CO-305

TWENTY YEARS OF THE WORLD WIDE WEB: PERSPECTIVES ON THE INTERNET TRANSITION IN CARTOGRAPHY

PETERSON M.
University of Nebraska at Omaha, OMAHA, NE, UNITED STATES

ABSTRACT

The World Wide Web Internet protocol was introduced twenty years ago in 1991. Although the system originally conceived by Tim Berners-Lee consisted only of text, its underlying hyperlinking structure unleashed a revolution in the dissemination of information. Two years later, Andreesen and Bina released the World Wide Web Mosaic browser that incorporated the display of graphic files and the Internet became a new medium for cartography. It is now clear that the Internet represents much more than simply a new way of distributing maps to users. The Internet has changed how map data is acquired, how maps are stored, manipulated, and used. The International Cartographic Association recognized this important change in creating the Maps and the Internet commission in 1999.

The Internet has highlighted the importance of maps to humans in a variety of ways. A considerable amount of Internet traffic can be associated with the distribution of maps. Portals and search engines have promoted maps and invested considerably in their creation and use. In addition to searching web pages, it is now also possible to search for features on maps with services like Google Maps. There are signs that Google maintains the individual tiles for its map in computer memory at its many data centers. It is estimated that the company invests as much as US$629 million to store all map tiles at each center.

The current Internet cartographic landscape has been shaped by a number of polarizing forces. Foremost among these is the battle between commercial and open-source (public) interests. Initially an open network, the Internet has been increasingly commercialized. Open-source software and open-sourced data are efforts to counter commercial interests. Another polarizing aspect of Internet maps is the form of delivery. While desktop computers have been increasing the size of the display, many maps are now delivered on small mobile units. Browser vs. non-browser forms of map delivery would represent a further polarizing aspect. Many map "apps" for mobile devices no longer use a browser.

The future of cartography is inextricably tied to the Internet. The current from of cartography on the Internet is a product of a series of polarizing interests. These countering forces will likely continue. If history repeats itself, the future will be much like the past.

I. INTRODUCTION

Technological change is not new to cartography. In many ways, the making of maps has always been subject to new developments in reproduction and distribution. But, the change brought by the Internet over the past 20 years, particularly since 1993, has been fundamentally different. To put it simply, the Internet has changed the map (Peterson 2003, Peterson 2007, Cartwright, et. al. 2006, Gartner, et. al. 2007).

The change was subtle at first because Internet maps were not that different than the maps that began to appear during the 1980s on CD. In some ways, Internet maps were not as useful because the Internet was slow at first and the maps took longer to display. But, maps through the Internet were free and CDs faded abruptly as a form of map distribution in the mid-1990s.

Internet maps have developed into a uniquely interactive products and have broadened their distribution to small, mobile devices. These devices have added yet another important for maps – a symbol indicating the current position of the user. Derived either from GPS or cell phone tower triangulation, the dot indicating the current position is meaningless without the map. In looking to the future of cartography, it is useful to look at the recent past and the progression of maps and the Internet.

II. EARLY DEVELOPMENT OF THE INTERNET

The beginnings of the Internet can be found in ARPANet – a computer network created for the Advanced Research Projects Agency and funded by the U.S. Department of Defense. The purpose of the network was to help scientists work together and to create a network with a redundantly linked structure that would continue to work even after a limited nuclear attack. ARPANet switched from the NCP protocol to the currently used TCP/IP (Transmission Control Protocol/ Internet Protocol) on January 1, 1983. Many view this date as the beginning of the Internet (see Figure 1).

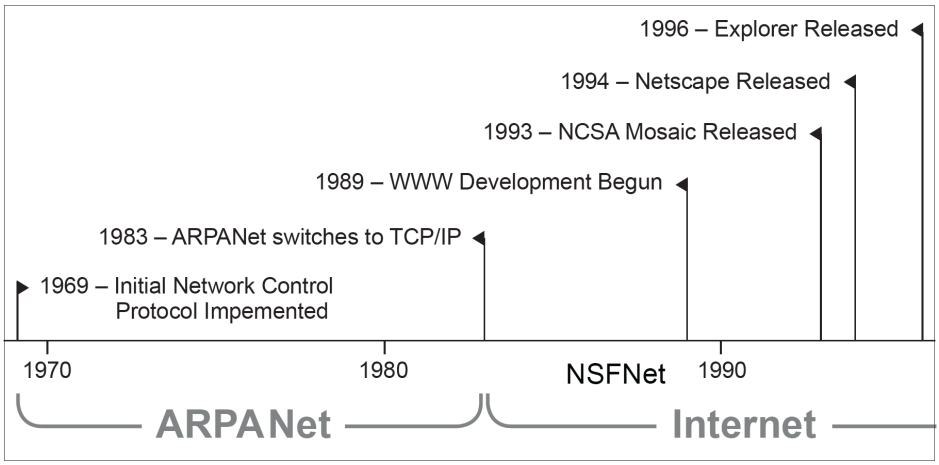

Figure 1. Timeline of early Internet development. Initial work was begun in 1969. The introduction of NCSA Mosaic in 1993 spurred the development of other browsers and the widespread use of the World Wide Web.

Conceived at the European Particle Physics Laboratory (CERN) located near Geneva, Switzerland, in 1989 and introduced in 1991, the World Wide Web was intended to assist researchers in high-energy physics by linking related documents. The developers wanted to create a seamless network in which articles on physics could be accessed in a simple and consistent way. The system could only depict the textual part of the articles.

On demand, web maps began appearing in 1993 soon after the introduction of the Mosaic browser that incorporated the display of graphics. One of the first of these online mapping programs was developed by Steve Putz (1994) at the Xerox Palo Alto Research Center (PARC). His Map Viewer program allowed the user's, client computer to create on-demand maps from a geographic database. Each interaction with Map Viewer would request a new map from the server that was zoomed in on a specific point. Individual maps were generated in a graphic file and embedded into a web page.

Almost exactly four years after the introduction of Map Viewer, a new era in online mapping was introduced with MapQuest's user-defined street maps in 1996 (see Figure 2). Using a simple client-server model, the site combined maps with route-finding. MapQuest quickly became the largest publisher of maps on the Internet responding to millions of daily requests. Developed by a map publishing company called GeoSystems, MapQuest was a major success and was purchased by the Internet giant AOL in 1999 for $1.1 billion dollars. MapQuest held the distinction of having the greatest market share among mapping sites until 2009, competing against mapping services from Yahoo!, Microsoft (Bing), and eventually Google. Google Maps, introduced in 2005, transformed the online map into an extension of a search engine thereby making it possible to search for features on the map.

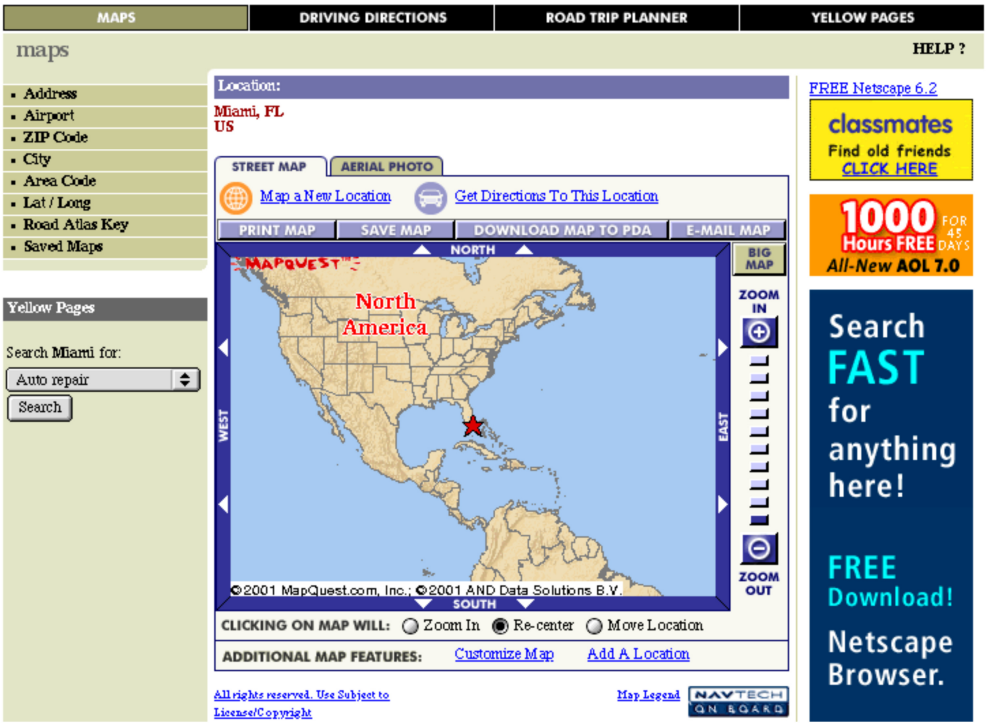

Figure 2. A 2001 version of the MapQuest webpage. Dominated by ads, the map constitutes only a small part of the page.

Businesses developed a new model for financing the production and distribution of maps. Initially, online maps were financed by small advertisements that appeared on the web page with the map. Google began to use maps to augment their search engine and provide a way for people to find the location of businesses. A search done through the map can result in income for Google as businesses pay to be found. The map itself is provided for free because it simply serves as a background to display sponsored information.

Figure 3 depicts the proportion of Internet traffic in the United States that can be attributed to the various protocols from 1990 to 2010. World Wide Web traffic grew rapidly throughout the 1990s and peaked in 2000. While absolute traffic is still increasing, the proportion of Internet traffic attributed to the web has declined as a result of the growth in the transmission of video. Internet video streaming has become a major way of distributing movies, surpassing video rentals on DVD. The size of video files is much greater than the amount of data required for a typical web page.

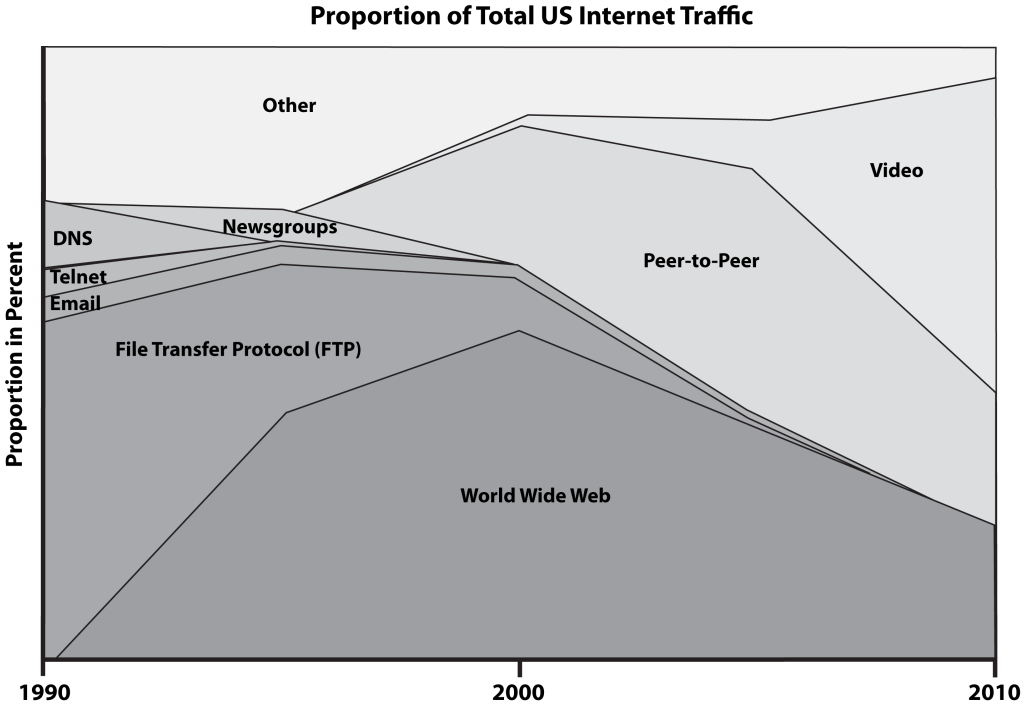

Figure 3. The proportion of Internet traffic from the various protocols. While World Wide Web traffic is still increasing, its proportion of Internet traffic in declining in comparison to video. Peer-to-peer is used mainly for distributing video content along with music.

III. WEB 2.0 MAPS

Beginning in about 2004, Web 2.0 represents a variety of innovative resources, and ways of interacting with, or combining web content. Web 2.0 includes the concept of wikis, such as Wikipedia, blog pages, podcasts, RSS feeds, and AJAX. Social networking sites like MySpace and Facebook are also seen as Web 2.0 applications.

A major part of Web 2.0 is Google Maps, introduced in 2005. The mapping service is based on two major ideas: 1) image tiling; and 2) AJAX. Image tiling had been used since the early days of the World Wide Web to speed the delivery of graphics. In comparison to text, images require more storage and therefore take longer to download. A solution was to divide the image into smaller segments, or tiles, and send each tile individually through the Internet (Sample 2010). These smaller files often travel faster because each can take a different route to the destination computer. On the receiving end, the tiles are reassembled in their proper location on the web page. With a moderately fast Internet connection, all of this occurs so quickly that the user rarely notices that the image is actually composed of square pieces. With slower connections, the individual tiles are clearly evident.

The second major innovation brought by Google Maps was the incorporation of Asynchronous JavaScript and XML (AJAX) in the relationship between the server and client. This was the culmination of many years of effort to re-shape interaction on the Internet. Essentially, AJAX maintains a continuous connection with the server – exchanging small messages in the background even when the user has not made a specific request. This allows for faster server responses when the user does make a request. AJAX might be thought of an application that works in the background of a browser to anticipate what the user might want and be ready to communicate with the server to respond to a request. Operations in Google Maps that are particularly assisted by AJAX include zooming and panning, a common form of interaction with maps.

All map tiles are 256x256 pixels and require about 15 KB a piece to store in the PNG format. Table 1 shows the number of tiles that are used in a tile-based mapping system for 20 levels of detail (LOD), or zoom levels, and the associated storage requirements and storage costs. With 20 LODs, there are a total of approximately 1 trillion tiles for the whole world. At an average of 15 KB per tile, the total memory requirement is 20 Petabytes, or 20,480 Terabytes. No single client computer could have this much storage capacity.

Table 1. The number of tiles, storage requirements, and storage costs used by a tile-based online mapping system to represent the world at different levels of detail (LOD) or zoom levels.

Levels of Detail (LOD)	Number of Tiles	Ground distance per pixel in meters	Storage requirements at 15 Kilobytes per tile	Disk storage costs at $100 per Terabyte	RAM memory storage costs at $30 per Gigabyte
1	4	78,272	60 Kilobytes (KB)	$0.000006	$0.002
2	16	39,136	240 KB	$0.00002	$0.007
3	64	19,568	968 KB	$0.0001	$0.03
4	256	9,784	3.75 Megabytes (MB)	$0.0004	$0.11
5	1,024	4,892	15 MB	$0.001	$0.44
6	4,096	2,446	60 MB	$0.006	$1.76
7	16,384	1,223	240 MB	$0.02	$7.03
8	65,536	611.50	960 MB	$0.09	$28.13
9	262,144	305.75	3.75 Gigabytes (GB)	$0.37	$112.50
10	1,048,576	152.88	15 GB	$1.46	$450.00
11	4,194,304	76.44	60 GB	$5.86	$1,800.00
12	16,777,216	38.22	240 GB	$23.44	$7,200.00
13	67,108,864	19.11	968 GB	$93.75	$28,800.00
14	268,435,456	9.55	3.75 Terabytes (TB)	$375	$115,200.00
15	1,073,741,824	4.78	15 TB	$1,500	$460,800.00
16	4,294,967,296	2.39	60 TB	$6,000	$1,843,200.00
17	17,179,869,184	1.19	240 TB	$24,000	$7,372,800.00
18	68,719,476,736	0.60	960 TB	$96,000	$29,491,200.00
19	274,877,906,944	0.30	3.75 Petabytes (PB)	$384,000	$117,964,800.00
20	1,099,511,627,776	0.15	15 PB	$1,536,000	$471,859,200.00
Total	**1,466,015,503,700**		**20,480 Terabytes or 20 Petabytes**	**$2,048,000**	**$629,145,600**

The cost of storing this much data on hard drives can be calculated based on a cost of about $100 per Terabyte drive, a price that does not include the housing or computer connection. To store the entire one trillion tiles would be about $2 million ($100 x 20,480 Terabytes). In order to achieve faster response times, there is some indication that Google uses faster random-access memory (RAM) to store the Google map tiles. If the entire map of the world were stored in RAM, it would cost the company more than $629 million.

These data storage requirements and costs are only for the map. The satellite view, with tiles in the JPEG format, requires approximately the same amount of storage space. Google maintains multiple data centers around the world and each would likely have a copy of the map and satellite image, and any other map that is provided – such as the Terrain view. Combining all of these data storage costs provides some indication of the importance placed on maps by Google and other companies.

Probably the major development in mapping during the first decade of the 21st century is the introduction of an online mapping tool in the form of the Application Programmer Interface (API). APIs are specialized libraries of computer code that are made available through the Internet. Soon after the introduction of Google Maps, the company made a library of routines available that would allow for the creation of custom online maps. Users could map their own points, lines and areas on a Google Map and make these maps freely available to others. The data used for mapping often came from other websites, thus the term "mashup" to indicate the melding of data and mapping tools to create new presentations of information.

An early application was the mapping of apartment listings from Craigslist, a free service for selling goods and services. New businesses were born simply by taking free data from one site and free mapping software from another and mashing them together. Websites like MapsKrieg and HousingMaps are examples of this type of combination.

IV. POLARIZING TRENDS

Stark contrasts in technology are inevitable in any immature technology but the differences are important to examine, both to understand the advantages and disadvantages of each approach as well as to envision future needs and developments. Contrasts are apparent in file formats, server/client relationships, Internet access, differences between users, and software development. How software is developed will likely have a major influence on the future of the Internet.

The polarization of software development is related to the distinction between open source and commercial software. While companies have been active in creating programs for the Internet, the Internet has cultivated an open source community of programmers around the world that contribute their time to the development of free software. Open source has vociferous private sector critics who suspect that any

software written by “idealistic nerds, and made available for free to anyone who wants to download it,” must be a plot against western business (The Economist 2007). Zealous believers in the open source movement, meanwhile, envision open source triumphing over the “evil empires” of commercial software. The clash is often depicted as an “epic struggle for supremacy between Linux and Microsoft's proprietary Windows operating system” (The Economist 2007).

The beginnings of the open source concept can be traced to the confrontation between Netscape and Microsoft in the mid-1990s. After Microsoft introduced its Explorer browser in 1996, the company made it very difficult for the Netscape browser to be installed under the Windows operating system. Eventually, Netscape, which had nearly 100% of the browser market share before 1996, rapidly lost its market share to Explorer and decided to release its source code as free software. The term “open source” was suggested to avoid the connotations associated with the word “free.” FOSS, or Free and Open Source Software, is the commonly used term.

Open source software is based on the work of small groups of people from around the world that work tirelessly together for little or no monetary benefit to create products that are ultimately provided for free to anyone who is able to download them. The installation procedure is usually not the easiest, the user interface is often not well-developed, and updating the software is somewhat problematic but the products are generally stable and work well once installed. A classic example in the mapping world is Minnesota MapServer, a product that has been implemented in all parts of the world. It implements a user interface that is now seen as clunky in comparison to Google, Yahoo and MSN map services.

In the last decade, most of the important online applications that have been introduced are open source – including the Apache web server software used by nearly 60% of all web servers (Netcraft 2007). Many online developers refuse to use any commercial applications – including commercial file formats like Flash SWF and Adobe PDF. The resistance to commercial software is particularly high in Europe and Canada. FOSS is not used as much in the less-developed world where the cost-saving aspect would be particularly beneficial.

V. FUTURE DEVELOPMENT

The Internet is still at an early stage of development. There will certainly be continual innovation that will likely influence maps and how we use them. There will also be the continuation of certain trends that we have seen in the past. The struggle between commercial and open source software will continue, with much of the innovation coming from the open source developments with the people involved receiving very little compensation for their efforts.

A critical and often over-looked area of online mapping is the speed of map delivery. This is one area where commercial interests have clearly surpassed non-commercial development. Even governments are not investing in the distributed data centers that increase the speed of access to the Internet. The sophisticated and highly-proprietary computer centers that are being built in many parts of the world increase the speed of access for Internet content, including maps. Governments and public agencies are not investing the equivalent of millions, perhaps billions, of US dollars to speed the delivery of maps through the Internet. As a result, these “slower” maps will have little interest to map users.

VI. REFERENCES

Cartwright, William, Michael Peterson, and Georg Gartner (2006) Multimedia Cartography. 2nd. Ed. Berlin: Springer.

Gartner, Georg, William Cartwright, and Michael Peterson. (2007) Location Based Services and Telecartography. Berlin: Springer.

Peterson, Michael P. (2007) International Perspectives on Maps and the Internet. Berlin: Springer.

Peterson, Michael P. (2003) Maps and the Internet. Amsterdam: Elsevier.

Putz, Steve (1994) Interactive Information Services Using World-Wide Web Hypertext. *Computer Networks and ISDN Systems* 27 (2) 273-280, Berlin: Elsevier Science BV.

Open Source Era. (2007) *The Economist.* (June 9, 2007).

Sample, John T. and Elias Ioup (2010) Tile-Based Geospatial Information Systems: Principles and Practices. New York: Springer.