconnected without due process. In the U.K., the Digital Economy Act, hastily passed in April, allows the government to order an ISP to terminate the Internet connection of anyone who appears on a list of individuals suspected of copyright infringement. In September the U.S. Senate introduced the Combating Online Infringement and Counterfeits Act, which would allow the government to create a blacklist of Web sites—hosted on or off U.S. soil—that are accused of infringement and to pressure or require all ISPs to block access to those sites.

In these cases, no due process of law protects people before they are disconnected or their sites are blocked. Given the many ways the Web is crucial to our lives and our work, disconnection is a form of deprivation of liberty. Looking back to the Magna Carta, we should perhaps now affirm: "No person or organization shall be deprived of the ability to connect to others without due process of law and the presumption of innocence."

When your network rights are violated, public outcry is crucial. Citizens worldwide objected to China's demands on Google, so much so that Secretary of State Hillary Clinton said the U.S. government supported Google's defiance and that Internet freedom—and with it, Web freedom—should become a formal plank in American foreign policy. In October, Finland made broadband access, at 1 Mbps, a legal right for all its citizens.

### Linking to the Future

As long as the web's basic principles are upheld, its ongoing evolution is not in the hands of any one person or organization—neither mine nor anyone else's. If we can preserve the principles, the Web promises some fantastic future capabilities.

For example, the latest version of HTML, called HTML5, is not just a markup language but a computing platform that will make Web apps even more powerful than they are now. The proliferation of smartphones will make the Web even more central to our lives. Wireless access will be a particular boon to developing countries, where many people do not have connectivity by wire or cable but do have it wirelessly. Much more needs to be done, of course, including accessibility for people with disabilities and devising pages that work well on all screens, from huge 3-D displays that cover a wall to wristwatch-size windows.

A great example of future promise, which leverages the strengths of all the principles, is linked data. Today's Web is quite effective at helping people publish and discover documents, but our computer programs cannot read or manipulate the actual data within those documents. As this problem is solved, the Web will become much more useful, because data about nearly every aspect of our lives are being created at an astonishing rate. Locked within all these data is knowledge about how to cure diseases, foster business value and govern our world more effectively.

Scientists are actually at the forefront of some of the largest efforts to put linked data on the Web. Researchers, for example, are realizing that in many cases no single lab or online data repository is sufficient to discover new drugs. The information necessary to understand the complex interactions between diseases, biological processes in the human body, and the vast array of chemical agents is

spread across the world in a myriad of databases, spreadsheets and documents.

One success relates to drug discovery to combat Alzheimer's disease. A number of corporate and government research labs dropped their usual refusal to open their data and created the Alzheimer's Disease Neuroimaging Initiative. They posted a massive amount of patient information and brain scans as linked data, which they have dipped into many times to advance their research. In a demonstration I witnessed, a scientist asked the question, "What proteins are involved in signal transduction and are related to pyramidal neurons?" When put into Google, the question got 233,000 hits—and not one single answer. Put into the linked databases world, however, it returned a small number of specific proteins that have those properties.

The investment and finance sectors can benefit from linked data, too. Profit is generated, in large part, from finding patterns in an increasingly diverse set of information sources. Data are all over our personal lives as well. When you go onto your social-networking site and indicate that a newcomer is your friend, that establishes a relationship. And that relationship is data.

Linked data raise certain issues that we will have to confront. For example, new data-integration capabilities could pose privacy challenges that are hardly addressed by today's privacy laws. We should examine legal, cultural and technical options that will preserve privacy without stifling beneficial data-sharing capabilities.

Now is an exciting time. Web developers, companies, governments and citizens should work together openly and cooperatively, as we have done thus far, to preserve the Web's fundamental principles, as well as those of the Internet, ensuring that the technological protocols and social conventions we set up respect basic human values. The goal of the Web is to serve humanity. We build it now so that those who come to it later will be able to create things that we cannot ourselves imagine.

---

© 2010 Scientific American, a Division of Nature America, Inc.

All Rights Reserved.