Internet Structure and Development: On Strategic Uses of the Archetypes of the Networked Mind

Dr T. Matthew Ciolek,
Research School of Pacific and Asian Studies, Australian National University, Canberra ACT 0200, Australia

0. Abstract

The paper identifies some 25 assumptions Internauts make (sometimes explicitly, sometimes not) about the best ways to establish and manage online information resources. The paper postulates that the logic and behaviour of newly developed online information systems, in order to gain indispensable visibility and following on the Net, need to be congruent with those basic "archetypes" of the networked mind.

1. Introduction

In 1746 Frederick the Great wrote *Principes generaux de laguerre* to embody the experience he gained during two Silesian wars (Palmer 1943). The text was an attempt to distil his understanding of fast-paced war and of the ways a combating party should conduct itself. The document spelt out how to win not only individual battles and campaigns, but also how to secure, by the end of such campaigns, a sustainable geo-political advantage.

This paper is offered in the spirit of Frederick's memorandum. It is written by a man who has not led any armed troops. Instead, for the last seven and half years, he has immersed himself in the Internet. It aims to provide practical advice on the nature of the networked environment and on the ways of taking sustainable advantage of it.

In August 1991, the Internet comprised some 550,000 hosts and was used approximately by 5 mln people (Zakon 1988). Seven and half years later the Internet has grown into the global network of networks linking some 36,740,000 machines (in July 1998) and over 160 millions of people world-wide (Euro-Marketing 1998, Network Wizards 1999). Moreover, the Internet has become very rich as well. In December 1998, the Net comprised approximately 5,000 Gopher sites, 10,000 anonymous FTP sites, 30,000 USENET discussion groups, 39,000 IRC channels, over210,000 mailing lists, and 3.7 mln web servers with some 420 mln online documents (Kahle 1996, Bharat and Broder 1998, L-Soft 1999, Southwick 1998).

Previous advice on how to create and manage online information resources (in the form of

"The Seven Golden Rules of the Asian Studies WWW Virtual Library") did not say (Ciolek 1998a) why certain lines of electronic conduct should be favoured, whereas others should be shunned.

Therefore, a set of strategic-level, general-purpose notes is presented hereto my online colleagues. I think especially of those who work on such projects as the Electronic Buddhist Text Initiative (EBTI) (www.human.toyogakuen-u.ac.jp/~acmuller/ebti.htm), the H-Netelectronic forum for Asian History and Culture (H-ASIA@h-net.msu.edu), Scholars Engaged in Electronic Resources (SEER) (titus.uni- frankfurt.de/seer/index.htm), and the Electronic Cultural Atlas Initiative (ECAI) (http://www.ias.berkeley.edu/ecai).

When I think about the future I can sense plenty of online "battles" which await all of us. Some of the skirmishes will be about particular formats for our data, others will be fought to secure funds and necessary resources, still others will be about the quality of our products. The main battle, however, is inevitably that for global online acceptance of our work and a strong networked presence.

The starting assumptions of this paper are simple:

The Internet, as we know it, is made up of networked data and tools for their manipulation. Regardless of its specific application, each tool has a number of structural (i.e. technical), functional and, finally, social characteristics. If these characteristics are consonant with the prevailing (both tacit and overt) expectations and needs of its users, the tool gets accepted and proliferates. If not, it struggles briefly for public attention and then fades away. The most visible parts of today's Internet are, therefore, the successful products. Their very success forms the catalytic, self-referential milieu into which new products and ideas are constantly being born. The emergence and flow of networked inventions and activities inevitably establishes precedents, the elaborate patterns of relationships and expectations, which are then followed, more or less deliberately, by most of the subsequent inventions and events.

Uncovering such archetypal patterns (Jacobi 1971:38-39) is a rewarding exercise.

2. The Internet and its Characteristics

The Internet by the late 1990s has evolved into a complex environment. Originally a military communication's network it is now routinely used for five types of operations: (i) long-distance transactions (e.g. e-commerce, form-filling, remote work, entertainment); (ii) interpersonal communication; (iii) data storage; (iv) research (i.e. data finding); (v) remote

data access and downloading.

The Internet is a dynamic and mercurial system endowed with a number of traits. These are:

1. Technological neutrality. The Internet joins together computers of various sizes and architectures. They may run on various operating systems and utilise a great variety of communication links (Rutkowski1994). Although a great number of machines connected to the Internet are Unix boxes (see below), Internet technology does not require them to be so. The Internet, in a sense, is a giant multi-tasking environment, capable of accommodating a great range of participants and inputs.

2. Built-in piecemeal change and evolution. The Internet is not a one-off development. It is an energetic, polycentric, complex, growing, and self-refining system. It is a network which is geared to expansion and growth. It is a system which scales up extremely well. In December 1969 it comprised four interconnected machines. Exactly ten years later, in December 1979 it spanned 188 hosts. In October 1989 it linked 159,000 machines. In July 1998, it did the same for 36,739,000 machines (Zakon 1998). The Internet is, to borrow a phrase from Gabriel (1996:93) "a set of communicating components that work together to provide a comprehensive set of capabilities." The Net harnesses the energies of four distinct but mutually reinforcing and catalytic processes:

(a) Western culture's fascination (Tarnas 1996) with scientific research, technology, and information;

(b) The creativity of the Net's culture. This creativity is the necessary consequence (Rutkowski 1994) of a very large-scale, unregulated and polycentric archipelago of people and their datasets, machines for managing the data, and tools for instantaneous communication;

(c) The vanity of humans. Major technical discoveries and implementations are frequently carried out for no other reason than to give its inventors a spate of online fame (Dodd 1992, Rheingold 1994);

(d) Human greed. New technologies provide new financial vistas to entrepreneurs who put inventions to daily commercial use (Gabriel1996). Some new technologies reduce the cost of routine operations (i.e. to use an Internet telephone costs less than to use a traditional telephone network). Also, new technologies create totally new products, services and markets. The emergence of online advertising, network connectivity services, web hosting and

cyberstores illustrates this point (Reid 1997, Yesil 1997).

3. Robustness and reliability. All basic technical features of the Netsuch as the TCP/IP (transfer control protocol/internet protocol) (Kessler and Shepard 1997), the multiplicity of routes followed by the packet-switched data, and the sturdiness of related software are designed to eliminate errors, to handle unexpected interruptions and interferences, to advise users of encountered difficulties and to recover gracefully from any disasters and down-times.

4. Low cost. The Internet makes new uses of old technologies (standalone computers, operating systems, telecommunication networks). Whenever possible, Internet operations piggyback on already existing solutions. They rely on modularised, configurable, easy-to-replace, and easy-to-upgrade off-the-shelf software and hardware. The Internet not only creates new tools and resources, but also it recycles and rejuvenates several of the old ones (Levinson 1997).

5. Ubiquity. The robustness, modularisation and low cost of the system is coupled with the growing densities of dedicated computer lines, network backbones, as well as wired and wireless phonenet works. This means that Internet-enabled tools are deployed in ever growing numbers in an ever widening range of environments (Gilster 1997, Kelly and Reiss 1998). The history of the Internet is that of a long-term move from experimental laboratories to university computing centres, and then to offices and business, schools, homes, museums, taxis, supermarkets and, finally, street-based consoles, kiosks and hand-held devices.

Powerful and omnipresent is the Internet. However, it should not be regarded as an autonomous force or mysterious repository of knowledge. To do so would be an error akin to bowing reverentially towards as tack of old newspapers.

It always worth remembering that the Internet is simply a man-made infrastructure for handling the data supplied by the people themselves. In other words, all networked information resources are like biscuit-tins: what's in them is no more and no less than that which has been put into them by a concrete person. The type of digital materials placed on the Net, their trustworthiness and accuracy, the frequency of updates, the data formats and page layouts are always the outcome of human action. Such action may be thoughtful and deliberate, or spontaneous and irrational, but it always remains a prerogative of one, or a couple of individuals. The networked information tools are always a product of human work. Likewise, all the information galloping across the Net is a product of the daily habits, fashions, and politics of people who make daily use of those tools.

3. The Internet Tools and their Characteristics

The evolution of the Internet is punctuated by the introduction and mass acceptance of such key resources and tools as Unix, Email, Usenet newsgroups, Telnet, Listserv Mailing List Software, File Transfer Protocol, Internet Relay Chat, WAIS, Gopher, WWW, and more recently by the Altavista search engine, Java language, and finally, the Google search engine.

Table 1 The timeline of some major Internet and non-Internet e-publishing/e-communication tools

----

Date	Tool

1969	Unix operating system# (a)
1972	Email (b)
1977	UUCP Unix messaging and file-transfer tool# (c)
1978 Jan 16	Computer Bulletin Board System# (d)
1979	Usenet news groups# (b)
1980 Jun	Telnet (e)
1981	Listserv mailing list software (b)
1984	Unix OS supports Internet connectivity (t)
1985 Oct	File Transfer Protocol (FTP) (f)
1986	Hypercard (Macintosh) software# (w)
1987	NNTP (Network News Transfer Protocol) links Usenet and the Internet (x)
1988	Internet Relay Chat (IRC) (b)
1990	Archie FTP semi-crawler search engine (b)
1990 Dec	WWW server (prototype) (g)
1991 Apr	WAIS publisher-fed search engine + full text databases (h)
1991 Apr	Gopher (i)
1991 May 17	WWW server (production version) (g)
1992	Veronica crawler search engine (b)
1992 Jul	Lynx ascii WWW browser (j)
1993 Oct	Mosaic graphic WWW browser (g)
1993 fall	Jughead Gopher crawler search engine (b)
1994 Feb 14	Labyrinth graphic 3-D (vrml) WWW browser (k)
1994 Apr	Aliweb WWW semi-crawler search engine (l)
1994 Oct 13	Netscape WWW browser (m)
1995 Apr	RealAudio narrowcasting (n)
1995 May 23	Java programming language (o)

1995 Dec	Altavista WWW crawler search engine (p)
1995 Jun	Metacrawler WWW meta-search engine (q)
1996 Apr	Alexa WWW intelligent navigation adviser (u)
1996 Jun	Internet Archive full text database (r)
1998 Apr	Google WWW crawler intelligent search engine (s)

Note: This table is based on Table 1 in Ciolek 1998b and supplementary data from Ciolek 1999b. Non-Internet technologies are marked with "#."

Sources:
(a) Hauben and Hauben (1995); (b) Zakon (1998); (c) Rheingold (1994:116); (d) Rheingold (1994:133); (e) Postel (1980); (f) Barnes (1997); (g) Cailliau (1995); (h) St. Pierre (1994); (i) La Tour (1995); (j) Grobe (1997); (k) Reid (1997:175); (l) Koster (1994); (m) Reid (1997:33); (n) Reid (1997:69); (o) (Harold 1997); (p) Compaq (1998); (q) Selberg (1997); (r) Kahle (1996); (s) (Google 1998); (t) Severance (nd); (u) Kahle & Gilliat (1996); (w) Goodman (1987); (x) Laursen (1997)

The review of major aspects of these tools will be carried out briefly. Basic information on the history as well as the scale of their contribution to the growth of the Net has already been provided elsewhere (Ciolek 1998b).

UNIX

The foundations of an operating system called Unix were laid at AT&T Bell Laboratories in 1969. On the 1st September the same year ARPANET, the pilot version of the future Internet, was first used to connect two computers at UCLA, and UCSB (Laursen 1997). Since the very outset Unix has become increasingly popular, first in the academic world, and subsequently as a de facto non-proprietary standard OS for tens of thousands of multi-user workstations and microcomputers. Technical details of the system can be found in a number of publications, including Frisch (1991) and Todino and Strang (1989). The system's early history is described in Hauben and Hauben (1995).

Unix is not a product of Internet culture. It is its catalyst and cornerstone. Internet culture owes Unix a major debt in the four areas. These conceptual and procedural debts are: multitasking, community fostering, openness and extensibility, and public access to the source code. Let's briefly look at each of these debts.

Unix was one of the first operating systems which embodied the principle of multitasking

(time-sharing). In most general terms it means that several users could simultaneously operate within a single environment and that the system as a whole coped well with this complicated situation. Unix was the first operating system which demonstrated in practical terms robustness and tolerance for the variety of it's users simultaneous activities.

The phenomenon of multitasking also had another important consequence. It facilitated the emergence of a self-aware community of computer users. People no longer competed with one other for precious time on the system. Their work was no longer handled in a sequence of discrete batch operations. With the advent of Unix, people making use of the operating system started sharing the same activity space. They could browse through all parts of the machine's file structure and they could invoke all available commands. They were simultaneously affected by the strengths and limitations of Unix. Unix users thus constituted a group which had a reason to argue and lobby for the further growth and extension of their jointly used work platform.

The evolving sense of the community received a further boost in 1977 (Rheingold 1994:116) with another of AT&T's inventions, the Unix-to-Unix-Copy (UUCP) utility. This software was made available world-wide along with new versions of the operating system. UUCP made possible for any computer running Unix to automatically connect via modem with any other computer using Unix and to ship messages and document files from one machine to another. This means that from 1977onwards the Unix users were encouraged by the UUCP software to start forming professional contacts with other Unix users, regardless where they might reside. The UUCP was the stimulus which led to the establishment of Usenet (see below).

Another influential characteristic of Unix is the public-domain status of its source code. The code was made available by its AT&T creators to anyone, anywhere, practically free of charge. This was a key strategic decision. It meant that Unix's universal availability broke a spell which hitherto kept innovation proprietary and chained to the company coffers. The free status of the basic Unix software (but not of the specialist data which could be managed by that system, an important distinction) meant that consecutive refinements could now be easily embarked upon (Xenix, Berkeley Unix (BSD), SunOS, System V) and improvements could flow spontaneously from one laboratory to another. This revolutionary approach was soon adopted by a number of other software developers. Its strategic value for capturing the major share of the installed base, hence of the user base, and hence of the market for related products and data, was obvious. This new strategy has been vindicated several times, most recently in 1998 by decisions to publicise source codes for the Linux operating system (the latest incarnation of Unix), Netscape's WWW Navigator browser, and the Sun Microsystem's

Java language (Effinger and Mangalidan 1998).

The fourth key feature of Unix is its structural openness and amenability to piecemeal improvement. Unix was deliberately designed to foster "a professional community of programmers who used the Unix toolbox to create new tools that all the other Unix toolbuilders could use" (Rheingold 1994:117-118). From the very outset anybody could contribute a variation to any of the Unix component software modules. In this way several tens of hundreds of UNIX modifications were implemented all over the world. In this way the University of Berkeley version of Unix (version 4.2BSD) was augmented in 1984 so that it could handle the TCP/IP protocol suite, the language of all Internet operations. The original networking support included remote login (Telnet), file transfer, and electronic mail (Rheingold 1994:83, Severance nd).

The incremental modifications to the software meant, on one hand, a state of anarchy and confusion. On the other, it was the beginning of a culture of individualistic creativity. A Unix programmer with an idea and skill could now try out a gamut of technical solutions, without asking anyone's permission. If she failed in her projects, she would do so in the privacy of her local system. Yet, if she succeeded, she could report the new invention to all who cared to hear about it. The free distribution of Unix source code meant the onset of a culture of public success (naturally, only the effective patches and modules would be announced) and that of parallel private tinkering, blundering and messing up. This was a liberating development. All sorts of generally useful utilities and applications were now written within the larger constraints of the Unix framework. Unix, therefore, is an early champion of the principles of natural selection within the competitive/cooperative world of computer software.

Email

Electronic mail was first introduced in 1972. An extensive analysis of email technology, and its social and political applications is provided in Anderson et al. (1995).

Email is the first of the Internet's tools dedicated to the provision of fast, simple and global communication between people. This revolutionary client/server software implied for the first time that individuals (both as persons and roles) could have their unique electronic addresses. Within this framework messages were now able to chase their individual recipients anywhere in the world. The recipients, in turn, were brought in close contact with each other and could form one-to-one communication links and friendships, independently of the official relations between their respective employers. This was a momentous development, for frequent and intensive communication forms social groups, and the groups, in turn, form an environment

in which innovative products are considered and created.

The initial format of email communication was that of a one-to-one exchange of electronic messages. This simple function was subsequently augmented by email's ability to handle various attachments, such as documents with complex formatting, numbers and graphic files. Later, with the use of multi-recipient mailing lists (a development which prompted the development of the Listserv software, see below) electronic mail could be used for simple multicasting of messages in the form of one-to-many transmissions.

Finally, email is important because it has disseminated and popularised an awareness that each networked computer has, in fact, a unique address, a world-wide recognisable digital identity. This identity was constructed according to a set of simple, and generally accepted rules. Hence the hitherto amorphous and anonymous mass of digital devices became a consciously recognisable and visible lattice divided into international and country-specific domains. Moreover, each domain was further sub-divided into subsets of specialised networks, which in turn were comprised of numerous lower-level networks of computers associated with the activities of such entities as hospitals, research institutes, media, business corporations, universities, and so forth.

In short, email is important not only as a tool for interpersonal communication, but also as a tool which stressed the notion of unambiguous and explicit targeting between entities involved in online transactions. This notion received a further boost in the form of hypertext linkages (see section on WWW systems below).

Usenet Newsgroups

Usenet (Unix Users Network), the wide-area array of sites collating and swapping UUCP-based messages was pioneered in 1979. Usenet was originally conceived as a surrogate for the Internet (then called ARPANET). It was to be used by people who did not have ready access to the TCP/IP protocol and yet wanted to discuss their various Unix tools. It was only in 1987 that the NNTP (Network News Transfer Protocol) was established in order to enable Usenet to be carried on the Internet (i.e. TCP/IP) networks (Laursen 1997). The history of Usenet is told in Hauben and Hauben (1995), Anonymous (nd), Bumgarner (1995), Schnierle (1995), and Rheingold 1994:110-131). The Usenet was influenced by, and in turn has influenced, other pre-internet technologies such as FidoNet and the dial-up, computerised bulletin board systems (BBS) (Dewey 1987, Allen 1988, Dodd 1992, Rheingold 1994: 131-144).

Usenet newsgroups taught the Internet three major lessons.

Firstly, the newsgroups proved the practical usefulness of distributed systems for the production of large volumes of online information. In1992, when the WWW comprised no more than 26 servers and a few hundred documents (Ciolek 1998b) Usenet already had 4,300 groups exchanging some 17.5 thousands messages a day from 63,000participating sites (Zakon 1998). Usenet was the first large-scale system where information could be created locally, by anybody who had the freely available client software and an interest in a given topic.

Secondly, like email, IRC groups and Listserv electronic agoras, the Usenet proved itself a great tool for promoting the growth of online communities. Again, the emerging pattern was clear: a valid topic prompts ample electronic communication about it. Spontaneous electronic communication leads to the formation of an invisible college of people with a vested interest in the issue, as well as an ingrained interest in informing and impressing one another.

Thirdly, Usenet also demonstrated, like the 100,000 conferences of the BBSers (Allen 1998) before it, that once a certain level of user participation is reached, public messaging systems require for their very survival some form of moderation of transactions. However, this was difficult to achieve since the Usenet, originally a professionals' forum for Unix troubleshooting, was intentionally designed as an "anarchic, unkillable, censorship-resistant" (Rheingold 1994:118) electronic meeting place for millions of people in dozens of countries. In terms of its capacity to cope with the flood of data the Usenet has gone through a series of crises and restructurings (Bumgarner 1995). It also had several upgrades made to the logic of its operations and to the networking technology. However, it continues to suffer from inability to handle adequately the uneven content of swapped news (all sorts of messages including drivel, flame-wars, spoofs, and deliberate spams are regularly posted on Usenet). This failure suggests that larger (say 20+) and quasi-anonymous online groups do not rely on commonsense and are not subject to self-regulatory processes.

This means that in order to be viable and productive, online resources require filtering of all publicly generated information before such information is fed-back into the communication loop.

Telnet

The networking tool called Telnet was invented in 1980 (Postel 1980). It allowed people (with adequate access rights) to login remotely into any networked computer in the world and

to employ the usual gamut of computer commands. Thereby files and directories could be established, renamed and deleted; electronic mail read and dispatched; Usenet flame wars indulged in; and statistical packages run against numeric data- all at a distance. Moreover, results of all these and other operations could be remotely directed to a printer or via FTP (see below) to another networked computer. In short, Telnet gave us the ability to engage in long distance man-machine transactions, that is, ability to do the work as telecommuters.

Listserv Mailing List Software

Listserv technology, which was subsequently followed by the introduction of such software as Majordomo and Listproc, appeared on the Internet scene in 1981 (Liu, C. et al. 1994). The Listserv software, based on a client/server principle, automated the most cumbersome tasks of the one-to-many email communication.

Experiences with the use of mailing lists (in late 1998 there were approximately 210,000 email-based communication loops) (Southwick 1998) confirm a lesson learned already from Usenet and BBS operations. That lesson is simple. People who receive regular feedback from each other tend to form lasting communities. Communities foster quicker development of products, in the form of data and tools. Tangible products further stimulate the flow of ideas and commentaries. However, information needs to be closely adjudicated and edited if a computer-mediated communication system is to function properly.

File Transfer Protocol

The FTP client/server technology was first introduced in 1985 (Barnes1997). Its usefulness to Internet culture is three-fold.

Firstly, the FTP was a first widely-accepted tool for systematic permanent storage and world-wide transmission of substantial electronic information (e.g. programs, text files, image files). Secondly, FTP archives promoted the use of anonymous login (i.e. limited public access) techniques as a way of coping with the mounting general requests for access to the archived information. That novel technique placed electronic visitors in a strictly circumscribed work environment. There they could browse through data subdirectories, copy relevant files, as well as deposit (within the context of a dedicated area) new digital material. However, the FTP software would not let them wander across other parts of the host, nor did the visitors have the right to change any component part of the accessed electronic archive.

Thirdly, the rapid proliferation in the number of public access FTP archives all over the world

necessitated techniques for keeping an authoritative, up-to-date catalogue of their contents. This was accomplished through the Archie database (Deutsch et al. 1995) and its many mirrors. Archie used an automated process which periodically scanned the entire contents of all known "anonymous FTP" sites and report findings back to its central database. This approach, albeit encumbered by the need to give explicit instructions as to which of the FTP systems need to be monitored, nevertheless integrated a motley collection of online resources into a single, cohesive, distributed information system.

Internet Relay Chat

Messages which are typed into a networked computer can be displayed, at the exact moment of their formulation, on screens of other interconnected machines. In 1988 this observation resulted in the development of the Internet Relay Chat (Southwick 1997). The IRC technology established electronic chat rooms or channels where strangers could congregate for the purpose of real-time, synchronous exchanges. There, typed messages forma series of interleaved dialogues between anonymous interlocutors with carefully managed electronic *personae*. Their exchanges are produced on screens in the course of one-to-one as well as many-to-many "chats".

The popularity of the IRC systems confirms the Internauts' hunger for interpersonal communication where each party has complete control over details of their real-life situation and identities. Also, the IRC confirms a principle already encountered in the context of Listservs, Usenet and FTP systems - anonymous Internauts if they are given a chance, will vandalise the very resource whose integrity attracted them to visit in the first place.

WAIS Search Engine

The Wide Area Information Server, with its name commonly abbreviated to WAIS, was introduced online in 1991 (St. Pierre 1994). As the tool for electronic publishing WAIS offered several novel features.

Firstly, it made intensive and sustained use of the already mentioned concept of distributed, that is locally published of data. Secondly, like FTP Archie before it, it made use of a central register of contents for those distributed data sets. Thirdly, WAIS was the first widely accepted information resource which made deliberate and explicit use of the meta-data. These were machine-readable electronic notes with summaries of the each of the published digital documents. Such meta-datahad to be manually provided by the publishers of online databases. Also, these identifiers had to be explicitly supplied (via email) to the WAIS central register of resources. In practice this meant that a number of WAIS databases might spring into

existence in various parts of the globe, without the WAIS headquarters (and thus the rest of the Internet community) having any knowledge of their whereabouts and contents.

Due to innovative programming, WAIS client server (unlike Usenet, BBS and FTP before it) enabled users to quickly locate and display on their PC screens any piece of required information, regardless of how big or how small it was. Meaningful and pertinent data could be now promptly retrieved regardless of where exactly they were stored. This was possible as long as the local WAIS database was up and running, and as long as the central register was kept up-to-date. In other words, WAIS was the first working example of global data find ability and of their transparent and hassle-free accessibility.

Gopher

Gopher client/server software was used for the first time in 1991 (La Tour nd; Liu, C. et al. 1994). It was a ground-breaking development on two accounts. Firstly, it acted as a predictable, unified environment for handling an array of other electronic tools, such as Telnet, FTP and WAIS. Secondly, Gopher acted as an electronic glue which seamlessly linked together archipelagos of information tracked by and referenced by other gopher systems. In short, Gopher was the first ever tool capable of the creation and mapping of a rich, large-scale, and infinitely extendable information space.

World Wide Web Server

The first prototype of the WWW server was built in 1991 (Cailliau 1995, Berners-Lee, nd; Berners-Lee 1998). The WWW server is an invention which has redefined the way the Internet is visualised by its users.

Firstly, the WWW server introduced to the Internet the powerful point-and-click hypertext capabilities. The hypertext notions of a home page and links spanning the entire body of data was first successfully employed on a small, standalone scale in 1986 in the Macintosh software called Hypercard (Goodman 1987). The WWW however, was the first hypertext technology applied to distributed online information. This invention was previously theoretically anticipated by a number of writers, including in the 1945 by Vannevar Bush of the *Memex* fame, and again in the 1965 by Theodor Nelson who embarked on the never-completed *Project Xanadu* (Nielsen 1995, Gilster 1997:267). Hypertext itself is not an new idea. It is already implicitly present (albeit in an imperfect because a paper-based form) in the first alphabetically ordered dictionaries such as *Grand dictionnaire historique*, compiled in

1674 by Louis Moreriego; or John Harris' *Lexicon Technicum* which was published in 1704 (PWN 1964). It is also evident in the apparatus, such as footnotes, commentaries, appendices and references, of a 19th century scholarly monograph.

The hypertext principle as employed by the WWW server meant that any part of any text (and subsequently, image) document could act as a portal leading directly to any other nominated segment of any other document anywhere in the world.

Secondly, the WWW server introduced an explicit address for subsets of information. Common and simple addressing methodology (Universal Resource Locater [URL] scheme) enabled users to uniquely identify AND access any piece of networked information anywhere in the document, or anywhere on one's computer, or - with the same ease - anywhere in the world.

Thirdly, the WWW provided a common, simple, effective and extendable language for document markup. The HTML language could be used in three different yet complementary ways: (a) as a tool for establishing the logical structure of a document (e.g. Introduction, Chapter 1, Conclusions, References); (b) as a tool for shaping the size, appearance and layout of lines of text on the page; (c) as a tool for building the internal (i.e. within the same document) and external (to a different document residing on the same or totally different server) hypertext connections.

The interlocking features of the hypertext, URLs and the markup language, have laid foundations for today's global, blindingly fast and infinitely complex cyberspace.

Moreover, the World Wide Web, like gopher before it, was also a powerful electronic glue which smoothly integrated not only most of the existing Internet tools (Email, Usenet, Telnet, Listservs FTP, IRC, and Gopher (but, surprisingly, not WAIS), but also the whole body of online information which could accessed by all those tools.

However, the revolutionary strengths of the Web have not been immediately obvious to the most of the Internet community, who initially regarded the WWW as a mere (and possibly clumsy) variant of the then popular Gopher technology. This situation has changed only with the introduction of PC-based Web browsers with user-friendly, graphics-interfaces.

World Wide Web Browsers

The principle of a client/server division of labour was put to work yet again in the form of a

series of WWW browsers such as Mosaic (built in 1993), Lynx (which is an ASCII, Telnet-based client software), Erwise, Viola, Cello, as well as, since 1994, several editions of Netscape and Explorer (Wolf 1994, Reid 1997:33-68). Each of the Web browsers, except for Lynx, which constitutes a deliberately simplified and thus very fast software (Grobe 1997), provided Internauts with series of novel capabilities.

These are: (a) an ability to handle multi-format, or multimedia (numbers, text, images, animations, video, sound) data within the framework of a single online document; (b) the ability to configure and modify the appearance of received information in a manner which best suits the preferences of the reader; (c) the ability to use the browser as a WYSIWYG ("what you see is what you get") tool for crafting and proofreading of the locally created HTML pages on a user's PC; (d) ability to acquire, save and display the full HTML source code for any and all of the published web documents.

The fact that one could simply copy and modify somebody else's promising HTML design, and incorporate it into one's own WWW-styled information system, and then to have the whole thing checked and double-checked through a Web-browser running in a local mode set-off a two-pronged explosion. Firstly, the volume of the Web-based information started growing exponentially, from a base of couple of tens of WWW pages in early 1991 (Ciolek 1998b) to approximately 420 mln pages in early 1999. Secondly, the great habitability (Gabriel 1996) of Web-based information resources meant that all of a sudden the Internet transformed itself from an elitist domain for sporadic email-mediated interpersonal contacts into a popular domain for continuous and large scale multimedia information storage and distribution.

WWW-Crawling Search Engines

The first four years (1991-1994) of the Web existed largely without a proper automated cataloguing systems (Koster 1994). Basic navigational services were instead provided by manually compiled distributed indices such as the WWW Virtual Library (Secret 1996) and the Yahoo-style, user-fed centralised databases.

The first software agents which started 'crawling' the web along the multitude of its hypertext paths, collecting data on the encountered web documents and reporting them back to a central database were introduced in 1995 (Selberg 1997; Lawrence and Giles 1998). In the late 1990s there were several hundreds of such systems, with the most prominent role being played by the simple (first-generation) search engines such Altavista, HotBot, Lycos, Infoseek, Excite and the Northern Lights, as well as the host of meta-databases living off the data collected by

the first-generation systems (Ciolek 1999a).

The Web crawling databases of Internet links signal a number of new and important developments.

The first of them is the introduction of pro-active data acquisition and cataloguing. A typical Web-search engine, unlike a static Yahoo catalogue, does not wait to be briefed or updated by a cooperative user. Instead, it takes, so to speak 'the reins of the Internet into its hands' and acts as its own supplier and evaluator of data.

The second development is the emergence of very large, but, nevertheless speedy and robust information service. For instance, in May 1998, Altavista (Compaq 1998) kept track of some 140 mln web documents, or 30-50% of the entire cyberspace (Bharat and Broder 1998). Major search engines such as Altavista or Infoseek are capable of handling several millions of accesses and data-queries a day. The exact nature of queries (a keyword-, a string-, and a boolean-query) and the subsection of the cyberspace to be searched (data from a particular server, or data from a particular domain, or those in a particular language) can be precisely tailored to the needs of the user.

There is also a third development. The overall content of the located documents can be estimated on the basis of the meta-data generated by the search engine itself. The reliance on the good will and skill of the data-producers, which was characteristic of the FTP and WAIS approaches, is thus eliminated. Admittedly, the machine generated meta-data are at present incomplete and rudimentary. They consist of details of the document's language and size, its title, the first 30-50 words of its content, the URL (which is quite effective for establishing particulars of the e-publisher) as well as the document's relevancy rank relative to the employed search terminology.

The fourth feature offered by some major search engines (e.g. Altavista) is their ability to provide free and real-time translations of the texts from one natural language to another.

The fifth capability is the incipient online search for non-text information. For instance, Altavista can scout subsections of the Web for a piece of static graphics, or even (since September 1998 and the Clinton/Lewinsky affair being documented on the Internet) for a keyword-referenced section of a full length online video.

Finally, web-crawling search engines are very effective in popularising the idea of virtual web-pages. The new technology enables the information to be deposited on a hard disk in a

simplified and generic form. This generic information is subsequently used for creation of more complex, one-time-only, on-the-fly documents. This means that detailed online information, together with all corresponding layouts and structures, is assembled each time afresh. It does not need to be treated as a snap-frozen whole any longer. On the contrary, it may consist of dozens of individually packaged and individually updated info-nuggets. These kernels of information are put together only on demand, so that they generate a synthetic document. This document is shipped according to requirements of a particular reader. Moreover, such information is highly configurable. It can be customised according to reader's identity, address, interests and other situational criteria.

Java

Java programming language, created by Sun Microsystems, gained visibility in 1995. It is a granular, general-purpose, multi-platform language (Harold 1997, Kelly and Reiss 1998, Sun Microsystems 1998). It is used for construction of tiny utilities which become operational for the duration of one's online transaction with a given information system. Such small-scale applets and beans (i.e. applications) are speedily downloaded into one's PC memory. They used as temporary Web client-enhancing devices for a wide range of supplementary processing. Therefore, Java provides specialist, disposable tools for manipulation, analysis and display of generic data obtained from Web and other servers. This on-the-fly supplementary processing extends the functionalities of an online resource without encumbering the data server itself or the network bandwidth. Once such local-mode data transformations and analyses are completed their results can be saved to a hard disk. The key aspect of Java technology is that it can be used without exposing the PC to the corrosive impact of a malfunctioning or malign applet. In other words, Java provides all the advantages of a sophisticated computing environment without the usual bad karma.

Google Database

The representative of the newest, third generation of WWW-search engines, called Google, was first launched in April 1998 (Google Inc. 1998). It is the result of a long-term research project conducted at Stanford University. The project is aimed at making information served by the existing WWW search engines more pertinent and more useful. The earlier generation of search engines (see above), however quick and powerful they were, did not cope well with the twin problems of the ongoing explosive growth of the Web and the rapid turnover in the content of networked documents (Frauenfelder 1998; Lawrence &Giles 1998; Notess 1998).

Despite the ability to track and catalogue of up to several million documents a day, the

traditional search engines had to, *volens nolens*, aim either at the completeness of their coverage, or the freshness of the gleaned data. Moreover, the sheer size of the Web-based information and the speed with which new materials were put into circulation made these databases increasingly unwieldy devices. In a situation where a single question such as "Asian Studies" can generate (in mid December 1998) up to 20,300 of possibly relevant answers, even the most comprehensive and most up to date register of links ceases to be a useful resource. Increasingly often online research would become a two-step procedure. Firstly, within a second or two the contacted database would dump on hundreds of leads to materials containing the matching keywords onto the PC's screen. Secondly, an investigator would spend endless minutes trying to filter out the irrelevant information, zoom on and finally check out the most promising links. In short, the crawler's strategy of brute-force does not scale-up well in the world where the volume of available information grows at an exponential rate.

This embarrassment of riches has been skilfully avoided by the Google search engine. Instead of treating the web's cyberspace as a homogeneous mass of online pages, Google treats it as anarchipelago of interlinked communities of documents. Each such group deals with a specific topic or a theme. Naturally, topical communities of pages are formed by the Internauts themselves, as they establish and cross-link their web documents. There are in existence many hundreds and thousands of such clusters. Some of them are distinct from each other, others may partially overlap. However, they all share a common pattern: a handful of high-quality documents are inevitably cross-referenced and linked-to by other sites. The high-quality sites are those which are regarded by the rest of the Web as the major clearing houses for a given subject matter.

A site which attracts a large number of hypertext links inevitably functions as an online authority which makes and unmakes the reputation of related sites. The most important resource for a given area of specialisation is the one which gains the greatest attention, in the form of web links, of other important (i.e. heavily linked-to) resources. The Google's algorithm for coping with the Web's size and complexity is, therefore, simple. First, the database actively collects online intelligence. Next, it subjects the data to an iterative mathematical analysis. This analysis quickly and almost always reliably uncovers which materials, according to their informed peers, are the best source on a given topic.

Google's arrival on the Internet scene is significant for two reasons. Firstly, it points out that while small volumes of information can be handled manually reasonably well, and the intermediate volumes need to go through an automated keyword search; the large volumes of data and meta-data need the assistance of intelligent and scalable search programs.

Secondly, the operation of the Google search engine suggests that even the uncoordinated, anarchic and free-wheeling environment of the cyberspace is in fact a self-organising environment. This unregulated and unmanageable system seems to be, surprisingly, hierarchically ordered on the basis of merit and hard work of its authors. However, an Internaut's greatest asset, her/his online presence and visibility as well as the corresponding stature and prestige, is earned slowly. The visibility and prestige arise only if a site wins the confidence and approval of its peers. If that happens, the quality of information, and the quality of its online organisation are - in the long term - recognisable (via ample hypertext linkage) even by the dumbest member of the milling online crowds.

Google analyses indicate that online links and cross-references act as an online currency with which debts of gratitude are settled. Moreover, these links act as electronic citations and recommendations. Therefore, one may conclude, that the whole structure appears to observe a simplified version of the peer review principle, the very one which forms the methodological foundations of modern science (Popper 1969, Tarnas 1996).

4. The Archetypes of the Networked Mind

This paper gradually approaches its conclusion. One way of looking a tall the discussed technical solutions is to view them as a series of innovations which enable and speed up our interchanges across the Net. However, a more fruitful approach would be to view them as a series of archetypes of successful online conduct. Each of the reviewed online tools represents a body of practical wisdom. Each of them tells us about their relationship to other tools and happenings on the Internet. Each of them offers an example, a set of principles, are commended strategy for the online existence.

These principles, or archetypes, form three broad clusters.

Firstly, there are *social* archetypes. A tool for the work on the Internet, or a data set thrives on visibility. In the case of software, it is the size of the user base which matters. In the case of information, it is the number of permanent electronic links, or connections, which are made to such data. Therefore, in the light of our review of the Internet tools, it can be postulated that successful networked resources are those which are:

(1) Ubiquitous. They are readily available, available in sufficiently large numbers to meet demand, and available from a variety geographic environments and online contexts;

(2) Freely accessible and inexpensive to use. Profit, or cost recovery, if any, is to be made

from provision of the supplementary tools, functionalities and services which arise in response to the mass demand for the free-of-charge primary ones;

(3) Have a clear electronic identity. They have short, simple and easy to remember (and if possible, eponymous) electronic address;

(4) Foster the establishment and growth of a community of users;

(5) Have visible and frequent bi-lateral contacts with that community. These contacts are open to the members of the public. However, all transactions are moderated;

(6) Have a clearly defined team of maintainers who receive, process and acknowledge inputs from users;

(7) Are easily copyable, replicable and modifiable by their users. Users are regarded as partners vitally interested in the resource's success.

Secondly, there is a cluster of *structural* archetypes:

(8) Technological neutrality. The product runs on any platform (including old ones with limited memories and speeds) and any operating system, but above all on the most popular ones;

(9) The client/server principle of division of labour;

(10) The served data is fully configurable, customisable at both the publishing and consumption ends. The user's view of the data can be made as simplistic and as detailed as required;

(11) The server supports the presence of virtual documents, created on-the-fly in response to a particular query.

(12) The client supports the use of additional tools for the purpose of local data manipulation and analysis;

(13) Robustness, sturdiness and reliability. If the product collapses, it does so gracefully, so that the afflicted user can recover from the disaster with minimum fuss and damage to his work;

(14) Endless improvement. The information resource can grow and evolve. All improvements can be carried out in an ad-hoc, incremental fashion;

(15) Meta-data are generated through an automatic process. Whenever possible meta-data are extracted from the resources themselves and from the patterns of its use. The manual input of metadata information should be possible but not indispensable;

(16) Authoritative and frequently updated catalogue of resources forms the hub for a an archipelago of distributed data sets;

(17) Scalability. The system is able to cope with its own growth and popularity;

Finally, there are a number of *functional* archetypes:

(18) The system can handle large numbers concurrent users;

(19) Each piece of information has a unique address. This address helps to identify it in the universe of other related and unrelated pieces of information, to repair it and to make a hypertext link to it;

(20) Each piece of information can be easily marked up to represent its internal structure as well as semantic content;

(21) The system handles all varieties of information (numbers, text, vectors, graphics, sound, moving images, virtual realities and simulations);

(22) All data sets have secure, backed-up and permanent online storage;

(23) All data sets are quickly and globally findable;

(24) All data sets can be accessed in a simple, intuitive and 'transparent' manner;

(25) All data sets can be transposed in real time, from one format to another. Texts can be automatically translated from one language to another;

Clearly, not all of these 25 social, structural and functional archetypes need to be simultaneously present in each of our online projects. However, the common sense suggests that since each of these patterns is now a part of the Internet's tradition, we would do well to

pay heed to their existence. How we do that depends on assumptions we make about the actual implementation of our electronic work.

5. Establishing the User Base: Two Schools of Thought

Richard Gabriel in his seminal analysis of the ways software applications flourish and fail suggests the existence of two basic strategies (1996: 215-229).

The first one, which he calls "Worse is Better", is founded on the idea that the main objective of the developers' team is the early capture of the large user-base. For this to happen they have to move quickly. They have to deliver results of their work as early as possible; capture the attention as well as loyalty of the potential users; and thus pre-empt and block any possible competitors.

The second strategy is called "The Right Thing". This strategy is founded on the idea that the main objective of the developers' team is the construction of a superior product, something that then can be truly proud of. If the product is well designed and provides a good service, people sooner or later will learn about it, start using it and, eventually, start loving it too. For this to happen, the developers must design their product carefully; they should make it available to the online community only when all the work is fully completed, and release it to the accompaniment of testimonials from experts.

Products developed according to "Worse is Better" philosophy have, accordingly, the following characteristics (listed in order of importance):

(a) The design is simple in implementation. The implementation is fast and lean, so it can run on the smallest of computers. The product is fully interoperable with the programs and tools that the expected users are already using;

(b) The product is bug-free even if this requirement means implementing fewer functionalities;

(c) The interface is adequate yet simple;

(d) The offered functions cover only the essential requirements. The completeness of the product is less important than the speed and simplicity of its implementation;

(e) The design is internally consistent as far as possible, even if this requirement means fewer

functionalities and rudimentary interface.

The key assumptions of this "rapid-fire", "populate or perish", and the "learn-from-correctable-mistakes" paradigm are: if the product which has some substantial value has been delivered early, while the need for it is acute and unsatisfied by other products, it will be noticed, and heavily used and will tend to spread like a virus, "from one user to another, by providing a tangible value .. with the minimal acceptance cost". Once such a readily available though still imperfect product becomes popular, "there will be pressure to improve it, and over time it will acquire the quality and feature-richness" characteristic of the quality-orientated "right-thing" approach (Gabriel 1996:220).

In contrast, products informed by "The Right Thing" design philosophy display the following characteristics:

(a) The interface is simple;

(b) Simplicity of implementation is irrelevant;

(c) The design provides as many important functionalities as possible;

(d) The product is 100% bug-free, even if it means delays in making it available to the community of possible users;

(e) The design is 100% consistent, even if it means that interface and implementation are more complicated.

Clearly, the "Right Thing" approach has evolved during the last 500 years of book production and dissemination. It is an approach which makes sense only when the publication of a piece of work is so a monumental and costly an event that it has to be deferred until all intricate aspects of content and format are fixed and double-checked. It is an approach shaped by the abhorrence of shortcomings. These shortcomings, if ever allowed, in the world of printed publications were almost impossible to repair.

Here the key assumption is that the world is populated by knowledgeable, discriminating and rational people, who will not fail to notice the arrival of the truly good, nay, advanced product and who will, moreover, change the existing working routines and habits, abandon any previous investment (both monetary and emotional) in systems they depended on so far, and will switch to "The Right Thing" software or data (or metadata) format.

However, there is a problem with such patient, sensible, and "one-shot-only" philosophy. As the sad outcome of the competition between the superb Mac (Apple) and mediocre Windows (Microsoft) operating systems eloquently testify, "The Right Thing" strategy does not happen to work for the realm of standalone computers.

It will not work for the Internet, either. As our review of the networking tools has demonstrated, the Internet is not about superior technology. It's about superior relationships.

6. About the Author

Dr T. Matthew Ciolek, a social scientist, heads the Internet Publications Bureau, Research School of Pacific and Asian Studies, The Australian National University, Canberra, Australia. Since December 1991 he has been responsible for making the RSPAS' electronic research materials available to the Internet community via FTP-, WAIS-, Gopher-, Web- and email-based technologies and is one of the world's pioneers in electronic communication regarding the Asia-Pacific region. Since June 1994 he has been a designer and editor of an electronic journal "Asian Studies WWW Monitor"(coombs.anu.edu.au/asia-www-monitor.html) and a number of online guides to the Internet, including the influential Asian Studies WWW Virtual Library (coombs.anu.edu.au/WWWVL-AsianStudies.html). He also serves as a co-editor of the H-ASIA@h-net.msu.edu electronic forum, and as a member of the Steering Committee of the Electronic Cultural Atlas Initiative (ECAI) (www.ias.berkeley.edu/ecai), University of California, Berkeley, USA. His work and contact details can be found online at http://www.ciolek.com/PEOPLE/ciolek-tm.html

7. Acknowledgments

I am grateful to Monika Ciolek and Abby Zito for their critical comments on the earlier version of this essay.

8. References

[The great volatility of online information means that some of the URLs listed below may change by the time this article is printed. The date in round brackets indicates the version of the document in question. For current pointers please consult the online copy of this paper at http://www.ciolek.com/PAPERS/pnc-taipei-99.html

- Allen, Thomas B. 1988. Bulletin Boards of the 21st Century are Coming of Age. Smithsonian, September 1988. Vol 19(6), pp. 83-93.

- Anderson, Robert H. et al. 1995. <u>Universal Access to E-mail Feasibility and Societal Implications. RAND Corporation.</u> (v. 1995).
  www.rand.org:80/publications/MR/MR650/

- Anonymous. nd. <u>Usenet: History</u>. (v. Sep 1995).
  www.vrx.net/usenet/history/

- Barnes, Jesse. 1997. <u>History and Overview of FTP</u> (v. Sep 1997).
  www.humboldt.edu/~jsb7/C/TCPIP/FTP/overview.shtml

- Bharat, Krishna and Andrei Broder. 1998. <u>A technique for measuring the relative size and overlap of public Web search engines</u> (v. Apr 1998).
  www7.conf.au/programme/fullpapers/1937/com1937.htm

- Berners-Lee, Tim. 1998. <u>Frequently asked questions by the Press</u>. (v. Dec 1998).
  www.w3.org/People/Berners-Lee/FAQ.html

- Berners-Lee, Tim. nd. <u>The World Wide Web: A very short personal history</u>. (v. Dec 1998).
  www.w3.org/People/Berners-Lee/ShortHistory

- Bumgarner, Lee S. 1995?. <u>USENET: The Great Renaming 1985 - 1988</u>. (v. 1995?).
  www.vrx.net/usenet/history/rename.html

- Cailliau, Robert. 1995. <u>A Little History of the World Wide Web</u> (v. Oct 1995)
  www.w3.org/History.html

- Ciolek, T. Matthew. 1998a. <u>The Seven Golden Rules of the Asian Studies WWW Virtual Library</u>. The Australian National University, Canberra. (v. 10 Apr 1998).
  coombs.anu.edu.au/WWWVLAsian/VLRules.html

- Ciolek, T. Matthew. 1998b. <u>Asian Studies and the WWW: a Quick Stocktaking at the Cusp of two Millennia</u>. In: Shen, Stella (ed.). 1998. Proceedings of the Pacific Neighbourhood Consortium (PNC) Annual Meeting, Academia Sinica, Taipei, Taiwan, 15-18 May 1998.
  www.ciolek.com/PAPERS/pnc-taipei-98.html

- Ciolek, T. Matthew (ed.). 1999a. <u>Annotated Guide to WWW Search Engines</u>. (v. Jan 1999).
  www.ciolek.com/SearchEngines.html

- Ciolek, T.M. 1999b. <u>Global Networking: a Timeline</u>. (v. Jan 1999).
  www.ciolek.com/PAPERS/milestones.html

- Compaq Corporation. 1998. <u>Altavista- The Internet's Home Page</u> (v. Dec 1998).
  www.altavista.com

- Deutsch, Peter, Alan Emtage and Bill Heelan. 1995. <u>Archie(*) - An Electronic Directory Service for the Internet</u> (v. Feb 1995).
  ftp://archie.mcgill.ca/pub/bunyip-docs/archie.whatis

- Dewey, Patrick R. 1987. Essential Guide to Bulletin-Board Systems. Westport, Connecticut: Meckler.

- Dodd, Carol Anne. 1992. <u>What is a FidoNet? Discussion of the Growth and Development of an Amateur Computer Network</u>. Carleton University March, 1992. Working Papers in Public Access Networks (v. Mar 1992).
  www.carleton.ca/~jweston/papers/0992.dod

- Effinger, Anthony and Mylene Mangalidan. 1998. Sun to set Java free. The Australian. 15 December, 1988. pp.50.

- Euro-Marketing. 1998. <u>Global Internet Statistics by Language</u>. (v. Dec 1998).
  www.euromktg.com/globstats/

- Frauenfelder, Mark. 1998. <u>The Future of Search Engines</u>. The Industry Standard, 25 Sep 1998. (v. Sep 1998)
  www.thestandard.com/articles/display/0,1449,1826,00.html?01

- Frisch, Aeleen. 1991. Essential System Administration. Sebastopol, Ca.: O'Reilly & Associatess, Inc.

- Gabriel, Richard P. 1996. Patterns of Software: Tales from the Software Community. New York and Oxford: Oxford University Press.

- Gilster, Paul. 1997. Digital Literacy. New York: Willey Computer Publishing.

- Goodman, Danny. 1987. The Complete Hypercard Handbook. Toronto-New York: Bantam Books.

- Google Inc. 1998. <u>About Google!</u> (v. Dec 1998).
  www.google.com/about.html

- Grobe, Michael. 1997. <u>An Early History of Lynx: Multidimensional Collaboration</u>. (v. 10 Apr 1997)
  www.cc.ukans.edu/~grobe/early-lynx.html

- Harold, Elliotte Rusty. 1997. <u>The comp.lang.java FAQ List</u>. (v. 6 Oct 1997).
  metalab.unc.edu/javafaq/javafaq.html

- Hauben, Michael and Ronda Hauben. 1995. <u>Netizens: On the History and Impact of Usenet and the Internet</u>.
  www.columbia.edu/~hauben/netbook

- Jacobi, Jolande (ed.). 1971. C.G.Jung - Psychological Reflections: ananthology of his writings 1905-1961. Henley on Thames: Ark Paperbacks

- Kahle, Brewster. 1996. <u>Archiving the Internet</u>. (v. 11 Apr 1996).
  www.archive.org/sciam_article.html

- Kahle, Brewster & Bruce Gilliat. 1996. <u>Alexa Internet Vision - Company</u> (v. Apr 1996)
  www.alexa.com/company/index.html

- Kelly, Kevin and Spencer Reiss. 1998. One Hudge Computer. Wired 6(8), August 1998, pp. 128-133, 169-170, 182.

- Kessler, Gary C. and Steven D. Shepard. 1997.<u>A Primer On Internet and TCP/IP Toolsand Utilities (RFC 2151)</u>. (v. Jun 1997).
  www.hill.com/library/rfc2151.html

- Koster, Martijn. 1994. Personal communication. London, May 1994.

- L-Soft International, Inc. (eds.). 1999. <u>CataList, the officialcatalog of LISTSERV lists</u>. (v. Jan 1999).
  www.lsoft.com/lists/listref.html

- La Tour, Adam Michael. (nd). <u>Gopher Presentation</u>. (v. March 1995).
  www.cox.smu.edu/class/mis6386/people/latour/amlpres.html

- Laursen, Jesper Vissing. 1997. <u>The Internet: Past, Present and Future - Internet & WWW History</u> (v. 7 Oct 1997).
  www.vissing.dk/inthist.html

- Lawrence, Steve and C. Lee Giles. 1998. Searching the World Wide Web. Science, Vol. 280 (3 April, 1998). pp.98-100.

- Levinson, Paul. 1997. The Soft Edge: A Natural History and Future of the Information Revolution. London and New York: Routledge.

- Liu, C. et al. 1994. Managing Internet Information Services. Sebastopol, Ca.: O'Reilly & Associatess, Inc.

- Network Wizards. 1999. <u>Internet Domain Survey</u>. (v. Jan 1999)
  www.nw.com/zone/WWW/top.html.

- Nielsen, Jakob. 1995. <u>Short History of Hypertext</u> (v. Jun 1995).
  www.useit.com/alertbox/history.html

- Notess, Greg, 1998. <u>Search Engine Showdown</u>. (v. 27 Oct 1998).
  imt.net/~notess/search/

- Palmer Robert R., 1943. Frederick the Great, Guibert, Bülow: From Dynastic to National War. pp. 49-74,In:Earle, Edward Mead (ed.). 1943. Makers of Modern Strategy: Military Thought from Machiavelli to Hitler. Princeton: Princeton University Press.

- Panstwowe Wydawnictwo Naukowe (PWN). 1964. Encyclopedia. pp.426-429. In: Wielka Powszechna Encyclopedia PWN, Vol. 3 (Dep-Franc). Warszawa: Panstwowe Wydawnictwo Naukowe.

- Popper, Karl Raimund. 1969. Conjectures and Refutations: The Growth of Scientific Knowledge. London: Routledge and Kegan Paul.

- Postel. J. 1980. Telnet Protocol Specification (RFC 764). (v. Jun 1994)
  www.cis.ohio-state.edu/htbin/rfc/rfc764.html

- Rheingold, Howard. 1994. The Virtual Community: Finding Connection in a Computerised World. London: Secker & Warburg.

- Reid, Robert. H. 1997. Architects of the Web: 1,000 days that Built the Future of Business. New York: John Wiley & Sons, Inc.

- Rutkowski, Anthony-Michael. 1994. The Present and Future of the Internet: Five Faces. Keynote Address. Networld+Interop 94 Tokyo 27-29 July 1994. (v. Jul 1994)
  info.isoc.org/speeches/interop-tokyo.html

- Schnierle, Matt. 1995. alt.conspiracy.usenet-cabal Frequently Asked Questions (FAQ). (v. 7 Sep 1995).
  www.landfield.com/faqs/usenet/cabal-conspiracy-FAQ

- Secret, Arthur. 1996. Coordination of the WWW Virtual Library. (v. 1996)
  www.w3.org/vl/coordination.html

- Selberg, Erik. 1997. A brief history of MetaCrawler (v. early 1998)
  www.metacrawler.com/help/selberg-history.html

- Severance, Charles. nd. A Brief History of Unix (v. Jan 1999).
  www.ocean.odu.edu/ug/unix_history.html

- St. Pierre, Margaret, et al. 1994. WAIS over Z39.50-1988 (RFC 1625). (v. Jun 1994)
  www.cis.ohio-state.edu/rfc/rfc1625.txt

- Sun Microsystems. 1998. Java Technology Home page. (v. Dec 1998).
  java.sun.com

- Southwick, Scott. 1997. What is IRC? (v. Dec 1998).
  www.liszt.com/www.liszt.com/chat/intro.html

- Southwick, Scott (ed.). 1998. Liszt, the mailing list directory. (v. 31 Dec 1998). www.liszt.com

- Tarnas, Richard. 1996. The Passion of the Western Mind: Understanding the Ideas Which Have Shaped Our World View. London: Pimlico Random House.

- Todino, Grace and John Strang. 1989. Learning the UNIX Operating System. Sebastopol: O'Reilly & Associates, Inc.

- Wolf, Gary. 1994. The (Second Phase of the) Revolution Has Begun (v. Oct 1994). www.wired.com/wired/2.10/features/mosaic.html

- Yesil, Magdalena. 1997. Virtual Store: Taking your Web Site from Browsing to Buying. New York: John Wiley & Sons, Inc.

- Zakon, Robert H. 1998. Hobbes' Internet Timeline. (v. 31 Dec 1998) info.isoc.org/guest/zakon/Internet/History/HIT.html

---

This WWW server is provided by the www.ciolek.com, Canberra, Australia.

Maintainer: Dr T. Matthew Ciolek (tmciolek@ciolek.com)

URL http://www.ciolek.com/PAPERS/pnc-taipei-99.html

[Asian Studies WWW VL][www.ciolek.com][Buddhist Studies WWW VL]