The World Wide Web, a Paradigm of Innovation

Peter J. Denning

10/12/03

For the CS4900 course

CS4900 teaches the practices of innovation. Learning these practices is part of the educational requirements for the entire NPS. Because innovating is one of the four central practices of computing professionals, the CS Faculty have included the lab courses 4900 and 4901 in the required curriculum. The story of the World Wide Web, a famous innovation in which computer scientists participated, richly illustrates the key elements of innovation practice. This story will help you understand how we came to the course content, syllabus, and learning objectives.

The WWW Story

In 1960 J C R Licklider of MIT published a visionary paper describing a future in which a worldwide network would connect all computers. He envisaged resource sharing, ubiquitous computing, computing dispensed as utilities, new approaches to research, and new kinds of businesses. In 1964 Paul Baran of the Rand Corporation prepared a series of classified papers on the architecture of a new kind of network of computers, a packet-switched network, that would survive enemy attack. In 1967 ARPA undertook a research project to build a network using Baran's architecture that would realize Licklider's vision. The first four nodes of the ARPAnet became operational in early 1970.

By 1981 the ARPAnet had grown to about 200 nodes and developed the basic suite of protocols for an Internet (TCP/IP, FTP, Telnet, SMTP). During this time the Europeans undertook their own networking projects and spearheaded the development of the ISO (International Standards Organization) seven-layer model of protocol architecture. In the late 1980s, the NSF began partnering with ARPA to provide a backbone connecting the NSF supercomputing centers and to establish a new policy that opened the resulting network to commercial traffic. By 1990, the emerging Internet had grown to over 150,000 computers and was expanding exponentially. Through decades of tireless effort by the Internet Society and its IETF (Internet Engineering Task Force), the ARPA protocol model became the world standard for data networks.

In the 1970s, network visionary Ted Nelson envisioned authors using networks for publishing, replacing paper and transforming all systems of publication. He proposed a method he called hyperlinks between documents that enabled a reader to jump among documents as easily as moving to the next paragraph in the same document. Nelson's vision was spoken of frequently in academic circles as part of a future Internet, but remained in the status of a dream well into the 1980s, when AutoDesk bought his company, Xanadu, and made the software public.

In 1980, the American National Standards Institute (ANSI) formed a committee on computer languages for processing of text. That group proposed a Standard Generalized Markup Language (SGML) in 1985. Publishers traditionally marked up author's manuscripts with special symbols to tell the printers how to set up the typesetting machines properly. SGML is a meta-language used to describe the grammars of specific markup languages. It offers a standard notation and distinguishes between text objects and format specifiers. With SGML, it became possible for authors, editors, and publishers to process and exchange documents automatically.

Early in the 1980s researchers in high-energy physics (HEP), who were early ARPAnet users and were advocates of high-bandwidth connectivity among supercomputers, established bulletin-board services to exchange preprints of their research papers. This open exchange greatly accelerated the progress of HEP research. Within a few years they had an extensive repository of research papers stored on servers in several countries. As the usage expanded, the method of obtaining a paper from a download server became unwieldy and caused many complaints. To obtain a paper, a researcher wrote down the citation from another paper obtained on-line; obtained the IP address of that server; searched the directories of that server for the proper file; opened an FTP (file transfer protocol) session to that server, navigated to the proper directory, and downloaded the cited paper, closed the FTP connection; and finally opened the file in a word processor. To many physicists, the download method was becoming a hindrance to effective communication.

The intersection of these four communities of practice -- networking, hypertext, digital publishing, and research paper exchange in HEP community -- was the crucible in which the World Wide Web was born. The HEP need for a better method of sharing papers became the driving force for a solution derived from the combination of networking, hypertext, and digital publishing. Tim Berners-Lee was the man who did this.

Berners-Lee was a computer scientist who specialized in networking, was imbued with the visions of Licklider and Nelson, participated in the SGML users group, and worked at the CERN high-energy physics lab in Switzerland. Around 1988, he saw a way to create a new network protocol that would take a citation and automatically download the cited paper from a remote server and display it locally. He viewed a citation as a link to another document and called the protocol HTTP (hypertext transport protocol). He invented HTML (hypertext markup language) to permit authors to markup up their research papers so that hyperlinks could be located automatically by HTTP. He proposed to build such a system and thereby solve the problem his physics colleagues' paper exchange problem. He and his team called their system the World Wide Web.

Selling research authors on the idea that they should insert tags into their texts to identify hyperlinks was easy. Many researchers were already using TeX (Knuth's mathematical typesetting language), which inserted many tags into their documents.

During 1989 to 1992, the WWW spread among the HEP community and into some other research communities. One of research centers using the WWW was the National Center for Supercomputing Applications (NCSA) at the University of Illinois. Under the direction of Larry Smarr, himself a physicist, the NCSA emphasized interdisciplinary research projects that had strong graphics components. One of the graduate students, Marc Andreesen, decided to build a graphics interface for the WWW. He invented the Mosaic browser and offered it as a means to share documents anywhere in the Internet. His browser showed hyperlinks in a highlighted color and translated a mouse click on a hyperlink into a call on the HTTP protocol for following hyperlinks. Andreesen left NCSA and founded the Netscape company (originally Mosaic) to sell the browser. The browser enabled many people to rapidly use the WWW, leading in a few years to Internet and Web becoming household words and "http://" addresses on every business card and advertisement.

The story does not end with Netscape's success. Berners-Lee, conditioned by his experiences with international standards bodies and the Internet Society, saw that the WWW could quickly become chaotic if all the new players could not agree on standards for the new extensions and applications they were building. Following the example for the Internet Society, which was a grass-roots quasi-governance organization for the evolution of the Internet, Berners-Lee founded the World Wide Web Consortium (W3C) in 1994 and has chaired it since its inception. The W3C, which has 400 member organizations, develops common protocols that promote the Web's evolution and ensure its interoperability. The W3C has taken a leadership role in developing standards for Web protocols and for the extensions of the HTML now known as XML (extended markup language). The W3C has been a steadying influence helping the WWW grow while remaining open and available to all comers.

Ingredients of Innovation

From the story we conclude that these conditions made for the innovation:

- WWW occurred at the intersection between the domains of high energy physics, computer science, hypertext documents, and digital publishing. The WWW combined practices and knowledge from these four.
- Berners-Lee embodied the four domains from having practiced in each of them. He appropriated practices from each into the WWW.
- As with many other innovations, the WWW created a space in which many other new innovations could appear. For example, the innovation was extended at NCSA by another innovator, Andreesen, who himself stood in multiple domains: graphics, physics, and WWW. He created the Mosaic browser, a GUI for the WWW. Within a few years, the "dot com"

boom started, an explosion of business interest in using the WWW and Internet as a medium for marketing, buying, and selling products.

Berners-Lee demonstrated all five of the components of the innovation practice identified by Peter Drucker:

- *Seeing opportunity:* he saw an opportunity in the process need felt by many within the HEP community.
- *Analyzing it:* he designed a markup language HTML, to enable marking and automatic display of documents; a universal system to name objects globally; a protocol http to move hyperlinked documents when clicking a mouse on a citation link. He also proposed ways to prototype and deploy his system.
- *Listening to the community:* he listened that the concern for effective paper sharing was deeply embedded into the HEP community's psyche. He listened to feedback he received from his proposals and prototypes. He established the W3C as an institutional means to continuing listening to the community.
- *Focusing:* He kept himself and his team focused on building the system that would solve the original problem. He did not allow himself to be sidetracked by other interesting problems or by premature opportunities to expand the WWW functionality. He did not overextend himself or overestimate his capabilities. He simplified the set of concepts that everyone had to understand and did not deviate from the basic simplicity he advocated.
- *Leading:* He designed the WWW technology to appeal to many other research communities besides HEP. He established the W3C to provide ongoing leadership in setting WWW standards. He exhibited many qualities of personal leadership in advancing his proposals, mobilizing people to work on them, and in moving all projects toward successful delivery of their products.

In addition to following the elements of the innovation process, Berners-Lee exhibited mastery of certain basic linguistic-somatic practices. Linguistic means "in language", in other words causing action through language. Somatic means "in the body", in other words that the practice is learned, ingrained, and habitual and can be exercised without conscious thought. The linguistic-somatic practices are:

- Making and grounding assessments (of the situation, ofthe team's capabilities, and of one's personal capabilities)
- Getting and holding attention
- Conversations for possibility (setting context and purpose for action)
- Conversations for action (action loops that complete action)
- Earning trust
- Making powerful, embodied declarations

- Choosing the right team
- Managing commitments (team and personal)
- Learning, developing competence in practice
- Focusing (staying centered)
- Listening to the community
- Taking the lead

This course

The course contains two coordinated, parallel segments. The **conceptual track** focuses on the ideas involved, such as Drucker's 7 opportunities or 5 practices, Drucker's do's and don'ts for innovators, the role of knowledge creation, appropriation of practices across domain boundaries, the theory of action loops, and linguistic interpretations of key elements of innovation. The **linguistic-somatic track** focuses on exercises that teach the practices listed above, and helps embody them. These tracks are integrated: the practices support the concepts.

In addition we occasionally invite outside speakers who are innovators and can discuss their own practices of innovation.

The course is in two parts (4900 and 4901) because we need 4 contact hours to cover all the material and we want to allow two full quarters for students to engage with the basic practices and come to embody them. In both parts, the students will conclude with a declaration of their thesis area (or proposal).

These courses also support students preparing for doing an innovative thesis project. Innovation is not the invention of good ideas; it is work. We teach what the nature and content of that work is.

References

Licklider, J C R. 1960. Man-computer symbiosis. *IRE Trans. Hum. Factors Electron*. 1:4-11.

Baran, Paul. Collected papers on distributed communication. The Rand Corp. http://www.rand.org/publications/RM/baran.list.html

Nelson, Theodor. 1974. *Dream Machines.* Apeture Publishers.

Nelson, Theodor. 1993. *Literary Machines.* Mindful Press. Originally published in 1987.

Denning, Peter. 1997. The ARPANET After Thirty Years. In *Internet Beseiged* (D. Denning & P. Denning, eds.). Addison-Wesley.

Berners-Lee, Tim. 1996. "The Web Maestro: An interview with Tim Berners-Lee." *Technology Review* (July).

Berners-Lee, Tim. 1996. "WWW: Past, Present, and Future." *IEEE Computer 29*, 10 (Oct), 69-77.

SGML User's Group. 1990. History of SGML. http://www.oasis-open.org/cover/sgmlhist0.html

Drucker, Peter. 1986. *Innovation and Entrepreneurship.* Harper Business. (First published by Harper Perennial in 1986.)