COGNETICS CORPORATION

The Evolving Web

How Web 2.0 got here & where it's headed

Charles B. Kreitzberg, Ph. D.

2008

WWW.WEB2ANDYOU.INFO

This page intentionally left blank for double-sided copying.

Introduction

The World Wide Web was created in October 1990 and it's been growing ever since. With growth and maturity deep changes often occur. A teenager is not just a larger version of her younger self but a qualitatively different person. The same is true of businesses which change qualitatively as they pass through stages from start up to maturity.

A similar growth process has occurred with the World Wide Web. Over the past 16 years, millions of people, organizations and government have made immense intellectual, monetary and sweat investments in the Web. Now the World Wide Web has become a teenager – a Web that is different in its capabilities and potential from its predecessor, Web 1.0.

In the software industry, products are routinely labeled with version numbers. Version 1.0 is the initial release; when version 2.0 was released its users expect a new and improved product. Since the Web was not developed by a single company there is no official way to assign version numbers; nor would this make a lot of sense in the fluid, rapidly evolving World Wide Web. The term Web 2.0 is more a metaphor than a precise description. As with teenagers you may not be able to point to a precise date on which the transition from childhood occurred but it is clear that a new level of maturity has been reached.

The value of an Evolutionary View

Web 2.0 is a complex multi-dimensional concept. Our goal here is to present it in a comprehensive and actionable way. We'll start by looking at the process that brought the Web to its present state. The goal in doing this is not to document history but to identify the big ideas that are key to understanding where we are today.

How the term Web 2.0 was coined

The term Web 2.0 is generally attributed to a group of Web experts lead by publisher Tim O'Reilly[i]. While brainstorming about the Web, they realized that the nature of Web sites was undergoing a shift.

In the 1990's there was an initial "gold rush" as dotcom companies were created, often with little thought about their business model. Inevitably, there was a meltdown in 2001.

While many experts responded by predicting the demise of the Web as a viable business tool, forward thinking companies continued to develop new Web applications and sites. The dotcom meltdown cleared the landscape but the players that remained were developing rich Web sites that had a new focus. These sites typically share two things in common:

> Many people think of the Internet and the World Wide Web as the same but they are not.

1. *People went to these sites to do something.* The new sites were not just about delivering pages of text and graphics but about delivering services to their users.

2. *There was a shift in the way that people interacted with these new sites.* Web 1.0 sites were designed to deliver information to users where Web 2.0 sites obtained much of their data from users. With Web 2.0, the user often becomes a producer of information as well as a consumer of it.

O'Reilly and his colleagues started calling this new and improved Web environment "Web 2.0." The term took off rapidly. According to O'Reilly there were 9.5 million citations in Google within 18 months and, at the time of this writing, Google lists

about 163 million citations for “Web 2.0”. In typical Web 2.0 style the concept has spread virally and become a part of how people understand and strategize about the Web.

While different people may have somewhat different definitions of Web 2.0, it’s clear that a major shift has occurred. We are not dealing with “Web 1.1” but with something really new.

Web 1.0: The Hypertext Web

Computers have been around since the 1950’s but until the personal computer emerged in the early 1980’s, they were large, expensive back-office tools. Because these early computers were so expensive to produce and run, few organizations were willing to allow individuals to interact with them directly. Most processing was scheduled; as far as a possible run without human intervention; and printed reports were delivered to the user when processing was complete. There was always a desire for interactive processing but the economics did not favor it.

Over time, organizations began to develop programs which allowed users to remotely access the computer through “dumb terminals” (affectionately known as “green screens” for the black and green characters they displayed). Because these terminals were not computers, the only interactions that could be supported were well-defined business transactions like making airline reservations. Tasks like word processing were far too expensive for this hardware.

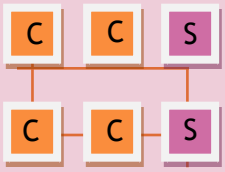
The situation changed radically when personal computers were introduced. Advances in chip manufacturing made it possible to produce low cost computers and the personal computer was created. PC’s made is possible for one user to work with an entire computer, rather than sharing it with others as had been previously required. Interacting with a program became the norm.

Networking

In the beginning, personal computers were stand-alone devices but they soon were networked together so that they could share files with each other. Early computer networks were very small.

Some networks were limited to a building or even to computers in the same room; others might link computers that were continents apart and were connected by very expensive communications lines. But each network was like an island, self-contained and isolated.

The stand-alone PC was quickly connected into networks. The network made a distributed computing model possible in which the PC's that supported the end user were called clients.



Through the client, a user could request files and data from connected computers, called servers.

Client Server Computing

Computer networking increased the power of computing in an interesting way. Once computers were networked, it was no longer necessary for a single computer to do all the work. One computer could send a message to another computer across the network, requesting that a program be run and the results delivered back over the network.

This very powerful model was called *client-server computing* because the computer requesting the program (the client) could call on the combined power and resources of all the computers (the servers) with which it was connected. It didn't matter how far away the client and server computers were physically. As long as they were on the same network the servers could deliver results to the client.

Client-server computing turned the personal computer into a powerful business tool. With it, the PC could interact with a company's back-office computers. And if development had

stopped there, there would be no World Wide Web today. But two big ideas came together and changed everything.

The birth of the Internet

The first big idea was the Internet. Many people think of the Internet and the World Wide Web as the same but they are not. The Internet is the technical infrastructure that allows computers to communicate with each other; the World Wide Web is built on top of the Internet which came first.

Did you know?

A similar interconnection problem existed in the early days of telephones. According to AT&T, "between 1894 and 1904, over six thousand independent telephone companies went into business in the United States, and the number of telephones boomed from 285,000 to 3,317,000….But … subscribers to different telephone companies could not call each other. This situation only began to be resolved after 1913."

Having each computer network being an isolated island was limiting. With such networks, for example, you could send an email to others on the same network but not outside. That was a serious limitation as business was starting to collaborate with partners. The isolated nature of computer networks made it expensive and complicated for businesses that wanted to exchange data with vendors, suppliers and customers. The solution came not from the business world but from the military.

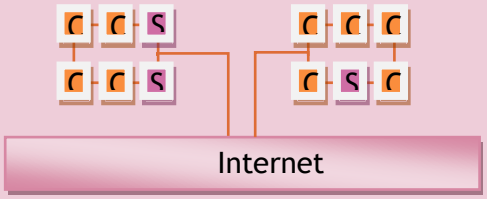
In 1973 a group of researchers launched a military project in 1973 to find a way for individual networks to communicate with each other. It was called the "Internetting Project" and the collection of connected networks that emerged from the project was called "The Internet[ii]."

Protocols

Connecting computers across different networks was less about the hardware than it was about everyone communicating in the same way. The core need was for a set of standards that would allow programs on one computer to communicate with programs on another. Engineers call these standards a *protocol* and devices that share the same protocol can communicate with each other.

Lack of a common protocol, for example, prevents many cell phones from working in both the United States and Europe.

The research team working on the Internet created a protocol for connecting computer networks which they called TCP/IP. TCP/IP is similar in many ways to the phone network. Each computer is assigned a unique phone-like number called a ten-digit *IP Address* like 68.12.188.233. If you know a computer's IP address, you can contact the computer and ask it for data much as you can contact another phone by dialing its number.



Networks were isolated islands at first - until a group of researchers created the Internet which allowed computers to communicate with each other. This effectively created a computer utility much like the telephone network.

Instead of phone numbers, each computer was assigned an IP address. Servers were given easy to remember "nicknames" (domain names) like www.server.com.

Because it is difficult to remember IP addresses, a system to use easier to remember domain names was invented in 1983. The Domain Name System (DNS) which is now part of the Internet allows people to use easier to remember names like www.mycompany.com; a software program looks up the appropriate IP address automatically and routes the message to the appropriate computer.

Hypertext and the World Wide Web

The Internet was created in the 1970's and while it found application in a lot of organizations, its overall impact was modest. In 1989, there were about 100,000 hosts (servers) on the Internet[iii] -- a significant number but a far cry from the 541,677,360 hosts counted on January 1, 2008[iv]. The problem was that the Internet was not easy to use.

While it had become straightforward to get two computers to communicate with each other, there wasn't a simple computer program that made it useful. As a result, the Internet in 1989 was still a tool of programmers and the back office. What was needed was a simple way for people to access it and that came from the work of Tim Berners-Lee, a scientist at CERN – a European organization dedicated to basic research in atomic physics. The breakthrough was based on a simple but elegant way to organize textual information known as *hypertext*.

Hypertext and its links – now familiar to every Internet user – was a curiosity in 1989.

Hypertext enables navigation

The term "hypertext" was coined by Ted Nelson in 1965 and only made sense for text displayed on a computer. In hypertext, key words or phrases are highlighted and treated as links; clicking on them takes the reader to related information. Hypertext and its links are, of course, familiar to every Web user today but in 1989 it was more of a curiosity. But a number of companies were working on hypertext programs and the idea was getting better known.

Hypertext was a great idea for navigating. As compared to menu and other types of control structures, embedded hypertext was easy to understand because every link was in context.

Hypertext was also an amazingly powerful knowledge tool. You can express an idea in a paragraph and link to related ideas, expansions, reference and resources easily. So hypertext

combined information delivery and screen navigation with a simple user action - clicking on a link.

The Browser is born

Tim Berners-Lee was intrigued by hypertext. He saw it as a convenient way to organize large amounts of information. In 1989, he wrote a proposal to CERN to develop a hypertext system. He decided to use a client server network and place the text on a server so any user could access it through a program on their PC. He called the program on the user's PC a "browser."

There were a number of other hypertext systems developed around the same time but Tim Berners-Lee's great insight was to combine hypertext with the Internet so any computer could connect to any server. In his proposal, he pointed out that the ability to access the data from any type of computer would be not only useful to CERN but would be "a boon for the world in general." That might qualify as one of the great understatements of the last century.

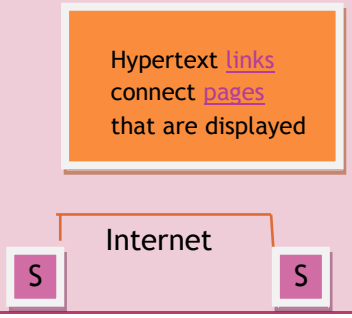
Most of the pieces for the system that Berners-Lee wanted to develop were in place. The Internet was well-established and easy to access thanks to a project by the National Science Foundation that made it available free to educational and research institutions. All that he needed to develop was a way to organize the hypertext. He accomplished that by defining a way to mark the links in text – he called this scheme the Hypertext Markup Language or HTML. He wrote a browser and a server program in 1990 and called his new system the "World Wide Web[v]."

The Hypertext Web takes off

Business was quick to see the value of the Web. Product catalogs and brochures could be placed on-line and delivered practically without cost to anyone who connected to their server. With a bit of technical legerdemain it was possible to link Web servers to back-end computers so e-commerce was possible; companies like Amazon and EBay were established. As the number of Web sites

proliferated, companies like Yahoo and Google that indexed the sites so users could search them, became major businesses.

The World Wide Web was designed as a way of using the Internet. It was based on hypertext, allowing the user to link from page to page. Each page comes from a server linked to the Internet.



By 1992, the 100,000 hosts on the Internet had grown to a million. Within the next 15 years, Web 1.0 – *The Hypertext Web* matured and became more capable. The Browser became a powerful and complex application. While the user interface has retained much of the point and click simplicity of the initial Web, the Browser has become a highly sophisticated presentation tool. Millions, tens of millions and ultimately hundreds of millions of servers were added to the Web. As of December 2007, there were 1.3 billion users representing 20% of the world population[vi].

Web 2.0: The Social Web

Web 1.0 is called "The Hypertext Web" because it is based on a model of presenting pages to the user using hypertext links. Over the years since its introduction, as the browser's ability has grown, the page display has become more sophisticated. A programming tool called Flash has allowed pages to contain animated graphics. Page layout tools have become more capable and some Websites offer tens of thousands of pages of information.

The sheer vastness of its information is awesome. But the basic model has not changed. It is largely a one-way Web in which the

user requests pages, the server delivers them and the browser displays them.

Of course there are many examples of Web sites that do more than just present pages. Companies like Amazon and EBay are built around more complex processing. But they have to use special techniques to accomplish this and create custom programs to connect their processing applications to their Web servers. It takes a lot of work and investment to make this happen.

Programming Platforms

The ability to create fully functioning software that runs over the Web changes the way applications are developed, delivered and used.

Programmers have been working to extend the capabilities of Web servers so that become powerful programming platforms that can run industrial strength applications. It's required a lot of time and investment to get there but here are now several of these programming platforms available. Two of the best known are Sun's Java (also supported by IBM) and Microsoft's .NET platforms. With either Java or .NET it is possible to create fully-functional software that runs over the Web.

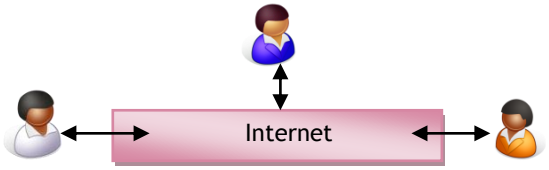
It is difficult to overstate the importance of this development. With these programming platforms (also called *application frameworks*) it is possible to write any desired program and run it over the Web within a browser. This is just an extension of the client server model that was used in Web 1.0 (and before); the browser continues to be the client while the application is run on the server. But the increased capability of the server to run applications changes the way that software is developed, delivered and used.

Traditionally, software has been distributed on electronic media such as floppy disk or CD (or perhaps downloaded) and installed on a computer's hard drive. If you wanted to share an application with hundreds, thousands or even millions of users, you would need to provide each one with a copy of the software and get

them to install it on their computer. This raises a host of logistical, security and user support issues that makes it almost impossible to deliver applications on a wide scale without a lot of expense. Perhaps more importantly, from a business perspective, installing a software product creates a hurdle that the customer would need to be willing and able to cross; it's a limitation that makes it really hard to distribute software widely.

Web 2.0 resulted from the maturation of Web technology, especially increased capability in the programming platforms. This makes it possible for Web sites to function as true applications.

As a result, users are able to perform tasks, create data and communicate and collaborate using the Web. On-line communities have become important both within organizations and externally.



Software as a Service

With Web applications, all this mess disappears. Users simply point their browser at a specific Web site and they can run the software without needing to install anything. If the software is upgraded, users automatically get the latest version. This model is called *Software as a Service* (SaaS). Many experts believe that it will become the universal way that people access software.

Google is offering a suite of Web applications called Google Apps that includes email, a calendar and a word processor. Microsoft is offering its own suite of Web applications called Office Live. We

are seeing the results of this technology advance all around us. Sites like YouTube allow users to upload videos and share them with the world. Flickr does the same for photos.

Collaboration tools, communities of all sorts and business tools are changing the nature of the Web. With Web 2.0, it's a new game.

Community

When people think about Web 2.0, it is usually the social aspect that comes to mind first. With Web 2.0 we have seen an explosion in blogs, wikis, discussion forums, webcasts, instant messaging and social networking sites like LinkedIn, Facebook and MySpace. Each of these involves some form of on-line community.

These tools are entering the business community with both external (for example, customers) and internal (employee) focus. Organizations are starting to think in terms of communities and neighborhoods and how they can relate to them.

Web 3.0: The Semantic Web

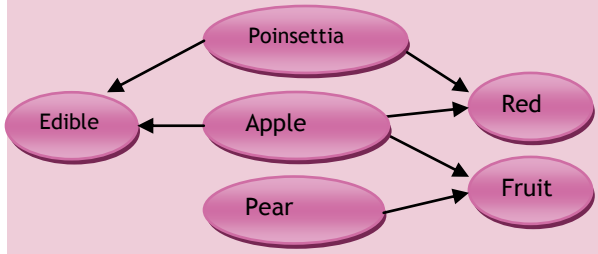
Tim Berners-Lee's vision for the World Wide Web involved more than simply the hypertext pages that were the basis for Web 1.0. But while hypertext shows the relationship among texts through links, it does not help you understand what the text actually means.

To see this limitation, type an ambiguous phrase like "press a suit" into a Web search engine. Here are three results that result:

> "Finally, the judge rejected the argument that the board would not **press a suit** in which…"
>
> "…a personal valet is available to **press a suit**, mend a button, draw a bath…"
>
> "…knights can deliver a rough challenge without ceasing to be chivalrous, and their ladies can **press a suit** without ceasing to be ladylike"

In the first example, the phrase “press a suit” refers to a lawsuit. In the second, it refers to ironing a suit of clothing; while in the third, the phrase refers to a romantic relationship. These are three quite different meanings; the search engine was unable to distinguish among them. This ambiguity is the problem that Web 3.0, the semantic Web, is being designed to overcome.

With some 500 billion pages on the Web, searching text without understanding what it means cannot be efficient. Of course, computers cannot really understand meaning. With so much information available, users want to quickly locate what is most relevant to them.



Web 3.0, the Semantic Web, reflects a new set of capabilities that are currently being worked out by researchers and are being translated into applications.

It will provide ways to make use of the meaning of data (its “semantics”) for searching and processing.

Like Web 1.0, the core of the Semantic Web is a way of marking up data so that computers can process it more efficiently. We believe that the Semantic Web will have far reaching effects on the way that information is managed and retrieved. But much of this is still in the very early stages. And marking up existing data will be an enormous task. So we don’t expect that Web 3.0 will be a factor that business needs to focus on for a few years.

References

---

[i] (O'Reilly, 2005)
[ii] (Cerf, 2007)
[iii] (Slater, 2002)
[iv] (Internet Systems Consortium, 2008)
[v] (Jacobs, 2004-2008)
[vi] (Miniwatts Marketing Group, 2007)

For more information, contact the author:

Charles B. Kreitzberg, Ph.D.

CEO

Cognetics Corporation

PO Box 386

Princeton Junction, NJ 08550

609.799.5005 ext 235

Charlie@cognetics.com