Regulation of the Internet

A Technological Perspective

"As we surge toward a new millennium, the Internet has become more than the overwhelming reality of the technology industry's current existence. It is the foundation for the Information Age, the environment in which we will all be living before long."[1]

Gerry Miller
Gerri Sinclair
David Sutherland
Julie Zilber

March, 1999

[1] Dan Gillmor, San Jose Mercury News, December 19,1998, http://www7.mercurycenter.com/business/top/069597.htm

Table of Contents

PREFACE

As is amply illustrated in this well-researched and well-written report, Internet is a global and rapidly growing phenomenon. From its early period as a research vehicle in the 1970s, Internet has already become a significant telecommunications infrastructure that is having a profound impact on the economy, education, personal communications, government, and the business of daily living. Despite its relative immaturity as a commercial enterprise (commercial Internet services were first offered by UUNET, an MCI WorldCom business unit, in 1990 for example), Internet is rapidly evolving business models that are changing the face of nearly every aspect of commerce worldwide

The historical development of Internet placed it outside the normal regulatory framework as a value-added or enhanced telecommunications service in nearly every country in which the system has taken root. That, alone, makes the Internet unique in the history of telecommunications, since virtually all other modes of communicating (television, radio, satellite, coaxial cable and even carrier pigeon) have developed in a regulated framework. This freedom has almost certainly been a key reason that Internet applications have developed so quickly and across such a wide range of services and functions.

What this report investigates, with considerable credibility and meticulous care, is the technical feasibility of controlling content or access to content on the Internet or, conversely, confining access to a selected set of source sites.

John Gilmore, one of the early founders of Sun Microsystems, is credited with observing that "Internet interprets censorship as damage and routes around it!" The authors of this report pose many technical possibilities for constraining use of Internet but virtually all of these suffer from serious implementation problems. I am in complete agreement with the general conclusion that local, optional, parental filtering of Internet content should be permitted but that it should not be made mandatory. Indeed, any such mandatory attempts would be met with great resistance by many if not all Internet users.

Moreover, any such attempts would be doomed to fail for the simple reason that there are technical flaws in any attempt to control contents on the web, in email, or other Net applications.

Internet Society believes that "Internet is for Everyone" and the report of this special Canadian group underscores the importance of such a goal as we approach the 21st Century.

Vint Cerf
Camelot, VA
April 1999

INTRODUCTION

This report discusses, from a technological perspective, issues arising from attempts to regulate content on the Internet and to control access by individuals to Internet sites and facilities. The discussion does not focus, except in passing, on the non-technical issues surrounding the regulation of Internet content.

The report was commissioned by Industry Canada. The authors are Gerri Sinclair and Julie Zilber from EXCITE, Simon Fraser University, and Gerry Miller and David Sutherland.

Gerry Miller is Executive Director of Information Services and Technology at the University of Manitoba. He was involved in NetNorth and was one of the founding member of CA*net in 1990. From 1992 to 1997 he was Chairman of the CA*net board of Directors and was instrumental in the success and growth of this founding Canadian Internet. He is also President of MBnet, an Internet service provider in Manitoba owned by the three Universities, chairman of MRnet, the Manitoba research and development network, and chairman if the CA*net Insitute, a granting agency for Internet related development projects founded by the CA*net community and Bell Canada. He is a member of the board and Secretary-Treasurer of CANARIE and was a member of both the Manitoba and federal Information Highway Advisory Councils. He is also a member of the Internet hall of fame.

Gerri Sinclair is the Executive Director of EXCITE, Canada's first multimedia research and production centre, which she established at Simon Fraser University in 1987. She is also the co-founder and President and CEO of NCompass Labs Inc., a high-tech Internet start-up company based in Vancouver, which develops enterprise-level Web design and content management software. Dr. Sinclair is a member of the Boards of Directors of BCT.Telus and the Canadian Foundation for Innovation. She is also a former member of both Canada's National Information Highway Advisory Council (IHAC) and the CANARIE board of directors. Sinclair has gained an international reputation for her pioneering work in developing interactive new media applications.

Julie Zilber is Director of Operations of EXCITE, and a University Research Associate (faculty researcher) in the Faculty of Education at SFU. A former lawyer, Zilber joined EXCITE in 1990, and since then has worked as part of a collaborative team of education, content and new

technology experts that has gained an international reputation for its innovative work developing telecommunications projects and interactive multi-media software applications for traditional and non-traditional learning situations. Zilber has been consulted in Canada and abroad on the impact of new technologies on the work, entertainment, and learning environments of the future. She has led research teams investigating issues ranging from human factors, interface, and technological design for on-line educational delivery systems to the design and development of interactive television applications.

David Sutherland is currently employed by CANARIE, and for many years was Director of Computer Services at Carleton University. He was a founding member of the first Freenet in Canada, and has been actively involved in many Internet activities such as SchoolNet. He was a member of the federal Information Highway Advisory Council.

Part 1 of this report sets the context by presenting a history of the Internet, a description of the technology used, recent trends, and various statistics.

Part 2 deals with various approaches towards the control of Internet content and the inherent difficulties in implementing them in a large-scale environment, like the Canadian Internet. It looks first at methods for restricting access to content and then at methods for the promotion of certain types of content.

There are appendices for further reference, as well as a glossary of terms attached to the report.

The authors wish to thank Lynette Miller and Jacob Zilber for their editorial assistance and Edwin Hargrave for valuable technical inisghts.

SUMMARY OF CONCLUSIONS

The authors have concluded that while a number of technologies exist that could be applied toward the regulation of Canadians' access to Internet content, none of these technological approaches would effectively prevent the Canadian Internet user from accessing content that violates pre-defined rules of acceptability, nor would they ensure that the user would be exposed to any measure of desirable content.

There are basically two technological approaches to restricting access to content on the Internet. These are:

- blocking requests for identified "unacceptable" content using a list of prohibited sites, and
- filtering of content by identifying prohibited text strings on the basis of partial or full-text searches or, by detecting rating labels attached to the content.

Even in relatively small-scale environments such as corporate Intranets, blocking and filtering are expensive to implement and maintain and they impose delays and inefficiencies in network performance. On a national scale, such measures would have enormous cost and performance implications and would, in effect, cripple the Canadian Internet and make it uncompetitive with the rest of the world. Imposing these costs on Canadian Internet service providers would drive some out of business and drive others to the US. The broader economic costs of imposing these measures is beyond the scope of this report, but they would be significant and would in all likelihood undermine Canada's ability to take full advantage of the economic and social opportunities offered by the wide spread use of Internet technology. Imposing blocking and filtering technologies would also have an adverse impact on investment in Canadian telecommunications resulting in a direct economic loss to the Canadian economy. Finally, the question may be moot, both because blocking and filtering technologies are of limited accuracy and because there are a number of technical means of circumventing blocking and filtering systems.

Approaches to content promotion considered in this report include:

- requiring Canadian ISPs (Internet Service Providers) to operate portal sites containing specified percentages of the desired type of content. However, users have complete freedom to decide if they wish to visit a portal site or make it their home page. If operators of portal sites find they are losing visitors because of the type of content they are legislatively required to include, they will either stop providing the service or move it to the United States where they will not face content legislation. There are also significant problems in defining type of content and in measuring the amount of that content at a particular site.
- special versions of Internet search engines that would prioritize the desired type of content for Canadian Internet users. For such systems to be technically effective in prioritizing Canadian sites when presenting the results of a search, three things are necessary. First, there must be a way of determining in which country the user is located. Second, there would have to be a way for the search engine to identify sites by their country of origin. Third, companies operating seach engines, most of which are located outside of Canada, would have to agree to implement this system in their search engine software. All of these requirements pose significant technical barriers.
- substitution of non-Canadian banner ads on web pages with Canadian banner ads. However, implementing such a facility would require intercepting all Internet traffic entering Canada in order to insert Canadian content, which would bring network performance to a crawl. As well, content providers gain revenue from placing banner ads on their pages. If Canada was stripping out the banner ads that had been paid for and replacing them with substitute material, the operators of these sites would in all probability bar access to their sites from Canadian sources to the extent that this is possible.
- "pushing" the desired type of content at Canadian Internet users through email or some other means. However, the user has complete freedom in deciding whether or not to read email and most client email systems allow the user to filter out unwanted messages.
- improving the ease of access to desired content by providing Canadians with high speed network connections to the desired content. This potentially viable approach to promoting content is in fact taking place as more investment is made in high-speed network connections to servers and in the implementation of web caching technology. Mandating it through legislation is not necessary.

While various methods might be considered for promoting certain kinds of content on the Internet, we have concluded that no purely technological approach will guarantee that Canadian Internet users will be exposed to that content.

Ultimately, it is the quality of the content and its interest to users that will determine whether Internet users decide to look at it. The steps the government is taking in supporting initiatives such as CA*net III , Schoolnet and the Community Access program, Strategis and other government web sites, as well as existing and emerging programs for the development of outstanding Canadian content, will do much more than any regulatory regime to ensure that Canadians access Canadian content on the Internet.

The authors note that it may be possible to control Internet content *to some extent*, but only if we are prepared to accept considerable costs in terms of technological infrastructure, human resources, enforcement mechanisms, and social and legal consequences. For instance, it may be possible to produce a non-comprehensive list of Web sites that violate Canadian legal standards and to require Canadian ISPs to filter for these prohibited sites. As both the technological discussion in this report and the experience of other countries indicate, such an approach would only restrict access to a limited number of offending sites. It would not guarantee that Canadians would be protected from other Internet content that violates Canadian legal standards and has not been screened by an authoritative body responsible for composing a list of prohibited sites. It also may be possible to require Canadian portal sites to display a number of banners containing specified content or to include a list of hyperlinks to other sites containing that specified content. The report shows that such an approach would fail to ensure that Canadians were exposed to a particular kind of "desired" content. However, if we are prepared to accept both the substantial costs (discussed in the report), as well as the consequent technological and operational problems (e.g., lack of accuracy, performance degradation, lack of scalability, administrative overhead, etc.), some might view this sort of ISP filtering as a "best-efforts" technological approach to regulating Internet content. While it might satisfy the concerns of some Canadians, the authors believe that the unreliable, hit-and-miss results of this approach would not justify its costs or the ensuing negative impact on Canada's place in the global Information Economy.

The authors also wish to point out that much content that is not suitable for children is legal for adults in Canada. Therefore, blocking of content that is not suitable for children is not appropriate at the level of the ISP because, by restricting material that is not suitable for minors, the ISP would also be denying legitimate access by adults.

For all of the foregoing reasons, we believe that the most promising technological avenue for regulating access to Internet content is self-regulation through voluntary client-side filtering (e.g., using software such as Net Shepherd or SafeSurf) combined with voluntary self-labeling of Internet content by content providers (e.g., using a PICS-compliant labeling system). Restricting access to some types of Internet content by children is an important issue that must be addressed. However, attempting to exert this control through a national regulatory framework

requiring blocking and/or filtering facilities is impractical and ultimately ineffective. Despite the limitations of filtering software discussed in the report, filtering software installed on the family PC may meet the majority of the needs of those parents who wish to restrict their children's Internet access. It is a good first step.

PART 1 SETTING THE CONTEXT

1. AN INTERNET PRIMER

1.1 WHAT IS IT?

The Internet is a global collection of networks connected and sharing information through a common set of protocols. Perhaps its most powerful feature is that it allows computers attached to networks to communicate openly and effectively regardless of make, architecture, operating system or location. All resources and network management are widely distributed. There is no central point of control.

1.2 WHO OWNS IT?

No one can completely own the Internet. Each network in the collection of interconnected networks is in charge of its own area, each is owned by distinct stakeholders, and all work together according to common sets of rules and standards. No one is forced to connect but it is in the interests of all to be connected and to enjoy global communication.

1.3 HOW DOES IT WORK?

The original research that led to the Internet was motivated by the desire to build a communications infrastructure that could survive nuclear attack. This of course implies that if a portion of the network is disabled, the rest of the network should survive. By definition therefore, there can be no central point of control.

The technique developed to ensure the flow of information over the Internet is called "packet switching." Unlike the telephone system, this technology delivers data between two points without a direct fixed connection or circuit. Data is broken into packets which contain addresses. The network delivers those packets to the destination by routing them through a succession of

interconnected computers, called routers, much as mail is passed through different postal facilities before being delivered. At the final destination the data is reassembled into its original form. Each packet may take a different route, and if part of the network is slow or unavailable, the packet is sent through a different route.

As an analogy, imagine a jigsaw puzzle being mailed, with each piece being put into a separate envelope. Each letter may take a different route to the same destination. Until all pieces are delivered and the puzzle reassembled, you don't know what the picture is. Intercepting one or a few envelopes and opening them does not give you any idea of the whole picture.

Transmission Control Protocol (TCP) ensures that packets are carried over the network without error. Internet Protocol (IP) ensures that packets are delivered to the correct destination. The two combined are known as TCP/IP and are the fundamental underlying architecture of the Internet.

Each computer connected to the Internet is assigned an IP number, which is its address. They are analogous in many ways to telephone numbers. Packets are delivered to their destination using that address. Domain names are aliases to IP addresses, and are more intelligible to humans. Thus, for example, people know the University of Manitoba web site as www.umanitoba.ca rather than 130.179.16.50. IP numbers are assigned in blocks to regions and organizations. If one had a concordance of IP number assignments, it might be possible to determine the location of a particular number. Although many systems now allocate IP numbers dynamically and only for the duration of a particular session, these IP addresses still fall within the domain of the host server. While dynamic IP addressing may make it more difficult to identify the exact geophysical location of a particular computer, it is still possible to determine the domain (and hence the region) to which the computer's dynamic IP address belongs.

1.4 WHO GOVERNS THE INTERNET?

The Internet has no single central governing body. There are, however, a number of organizations that work cooperatively to establish standards for interoperability.

The Internet Society (ISOC) is a non-profit, non-governmental organization that plays a support role in many Internet activities. It houses the Internet Architecture Board, The Internet Society Engineering Steering Group (which manages the standards work of the Internet Engineering Task Force). It also hosts the Internet Research Task Force as well as sponsoring training activity

in the form of international networks and various conferences including the annual international INET meeting.

IP addresses and domain names are assigned by independent bodies. Until 1998 the US Department of Commerce was responsible for issuing Internet addresses. Recently a transition to a new international organization was started.

There are also many organizations and trade associations in different countries who act on behalf of their region.

1.5 WHO ARE THE PROVIDERS?

The first level of provider is the Internet Service Provider (ISP), which delivers, in effect, Internet dial tone. Users, who may be individuals or companies, contract with this provider for a dial-in or dedicated connection to the provider's equipment, which then gives them access to the Internet. The ISP may be a private for-profit organization, a non-profit community organization, an educational institution, or a government agency. Any user may be a client or a host. In other words, they may be accessing information or supplying it.

The ISP may then connect to a Regional Network Provider (RNP) which operates a wide area network and provides Internet connections across a geographic market area.

RNP's then connect to the Internet backbone through Network Access Points (NAP's). The backbone is operated by service providers who operate the networks that route TCP/IP packets from point to point. These providers may use carrier facilities from telephone or cable companies, or may use their own facilities. Connected together, they are the global public Internet backbone.[1]

There are no regulations in Canada that govern who may be an ISP, just as there are no regulations that govern who can be a bookseller or who can build a library. In addition, there are no regulations that control interconnections between ISPs. It is possible to interconnect providers using regulated carriers such as telephone companies or by using unregulated technologies such as spread spectrum radio. Also, with the emerging satellite environment, an even greater number of infrastructure bypass options exist.

[1] See Appendix D for an Internet diagram.

1.6 IS THE WEB THE INTERNET?

No. The Internet is a network of networks made up of computers and network infrastructure, wired and wireless. It delivers packets of information anywhere in the world, typically in less than a second.

Many different kinds of software programs use the Internet to exchange information: electronic mail, for example, was around long before the global hypertext system called the World Wide Web. Now, videoconferencing and streamed audio channels are among other things which, like the Web, encode information in different ways and use different languages between computers ("protocols") to provide a service.

The Web is an abstract (virtual) space of information. On the Internet, you find computers -- on the Web, you find text, pictures, sounds, videos, etc. On the Net, the connections are made over network infrastructure between computers; on the Web, connections are hyperlinks, links between documents. The Web exists because of programs which communicate between computers on the Net.

The Web uses the Internet and makes it more useful because people can now get information from thousands of locations without having to know about the technical architecture of the network.

1.7 WHAT IS AN INTRANET?

Many companies and public sector organizations have realized that it is more cost effective to use the public Internet in their operations than to build private networks. They have also installed private Internets in their organizations for internal communications between regional locations. These are called Intranets. An organization's Intranet in one location will communicate with an Intranet in another location of the same organization over the open public Internet. In most cases the information sent from one location to another is part of the core business operations, and is vital to the organization's welfare. The essential difference is that an Intranet is closed and may only be used by those who are authorized to do so by the organization which owns the Intranet. The Internet is open to all.

1.8 WHAT ABOUT SECURITY AND HACKERS?

If an organization connects its Intranet to the public Internet, which may be accessed by anyone with a modem and computer, it becomes the organization's responsibility to take technical measures to protect its internal networks from unauthorized external access. This may be done using devices called firewalls, which filter all of the information going in and out of the Intranet. Unauthorized attempts at access and certain activities are denied. Sensitive information being sent from one Intranet to another, such as credit card numbers, SIN numbers, etc. may be encrypted to prevent theft.

2. A BRIEF HISTORY OF THE INTERNET

2.1 A TIMELINE OF EVENTS IN NORTH AMERICA

Hobbe's Internet Timeline, a chronology of the Internet in North America, which is maintained by Robert Zakon of the Mitre Corporation, is the definitive work on key events and technologies in the history of the North American Internet[2,3] A summary of key events, based on that timeline, follows. Canadian events are discussed in a subsequent section.

The first research into internetworking technologies occurred in the 1960's at various US universities and research institutions. These efforts were sponsored by the Advanced Research Projects Agency (ARPA) of the US Department of Defense (DoD). The research dealt with the possibility of developing packet switching networks with no single points of failure, cooperative sharing of computing facilities across telecommunication networks, and designing and building packet switching equipment. This research determined that such technologies were feasible and ARPA issued a request for a proposal to build a prototype in 1968. Awards were made to UCLA for network modelling and measurement, and to a company called Bolt, Bernak and Newman (BBN) for network management and building Interface Message Processors (IMPs).

The initial network called ARPANET had four nodes: UCLA, Stanford Research Instiitute, University of California Santa Barbara and University of Utah. The telecommunications lines, supplied by AT&T, had a bandwidth of 50 Kbps.

ARPANET grew slowly in the early 1970's and membership consisted solely of universities and research labs. Research into network management and protocols continued and in 1974 Vinton Cerf and Bob Kahn published a paper on a protocol for packet network connection which detailed the Transmission Control Program (TCP) standard.

Also in 1974 BBN established the first commercial packet switching network, Telenet.

---

[2] Hobbe's Internet Timeline © 1993-9 by Robert H Zakon. The current version is available at http://www.isoc.org/zakon/Internet/History/HIT.html .

[3] For a comprehensive history of the Internet, the reader is referred to "Where Wizards Stay Up Late:, Katie Hafner and Matthew Lyon, Simon and Schuster, ©1996.

In the late 1970's as the network grew, electronic mail standards were developed and email networks established among researchers in Computer Science and other disciplines. The growing popularity of email catalyzed network growth and ARPANET expanded.

In the early 1980's other academic networks were established to foster communication and sharing of resources. Two notable examples were BITNET (the "Because It's Time") network founded by City University of New York and Yale in 1981, and CSNet (Computer Science NETwork), established by a number of institutions with start-up money from the National Science Foundation. These networks expanded into Canadian universities, where BITNET became NETNORTH, established by Canadian universities with funding from IBM Canada. Network connections speeds were 56Kbps or less. The primary use was file transfer and email.

Around the same time, ARPA established the Transmission Control Protocol/Internet Protocol (TCP/IP) standard. This open non-proprietary standard permitted the interconnection of equipment from different manufactures across a common network and allowed the first "internet," a connected set of networks, to be formed. This protocol is still the one in use in today's Internet. Establishing open standards for internetworking was one of the seminal events in the Internet's history as it allowed different kinds and sizes of computers to talk to each other. This principle of open interoperability is a fundamental building block of the Internet and is necessary for its existence.

Another important parallel development was the development of name servers. Up to this time it was necessary to know the actual numeric IP address of the destination. These are obscure and have no inherent meaning to a user. Name servers allowed substitution of a name with some meaning rather than an actual address. This is why it is possible, for example, to connect to microsoft.com rather than 207.46.130.149. The Internet would be far less usable without this facility. This led to the introduction of another fundamental Internet building block, the domain name system (DNS).

1986 saw the establishment of NSFNET in the US. This network, funded by the National Science Foundation, interconnected 5 supercomputing centres at American universities over 56Kbps lines. This was truly a national Internet, and regional networks sprang up around these five nodes to allow other institutions to connect to the national backbone. This resulted in a rapid increase in connections from universities and other R&D organizations. It was, however, still a non-commercial network and would remain so for some time.

In the same year, the Internet Engineering Task Force (IETF) and Internet Research Task Force (IRTF) were established. Over the next few years, these two organizations would develop new technologies and standards, which allowed the growth of the Internet.

By 1989 the NSFNET backbone had been upgraded to T1 (1.544Mbps) speeds and the number of hosts on the network exceeded 100,000. Networks from Canada and Europe were connected to the US backbone.

Over the next two years, R&D networks flourished in the US and other countries until by 1992 the number of hosts exceeded 1,000,000 and backbone speeds were at T3 (45Mbps). Internet tools such as Gopher, Veronica and Archie appeared. The term "surfing the Internet" came into common usage. Countries from all around the world connected to NSFNET, and the global Internet started to appear.

In 1993, the National Center for Supercomputing Applications (NCSA) at the University of Illinois released software called MOSAIC, the first World Wide Web (WWW) browsing program. By the end of 1993, there were sixty-three Web servers in the world.

1994 was the twenty-fifth anniversary of ARPANET. The Internet was growing rapidly, and this year saw the connection of the US Senate and House of Representatives, the White House and other government services not only in the US but also Japan, Britain and others. Internet shopping malls or cyber-malls appeared. Community networks or freenets came on the scene. Cyberbanks opened for business. For the first time you could order a pizza online.

In 1995 the NSFNET commercialized the backbone and went back to funding R&D networks. Interconnected commercial network providers now operated the US national Internet backbone and commercial traffic proliferated. The WWW exploded and early in the year became the biggest source of traffic on the Internet. Companies started to see business opportunities, and many Internet related companies went public, resulting in some interesting stock activity. The Canadian government came online, and development of many government web sites commenced. Backbone speeds increased regularly, and by the end of the year the host count was over 6,000,000.

In 1996 growth, driven by the WWW and commercial use, continued exponentially. By the end of 1996 the host count was over 16,000,000. There was much discussion of the governance of domain names, as their commercial value became apparent. Governments in countries such as China, Germany, Malaysia and Singapore attempted to control their citizenry's access to the

Internet for political reasons, usually with marginal success. The Communications Decency Act, an attempt to control Internet content through legislation, was passed by the US Congress. It was declared unconstitutional the following year by the Supreme Court.

1997 saw the further commercialization of the Internet and continued exponential growth driven by the web and the emergence of electronic commerce. Most major companies were developing a web presence. By the end of the year the host count was over 25,000,000.

A major issue in 1998 was the privatization of the domain name system, managed up to that time by the US government. Many countries, including Canada, became concerned about U.S. private sector control over what many have come to see as an international public resource. The issue has still not been resolved. The number of pages on the web exceeded 300 million. The number of hosts reached 40 million. Electronic commerce grew rapidly, and business conducted on the web along with the wealth created by the information technology sector of the economy became a major contributor to GNP.

In summary, over the past three decades, the Internet has evolved from a secret, closed technology used by the academic and military communities to a pervasive, open, uncontrolled, flat system spanning the globe. The fact that it is global makes central control impossible.

The Internet has moved from a military tool to an academic tool to a populace tool to an economic tool. It will continue to grow and evolve.

2.2 THE CANADIAN EXPERIENCE (FROM R&D TO COMMERCIAL)

The history of the Canadian Internet closely parallels the American experience. In the 1970's, there were regional networks in a number of locations interconnecting Universities in the region. These networks used proprietary communications protocols, and, typically, interconnected large mainframe computers. The main use was for transferring large files of information.

At the start of the 1980's, newer networking technologies started to appear. CDNnet, a research network founded to develop email standards was established and connected a number of Universities in the country. NETNORTH, the Canadian equivalent of BITNET in the US, was established with the help of funding from IBM Canada by the University community as a national network, and was connected to similar networks in other countries. Email became a way of life for the academic community.

Towards the end of that decade, the first TCP/IP networks were established in Canadian Universities in Ontario and British Columbia. These were connected directly to the US backbone with cross-border links, and part of the Canadian academic community became members of the burgeoning Internet community.

In 1989 the NETNORTH board of directors, made up of representatives from the Canadian University community, developed a strategic plan to carry NETNORTH forward and transform it to a TCP/IP technology. Funding was sought from the federal government and a $2,000,000 start-up grant was awarded by the National Research Council (NRC). In-kind contributions were also received from IBM Canada.

The NETNORTH community incorporated a not-for-profit organization to operate the network, called CA*Net Networking Inc. The board of directors was made up of one representative from each province in the country as well as representatives from the University of Toronto, the network operator and from NRC. Most of the board members were from the University community.

At the same time, regional academic networks were established in each province:

British Columbia:	**BCNet**
Alberta:	**ARNet**
Saskatchewan:	**SASK#Net**
Manitoba:	**MBNet**
Ontario:	**Onet**
Quebec:	**RISQ**
New Brunswick:	**NBNet**
Prince Edward Island:	**PEINet**
Nova Scotia:	**NSTN**
Newfoundland:	**NLNet**

CA*Net interconnected these regional networks and provided three connections to the NSFNet in the US through Vancouver, Toronto and Montreal. The original connections were 56 Kbps, but the rapid growth of Internet traffic in 1990 and 1991 drove the need for increased network capacity.

In January of 1993, the federal government announced the formation of CANARIE, an organization created to stimulate industrial research and development on broadband network facilities and applications. One of its first initiatives was the upgrading of the CA*Net backbone to T1 speeds, or 1.54 Mbps. Similar upgrades were done in regional networks. At the same time, CANARIE funded the connection of Canada's north by funding links to regional networks in the Northwest Territories and the Yukon.

Internet growth in Canada paralleled the experience in other countries. It became exponential, and further upgrades were required to T3 speeds or 45 Mbps. In some cases multiple T3 connections were needed, particularly on the US links.

In 1995, the University of Toronto stopped operating the network, and, after a tender process, network operations were awarded to Bell Advanced Communications.

In 1996, it became evident to the CA*Net board of directors that the Canadian Internet had evolved beyond its origins as an academic research and development network to a fast-growing commercial network. The board then decided the time had come to transition the Canadian Internet to a commercial one, and after another tender process Bell Canada was awarded the network. It now operates as a commercial Bell offering. In recognition of the work of the founding CA*Net community, CA*Net and Bell Canada created the CA*Net Institute, a funding organization dedicated to promoting the use of the Internet in the spirit of the original CA*net.

This organization is in place and the first awards have been given to a wide variety of Internet-related projects.

At the present time, this backbone network is one of many in Canada. Companies such as Sprint, BCT.Telus, and MetroNet as well as Bell are installing and upgrading national Internet backbone networks, connecting to the global Internet through a number of locations. Speeds of these backbones are up to 655 Mbps, 12,000 times faster than the original CA*Net nine years ago. Theoretical speeds using new broadband network technologies are up to 1.5 Tbps, another large increase. Since the unit cost of bandwidth becomes cheaper as overall network speeds increase, the availability of higher speeds encourages network growth and its use by a widening clientele in both the public and private sectors.[4]

These networks are connected to the global Internet through cross-border connections to the US, Europe and Asia. As well, there are many private connections outside Canada for corporate Intranets. While the number of cross-border Internet connections is not easily determined, it is large and growing.

The volume of traffic on the Canadian Internet is growing at typical rates, doubling every 4-6 months. Since there are a number of national Internet backbones, and since such information is proprietary for competitive reasons, determining total traffic is difficult. However, it is certainly now in the hundreds of gigabits per second, and is rapidly approaching terabits per second.

The other fundamental change in the Canadian Internet has been the shift from research traffic to commercial use. Just a few years ago, the majority of the traffic was for research and education purposes. Now, of course, the traffic is overwhelmingly commercial.

[4] In BC for example, BCT.Telus, alone, now supports 1.5 million high speed nodes.

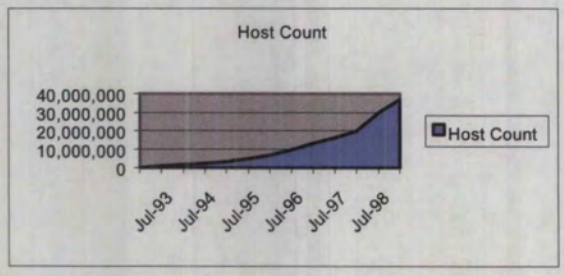
3. SOME INTERESTING STATISTICS

The following table of host counts with accompanying chart illustrates the exponential growth of the Internet in the last few years.[5]

Internet host counts

Data as of July 1998

Date	Host Count
Jan-93	1,313,000
Jul-93	1,776,000
Jan-94	2,217,000
Jul-94	3,212,000
Jan-95	4,852,000
Jul-95	6,642,000
Jan-96	9,472,000
Jul-96	12,881,000
Jan-97	16,146,000
Jul-97	19,540,000
Jan-98	29,670,000
Jul-98	36,739,000



[5] Source - Network Wizards. Data is available at http://www.nw.com

Appendix A presents a table[6] breaking down Internet hosts by domain.

It is interesting to note that Canada ranks fifth in number of hosts, after the US, Japan, Germany and the UK. This has been our traditional ranking for a number of years, indicating that Canada is keeping up with network growth and is ahead of many larger countries in use of the Internet, at least by this measure.

The following estimate from NUA Internet surveys[7] gives an idea of the present size of the Internet:

> **The art of estimating how many are online throughout the world is an inexact one at best. Surveys abound, using all sorts of measurement parameters. However, from observing many of the published surveys over the last two years, here is an 'educated guess' as to how many are online worldwide as of March 1999. And the number is 158 million.**

World Total	158 million
Africa	1.14million
Asia/Pacific	26.55million
Europe	36.76million
Middle East	0.78 million
Canada & USA	88.33 million
South America	4.63 million

[6] Source - Network Wizards. Data is available at http://www.nw.com

[7] NUA Internet Surveys, http://www.nua.ie/surveys/how_many_online/index.html

4. ELECTRONIC COMMERCE - A NEW BUSINESS PARADIGM

In the past year, the use of the Internet for business-to-business and consumer retail transactions, colloquially known as e-commerce, has come of age. According to Forrester Research,[8] US on-line business trade will explode from $43 billion to $1.3 trillion by 2003, and will surpass 9% of total business trade by that year. The leading industries using e-commerce will be computing and electronics, aerospace and defense, petrochemicals, utilities and motor vehicles. Industry adoption of e-commerce will be driven by the network effect, in which the value of participating increases dramatically as more and more companies join in.

> **Companies that are unprepared to compete on-line will be pushed aside by competitors who understand how to use the Internet to generate new values and efficiencies for customers.[9]**

Similar growth will occur in Canada. For example, according to an Ernst and Young survey[10], the revenue generated by the on-line advertising industry was expected to reach $US 13.4 million in 1998, and is expected to grow to $US24.4 million in 1999, an increase of 82 percent.

E-commerce will be an important contributor to national wealth. It is therefore critical that Canada has a viable and reliable Internet.

The IHAC[11] recognized the importance of e-commerce to the Canadian economy in its final report, as is evidenced in the following three recommendations[12]:

> **3.5 Strengthening the emerging role of the Internet as a platform for electronic commerce should be the central economic strategy for promoting the knowledge-based economy. It is imperative that the government reinforce its efforts, both nationally and internationally, in the following areas:**

---

[8] Data available at http://www.forrester.com
[9] Stuart D. Woodring, Vice President of Research at Forrester Research; http://www.forrester.com
[10] Data available at the NUA Surveys site; http://www.nua.ie/surveys
[11] Canada's Information Highway Advisory Council
[12] Preparing Canada for a Digital World- final report of the Information Highway Advisory Council, Industry Canada, September, 1997, pages 34, 35. Available electronically at http://strategis.ic.gc.ca/IHAC

a. **development and application of open networking standards for interoperability and interconnection;**
b. **clarification of market rules in areas such as privacy, security, and consumer protection; and**
c. **removal of the legal, policy and regulatory impediments to the conduct of electronic commerce.**

3.6 The government should clarify its intentions regarding formal regulation of Internet-based services. Currently, the Council questions the effectiveness of any form of licensing of Internet-based services or the imposition of formal content rules or quotas. At the same time, the Council believes the rapid development of information technology has surpassed the present regulatory framework. The government should explore other potential instruments for achieving policy objectives regarding the Internet.

3.7 With respect to taxation of the Internet, the government should avoid fiscal measures that may hinder the development of the Internet and its contribution to economic growth.

The federal government has since initiated the "Canadian Electronic Commerce Strategy" with the objective of making Canada a world leader in the development and use of electronic commerce by the year 2000. In October, 1998, the Minister of Industry hosted an OECD conference on electronic commerce in Ottawa. The government has followed up with policy and legislative initiatives to advance its strategy.

Aside from performance and reliability of the Canadian Internet, two important technologies in e-commerce are encryption and digital signatures. Encryption, the technique of encoding information to protect it from unauthorized access as it is transmitted, ensures the security of sensitive information such as credit card numbers as it is carried over an open network such as the Internet. A digital signature, a variation of encryption technology, is the digital analogue of a person's signature, and is crucial to e-commerce as it identifies without doubt the identity of the buyer and seller. Any attempts to hamper these technologies as part of a strategy to control Internet content will have a stifling effect on the conduct of e-commerce, and therefore on the growth of the Canadian economy. It should be noted that as of the time of this writing, Bill C-54 is before the House of Commons. This bill provides that the Governor in Council may make regulations prescribing technologies or processes for the purpose of securing electronic signatures. This is not the type of control referred to here.

5. THE RISE OF PORTAL SERVICES

A recent phenomenon on the Internet has been the rise of portal services. Portals are sites that provide a gateway to Internet services, and are run by companies such as Yahoo, Excite Inc.[13] and America On Line (AOL). Internet users can set their "home page" to one of these sites so every time they start up their browser software it opens that site. Typically, registration on the site is free; they make their money from advertising and other services.

Portals are also starting to deliver services that provide an alternative to the traditional Windows or Macintosh operating environments according to CNET:[14]

> **Portal sites are rapidly emerging as a computing alternative to the traditional Windows, and even Mac, desktop. Free email was the first service provided by portal sites that mimicked a standard PC application. As a result of the success of free email, users can now bypass the standard desktop application suite in favour of scheduling software, address databases, and other 'productivity' applications found free on various web sites. Yahoo, Excite and Lycos, among others, are pursuing these strategies.**

Portal applications are attractive alternatives because they are free and do not require large programs on the PC. As they mature and become more popular, and as more homes get high speed access to the Internet via cable modems and ADSL lines, the use of these portal sites will increase traffic on the Internet.

Canadian portals are also starting to emerge. Since an Internet user can go to any portal site in the world with ease, the Canadian sites must show added value to be competitive. This requires a high performance, unfettered Internet, and the ability to tailor content, without regulation on the site, to attract customers. Canadians prefer to go to Canadian sites where they can get local news,

---

[13] Please note that Excite Inc., which operates an Internet search service, is not related to Excite at Simon Fraser University, which was established in 1987 by Dr. Gerri Sinclair, one of the co-authors of this report.

[14] "Portals: the new desktop" Stephanie Mills and Michael Kanellos; CNET News.com http://www.news.com/News/Item/0,4,31162.html

weather, sports and other Canadian information. Canadian portal sites develop Canadian information to attract customers.[15]

This has been confirmed by the fact that US companies that have established Canadian portal sites, such as America On Line (AOL), have done so based on business cases. Their investment in establishing portals in Canada has been justified by the market.

[15] A similar argument was made by various organizations including AOL Canada and Rogers in their recent submissions to the CRTC New Media hearings. Details are available on the CRTC web site.

6. THE CHANGING TELECOMMUNICATIONS ENVIRONMENT

For many years the telecommunications environment in Canada consisted of regulated monopolies. Carriers were closely regulated by the Canadian government, and pricing and profit margins were predetermined.

In recent years, the progressive introduction of competition has occurred in a number of areas of telecommunications. The government followed this course to create an environment favourable to private sector investment in infrastructure and innovation, and to encourage the growth of the Information Economy. Competition is now allowed in most areas of telecommunications, and as a result the cost to the consumer has come down.

Another phenomenon has been convergence. Separate networks with separate technologies for voice, data and video transmission are no longer necessary. The Internet can carry all of these simultaneously over any type of network infrastructure. Transmission facilities no longer have to be landlines. Internet networks can use a number of carrier facilities, including fibre optic cable, microwave, radio and satellites. Wireless facilities are already in use for delivery of Internet services in many parts of Canada.

> **"Metamorphosing the Internet from a high tech toy delivering best-effort service into a favoured business tool of the 21st century is a world-wide priority.**
>
> **An Internet that can provide the high level quality of service demanded by real time business communications such as Internet telephony, video conferencing and on-line transaction processing is very much in demand."[16]**

The Canadian government has recognized this in the funding of CA*Net III through CANARIE. This optical Internet network, which will be the fastest in the world when deployed through 1999, is a research and development platform for very high-speed broadband network applications. As the technologies are developed and rolled out commercially, the use of traditional circuit switched networks for local and long distance telephony, data transmission and

[16] "The Internet Grows Up"; George Lawton, Global Telephony, December-January 1999, page 34.

video-conferencing will slowly fade away, as telecommunications converges to a single broadband Internet.

7. WHAT'S IN THE FUTURE?

"The world's population will be about 11.5 billion by 2047, compared to 5.8 billion in 1996. Internet will probably achieve penetration rates similar to television and telephony, at least in the parts of the world that have suitable power generation and other technology infrastructure. Indeed, by that time, penetration may exceed that of television, with the use of personal and vehicular devices adding to conventional office and residential units. Instant demand for communication capacity will be satisfied in large measure by a combination of fibre optics and optical switching as well as very broadband radio communication and perhaps infrared links over relatively short distances. Broadcast communication via digital satellite will also play a role, and conventional over-the-air media will carry Internet packets. Conventional television and radio may by that time have become as quaint as crystal radio is today."[17]

"If the average penetration of networking technology reaches thirty percent by 2047, this suggests on the order of three to four billion devices, possibly more if the "ubiquitous computing" applications predicted by Mark Weiser of Xerox PARC actually proliferate. There may be hundreds of such devices in a residence, vehicle, or office. Moreover, wearable devices could inflate the total even further. Such scales are dramatically more than the present day network of 600 million terminations, which has already had a material impact on all aspects of the global economy and social structure.

Data rates will have reached the limits of optical fibre technology in the 38 THz range per fibre. End user data rates will be in the gigabit range and backbone rates in the tens of terabits range. Optical switching will be the norm."[18]

Vinton Cerf was one of the prime architects of the Internet and its underlying technology. These predictions are if anything conservative; the growth of the Internet and its effect on commerce and society are unstoppable.

---

[17] Vinton Cerf, "When They're Everywhere," Beyond Calculation: The Next Fifty Years of Computing, Copernicus ©1997, page 38.

[18] *Ibid* page 39.

Canada has always been a leader in telecommunications technology. Three-quarters of Internet traffic is routed through equipment manufactured in Canada. Initiatives such as CANARIE's CA*net III, the fastest optical network in the world, will keep Canada at the forefront of this technology. It has been noted above that the penetration of the Internet in Canada is disproportionately high in comparison to larger countries. This is expected to continue, and the Internet will become a fundamental part of Canadian society and business, if it is nurtured and allowed to grow.

8. EXPERIENCES OF OTHER COUNTRIES IN ATTEMPTING TO CONTROL THE INTERNET

As the Internet grew, some countries became concerned over the ability of their citizenry to access information the government felt was harmful or subversive. Attempts were made to control access through technical means. Two prominent examples are China and Singapore.

The governments in those countries required licensing of Internet Service Providers and users. Modems were licensed, and all Internet traffic was routed through a very small number of gateways so as to monitor content. This was possible because the Internet was very small in these countries, and also because under their regimes the government was allowed to censor the flow of information to the populace.

The success of these attempts is marginal. The nature of the Internet is such that attempts to block or filter information may be circumvented with regularity, as will be discussed in part 2 of this report. Media coverage summarizing these attempts is provided in Appendix E.

There have been cases in other countries which resulted in unsuccessful attempts to block content. For example:

> **The deputy leader of Germany's reform communist Party of Democratic Socialism was indicted by German authorities in January 1997 for creating a hypertext link on her home page in Germany to the Dutch left-wing magazine *Radikal* that advocates the overthrow of government through terrorism. One of the largest German ISPs was forced to block access to the magazine, hosted on the Dutch Internet site XS4All, but freedom of speech advocates fought back by setting up mirrors (other sites around the world hosting the same information), posting the complete magazine in an Internet newsgroup, sending protest letters, and distributing the magazine by E-mail.[19]**

[19] from Government Interventions in the Freedom of Expression on the Internet, Dennis Cheong, http://users.wantree.com.au/~zylantha/freedom.html_Toc389652105

A government can only attempt to control content within its own jurisdiction. Setting up of mirror sites is relatively easy, and transmitting the "forbidden" content back into the country attempting to control it is a common means of circumventing those controls. Similar means have been used to counteract state-sponsored attempts at content control in other countries including Russia and the former Yugoslavia. There is no effective way of counteracting this.

To our knowledge, no country has been completely successful in controlling Internet content, and, given the nature of the Internet, it is unlikely any country will be.

> **"The nature of the Internet is such that once it gets in, as long as there's a human spirit, it will find a way to get around any attempts at blocking, or controlling content, and it therefore becomes a universal, neutral medium for the transmission of information."[20]**

[20] from "Efforts to Censor 'Net in Asia Doomed"
http://www.freedomforum.org/technology/1998/1/28asiasociety.asp

PART 2 - CONTROLLING CONTENT ON THE INTERNET

In this part of our report we will examine the question of whether it is possible to use technological means to control the content available to Canadians on the Internet, either through the restriction of access to certain content or, alternatively, through the promotion of specified types of content. We will first explore technological approaches to restricting access to certain types of content. We will then look at how technology can be applied in order to promote designated or preferred types of content. Since most people using the Internet use the World Wide Web, much of this discussion will concentrate on restricting or promoting access to Web pages. It should be noted, however, that content can be exchanged over the Internet in many other forms, including via email, ftp, IRC, bulletin boards, and multicasting. The restriction or promotion of content exchanged using these methods will also be addressed.

1. RESTRICTING ACCESS TO UNACCEPTABLE CONTENT

Let us consider a scenario in which one wishes to ensure that an Internet user in Canada cannot access content that is deemed by officials to be unacceptable. For the purpose of this scenario, we will largely ignore the existing legal framework in Canada, and will assume that the government is both willing and able to enact whatever legislation is required to implement the restrictions described. We will therefore address only the question of whether it is technologically possible to prevent a Canadian user from accessing unacceptable material over the Internet. We will consider both the situation faced when the content originates on a host server located in Canada, and that which exists when the host server is located outside the country. It is important to note that eighty to ninety per cent of Internet traffic (web surfing, email, downloading files from servers, etc.) in Canada accesses servers outside the country, creating a situation in which authorities in Canada have no jurisdiction over the originating server. We will also consider technological options available to restrict access to content by a user who does not wish to have his or her access to any kind of content restricted. What technological avenues might be available to prevent the "unacceptable content" from reaching the user in Canada?

1.1 CLIENT-SIDE APPROACHES

We will begin by examining approaches to restricting content that might be implemented on the user's computer. This is called client-side filtering.

Filtering Software

Filtering software is software that compares some or all of the contents of a data file[21] retrieved by a user against a pre-defined set of rules, and determines whether to permit the file to be received and/or displayed by the user's computer. Common rules used for filtering include:

- Blocking of selected files (e.g., web pages and newsgroups) or sites by comparing the URL,[22] name, or IP address of each item against a list of prohibited files or sites (commonly called "blacklisting");
- Blocking of all files except pre-approved files or sites by comparing the URL , name, or IP address of each item against a list of permitted files or sites (a less common practice, this is sometimes called "whitelisting");
- Filtering of selected files by scanning the header information of each file and comparing the contents of the header against a list of prohibited text strings (sequences of text characters);
- Filtering of selected files by scanning the full text of each file and comparing the contents against a list of prohibited text strings;
- Filtering of selected files or sites by comparing a "rating label" included in the header information of each file or site against a pre-defined set of rating criteria.

The first two types of blocking require that a human being examine each possible file or site and decide whether to add it to the list of prohibited or authorized items. The second two types of filtering require that human beings create the list of prohibited words or phrases, after which the screening process is automated. The last type of blocking requires that human beings establish rating criteria and rate each file or site, after which rated files and sites can be screened in accordance with their ratings.

Many software companies have created client-side filtering software that the owner of the client computer can choose to install on the client machine.[23] These products commonly use a

[21] A file may be a document, a newsgroup, or any other item that is stored in digital form and can be accessed on the Internet. Web sites generally consist of a large number of linked files.
[22] URL stands for Universal Resource Locator, and is the address of a document on the World Wide Web.

combination of blocking and filtering to restrict access to Internet content, using both a list of prohibited sites and text-filtering. Web pages and newsgroups are the type of content most commonly blocked by client-side content filtering software products. Some client-side filtering products which scan content go further, and will delete prohibited words contained in real-time chats or even in unencrypted email. In some cases, client-side filtering products can also restrict the disclosure of personal information such as addresses and phone number. Some parents and schools are interested in this type of software to prevent children from being exposed to material deemed unsuitable or from revealing personal information that might put them at risk.

Blocking and content filtering

As many experts point out, however, filtering software is far from perfect. The most common problem with filtering software is that it either prohibits access to a wide range of acceptable content or, conversely, allows unacceptable content to slip through. The first situation tends to occur when the filtering is done on the basis of partial- or full-text scanning of documents. Perhaps the best known example of this problem occurred when AOL decided to use the word "breast" as a criterion for filtering out pornography sites, and, as a result, inadvertently blocked all sites dealing with breast cancer as well as sites which included recipes for chicken breasts.[24] A study by EPIC (the Electronic Privacy Information Center), a public interest research organization based in Washington, D.C., of sites blocked by Net Shepherd Family Search[25] showed that 90% of all sites found by the underlying search engine (AltaVista™) were excluded by that filtering software.[26] This included such seemingly useful sites as Arbor Heights Elementary School, the San Diego Zoo, the National Aquarium, the Smithsonian Institute, and the vast majority of the sites about children's author Dr. Seuss.[27] Similar results, showing that content blocking or filtering software often denies access to acceptable content, have been found with a wide range of other products on the market.[28]

The converse situation, in which some (even the majority) of unacceptable sites will not be filtered, tends to occur when the filtering system functions by blocking documents using a predetermined list of prohibited sites. Given the vast number of sites on the World Wide Web

---

[23] See Appendix C for a list of a number of software-based filtering products.
[24] AOL changed their filtering criteria because of the resultant outcry from users.
[25] Net Shepherd Family Search is a web-based search engine located on the Internet at http://family.netshepherd.com.
[26] For more information on this study, visit http://www.epic.org/reports/filter-report.html. For more information on EPIC, visit http://www.epic.org.
[27] Of the 2,638 references to Dr. Seuss found by the AltaVista search engine, 2,630 were blocked by Family Search. Ironically, one of the sites not blocked was "a parody of a Dr. Seuss story using details from the murder of Nicole Brown Simpson." Ibid.

(estimated at over 300 million in 1998, and growing at an estimated rate of 40,000 per day), the task of compiling a comprehensive and timely list of unacceptable sites would involve an army of reviewers and be prohibitively expensive. A simple calculation explains both why filtering based on human review is cost prohibitive and why filtering software is so inaccurate. Using a conservative calculation of 2 minutes to review a site, it would take 177 people working 7.5 hours a day just to keep up with rating the estimated 40,000 new sites that are coming on line each day. To rate the estimated 300 million existing sites would take 10 million person hours, or 266,667 person weeks (at 37.5 hours per week). To accomplish this task in a year would take a team of 5,500 reviewers. At the minimum federal wage of $7.25/hour, salaries and basic benefits for these reviewers would be approximately $100 million per year. (Not including the costs of providing facilities, Internet connections, computers, etc.) This calculation also ignores the fact that many sites consist of thousands of pages and cannot possibly be reviewed in two minutes. (Because of the difficulties associated with comprehensive review of web sites, many filtering companies have adopted the practice of blocking an entire site if even one page at the site has content that fails to meet their screening criteria.) This form of blocking also ignores the constantly changing nature of content on the Internet. Unlike a book or a movie, the contents of which are fixed once it is published, the contents of an Internet site can be, and frequently are, updated monthly, weekly, daily, or even hourly. Furthermore, the ease with which the contents of a banned site can be replicated and moved to new, unlisted sites makes keeping a list of prohibited sites up-to-date even more difficult if not impossible.[29]

In the wake of the striking down by the Supreme Court of the Communications Decency Act in the United States,[30] the State of Texas has passed a law requiring that all ISPs advise their clients of where they can obtain filtering software. Interestingly, most Texas ISPs appear to have complied by providing a page with links to the web sites of companies that produce filtering software, accompanied by a critique of filtering software pointing out some of the problems described above.

Another limitation of filtering software is that all of the automated filtering products currently available (other than labeling systems, which require human review and rating) use text-based criteria to determine which materials to block. As the volume of content on the Internet consisting of graphics, audio, and video rather than text increases, the technical challenge of screening that content becomes virtually insurmountable. Software programs capable of

---

[28] A series of critical essays on many of the leading filtering products may be found at www.peacefire.org. Similar critiques of other filtering products can be found at www.censorware.org.

[29] Note that text-string filtering can also be defeated by content providers who use unusual spellings of common offensive terms: for example 4Q, forkyou, four queue, etcetera for the common f---you.

[30] Reno, Attorney General Of The United States, et al. *v.* American Civil Liberties Union et al., Appeal from the United States District Court for the Eastern District of Pennsylvania, No. 96–511, Argued March 19, 1997—Decided June 26, 1997

analyzing video for semantic content, for example, are still in the early research stage. Even the less daunting task of analyzing the content of a single image is an interpretive challenge well beyond the capabilities of current filtering software and computer hardware. If the filtering of text is inaccurate, filtering images based on more amorphous criteria is several orders of magnitude more problematic. Schemes suggested, such as blocking images based on the percentage of "flesh-tones" contained in the image, hint at the problems with this approach. What are flesh-tones? Are classical paintings containing nudes blocked? What about close-up pictures of faces (mostly flesh, after all)? Baby pictures? Medical images?

Content labeling

Labeling or rating schemes for Internet content have been proposed and developed by groups such as the RSACi[31] and SafeSurf. These schemes generally use either a simple age-based rating scale similar to that used for movies (e.g., General, PG13, and so forth) or a more sophisticated labeling system that rates material based on a number of dimensions (e.g., sex, violence, hate, language, and so forth). The PICS (Platform for Internet Content Selection) labeling system developed by the W3 Organization supports labeling schemes of either type, and has generated substantial interest, particularly in some European countries. Web browser manufacturers have indicated that they will make their browsers PICS compliant, ultimately allowing users to screen content in web sites based on PICS criteria. The latest version of Microsoft's Internet Explorer web browser is PICS compliant, allowing parents to turn on content screening. For example, a particular user might decide to screen out sites that have a PICS-compliant "violence" rating above 2, "sex" rating above 3, "offensive language rating" above 4, and so forth.[32] These systems rely on a combination of humans, who review and rate the content, and technology, which blocks content on the basis of the human-generated ratings.[33]

Which Rating System?

Several of the manufacturers of current filtering software use criteria that can be recognized by the PICS system. Examples of the PICS-compliant rating criteria established by RSACi, SafeSurf, and Net Shepherd are attached as Appendix B. The criteria established by each of these organizations vary widely. For the sake of consistency, one set of rating criteria and a labeling system capable of recognizing those criteria would have to be adopted for Canada. In essence, this would mean granting a monopoly in the Canadian market to one technological approach.

---

[31] Recreational Software Advisory Council on the Internet

[32] For more information on PICS, visit http://www.w3.org/PICS/.

[33] Note that WC3 has not established rating criteria for use by PICS, and PICS is not a rating scheme. Rather, it is a technology platform that will support the implementation of a range of "PICS-compliant" rating schemes, allowing Internet content to be rated either by the content providers or by third-party rating agencies.

One of the criticisms of the American decision to legislatively require the V-CHIP's inclusion in all televisions sold in the US is that, by eliminating competition between producers of rival products, the US government has frozen technology development in this field. While Canadian standards would probably not have the same impact on global technology as do American ones, official adoption of a single technological solution would have a similar anti-competitive tendency to stifle improvement and development.

Choosing a rating system is complicated by the fact that none of the existing rating criteria are defined in such a way as to clearly distinguish the type of content that might be prohibited in Canada. While it might be possible to develop a set of rating criteria specifically geared to Canada, Internet content developers around the world are highly unlikely to affix Canada-specific rating labels to their sites.

Who does the labeling?

In theory, a labeling system such as PICS would provide far more useful results than blocking or content-filtering software. As mentioned above, however, a labeling system is not a purely technological approach to content filtering. These systems rely on either voluntary compliance (self-rating) by content creators, or rating and labeling by third parties. Labeling or rating by a large number of diverse groups and individuals would obviously result in inconsistency. Standards-based (or subjective) rating systems would result in the same item receiving different ratings from different groups or individuals. In his excellent essay "Rating the Net," Jonathan Weinberg, Associate Professor at the Wayne State University Law School, uses examples of the rating systems used by a number of filtering products to illustrate some of the problems associated with standards-based rating systems:

> **The SafeSurf questionnaire, for example, requires the self-rater to determine whether nudity is "artistic" (levels 1 through 6), "erotic" (level 7), "pornographic" (level 8), or "explicit and crude" pornographic (level 9). The voluntary Content Rating self-rating system promoted by CYBERSitter is almost the model of a standards-based regime: it offers as its only guidance the instructions that self- raters should determine whether their sites are "not suitable for children under the age of 13," and whether they include material "intended for an audience 18 years of age or older." Specs for Kids raters are instructed to distinguish between sites that: (1) refer to homosexuality "[i]mpartial[ly]"; (2) discuss it with "acceptance or approval"; or (3) "[a]ctive[ly] promot[e]" it or "attempt[] to recruit the viewer." Each of these classifications requires more judgment on the part of the evaluator, and is not so**

hard-edged as the RSACi categories. Individuals with different perspectives and values may disagree as to where the lines fall. With respect to the Specs treatment of references to homosexuality, individuals disagree as to whether the categories are even coherent. These categories work only within a community of shared values, so that evaluators can draw on the same norms and assumptions in applying the value judgments embedded in the standards. [34]

On the other hand, rules-based (or objective) rating systems tend to obscure the kind of information that is often important in deciding whether access to a site should be prohibited. For example, the RSACi defines "objective" rating categories providing 5 rating levels (from 0-5) in each of four categories: nudity, sex, language & violence. Level 4 in the "nudity" category is described as "frontal nudity". Using objective criteria, Michelangelo's David would garner a Level 4 nudity rating using the RSACi criteria.[35] Thus, while it may be possible to achieve a greater consistency across multiple reviewers using a rules-based rating system, such a system is unlikely to provide the kind of value-based information that would be most useful in making a decision to block certain content. For this reason, many content developers are opposed to a requirement for self-rating. To quote again from "Rating the Net":

When an author evaluates his site in order to gain a rating from any PICS-compliant rating service, he must follow the algorithms and rules of that service. Jonathan Wallace, thus, in an article called Why I Will Not Rate My Site, asks how he is to rate "An Auschwitz Alphabet," his powerful and deeply chilling work of reportage on the Holocaust. The work contains descriptions of violence done to camp inmates' sexual organs. A self-rating system, Wallace fears, would likely force him to choose between the unsatisfactory alternatives of labeling the work as suitable for all ages, on the one hand, or "lumping it together with the Hot Nude Women page" on the other.[36]

Furthermore, it is doubtful that one could rely on the purveyors of hate literature (for example) to accurately label their content. On the other hand, labeling or rating of all the content on the

---

[34] Jonathan Weinberg, "Rating the Net," 19 HASTINGS COMM/ENT L.J. 453 (1997). Versions of this article appear in INTERCONNECTION AND THE INTERNET 225 (Gregory L. Rosston & David Waterman eds. 1997), and THE V-CHIP DEBATE: LABELING AND RATING CONTENT FROM TELEVISION TO THE INTERNET (Monroe E. Price ed. 1998). The article can be found on-line at http://www.msen.com/~weinberg/rating.htm.

[35] At least if viewed from the front!

[36] ibid.

Internet by a single third-party organization, as discussed above, poses practically insurmountable logistical problems. The task of perusing and rating each item would require the massive army of reviewers described in the discussion of blacklisting, above.

Screening dynamic content

Web pages are only one of the methods by which information is exchanged on the Internet. If rating web pages presents a monumental challenge, other methods of exchange, such as real-time chats, email, discussion groups, and so forth, are simply not amenable to labeling. The subject matter of these is constantly changing and is therefore impossible to categorize. Rating a real-time chat (whether text-based or audio) would be equivalent to attempting to rate a telephone conference call while it is occurring. Rating email, which is generally private correspondence between individuals, would require the interception and reading of all of the millions of email messages exchanged over the Internet in Canada each day. Rating of discussion groups could be done on the basis of the group's stated subject matter. However, the actual contributions to any discussion group, no matter how innocuous its theme, may occasionally veer into unacceptable areas. It would not be possible to predict this in advance for the purposes of labeling.

Blocking of dynamic content using lists of prohibited text strings is possible. Certain filtering products will, for example, block a discussion group, terminate a chat session, or delete an email message if it contains prohibited text strings. The problem with text string blocking, however, as discussed above, is that it tends to block legitimate content at a rate far in excess of the offensive content it blocks.[37] Defining text strings that would accurately identify types of content prohibited in Canada is next to impossible.

What if the user doesn't want to filter?

Aside from problems with accuracy, client-side filtering software relies on the computer owner's willingness to install and use it. Client-side filtering software is relatively easily evaded or disabled by the determined user. A number of websites and newsgroups provide instructions on uninstalling, disabling, or evading blocking and filtering programs.[38]

If a government were determined to prevent the user from accessing unacceptable material, even if he or she wanted to access such material, it could mandate the installation of some sort of

---

[37] Angry users brought to AOL's attention the fact that, in defining the word "breast" as a prohibited text string, it had excluded the breast cancer survivor's discussion group.

[38] A determined user could also reformat or partition the computer's hard drive to get rid of filtering. Some of the many sites that include information on disabling filtering programs are: Peacefire Youth Alliance Against Internet Censorship (www.peacefire.org) , Glen L. Roberts' web site (www..glr.com/nurse/), and Full Disclosure (fulldisclosure.org)

filtering or blocking program on all computers sold in Canada.[39] This would not affect those computers already owned by Canadians, but would have some initial effect on new computers purchased in Canada. Since current filtering software will prevent Canadians from accessing perfectly legitimate material, however, the imposition of requiring such software on their systems would inevitably be unacceptable to many Canadians.

Furthermore, as mentioned above, filtering and blocking programs are not difficult to disable or evade. A user who purchased a computer with filtering software installed could simply access readily available instructions for uninstalling the program, download a "clean" browser from a site on-line, or, if necessary partition or reformat the computer's hard drive to render the filtering software non-functional.

Without major modifications to current Canadian legislation, little else can be done on the client side to unilaterally constrain users' access to material.

Imposing Client-Side Restrictions

Other countries,[40] which have regimes that permit far greater restrictions on the rights of individuals than is the case in Canada, have attempted to impose client-side restrictions by legislatively prohibiting individuals from accessing unacceptable material and imposing severe penalties on those who violate the prohibition. A number of difficulties arise in attempting to do this: How does one define "unacceptable content" so the user will know which material to avoid? How do users know that the material they access meets the criteria of unacceptability before it has been downloaded? How do those charged with enforcing this legislation know what the user has accessed and whether it is prohibited? Let us consider each of these questions.

How do we define "unacceptable content" so the user will know which material to avoid?

This is not a technical question, but a question of definition that is not unique to the Internet. Prohibited materials (e.g., child pornography, hate literature, defamatory material, and so forth) have been legally defined in Canada. With respect to any individual document, however, the problem of classification remains. Before attempting to implement any technological solution to notifying the user of "unacceptable content," it must be possible to unambiguously assert that an

[39] Which filtering software, and what it would be set to block, is another question.

[40] Singapore and China are the most frequently cited examples of countries with restrictive Internet regimes. Most totalitarian regimes, however, have restrictive Internet access regulations. Restrictions tend to be more effective in countries where state terror combines with poverty and a meagre communications infrastructure to create hurdles to access that are virtually insurmountable. Even Cuba, however, with one ISP, a single 64Kb connection to the Internet and only 600 authorized Internet users, has outlaw Internet users.

item falls into a class of prohibited material. As numerous court cases have demonstrated,[41] however, no simple set of objective criteria will serve this purpose. Ultimately only a court can determine whether a particular item falls into a class of prohibited material.

How does the user know the material meets the criteria of unacceptability before it has been downloaded?

A comprehensive and consistent labeling system is the only way in which the user can know what the contents of an item are prior to downloading. As suggested above, there is no reliable set of objective criteria that can be used to identify prohibited types of content, nor is comprehensive and consistent labeling of all Internet content achievable. Consequently, partial- or full-text filtering and labeling will not serve to let the user know, prior to downloading, whether material requested on the Internet meets the criteria of unacceptability. It is only after having downloaded and viewed the material that the user would be in a position to assess whether it constituted "unacceptable content."[42] (At which point, the question would be moot.)

The user could know if specific content is prohibited if the definition of "unacceptable material" is limited to "those items included on a government-created list of prohibited files or sites." Such a list could be developed to work with a range of client-side filtering products, and the obligation to use a filtering product imposed on the user. There would remain, however, the problem of measuring compliance.

How do those charged with enforcing this legislation know what the user has accessed and whether it is prohibited?

Since most Internet usage occurs in the privacy of users' homes or offices, determining what a user has viewed is not an easy matter. Clearly, any attempts by the government to monitor what people are doing in their own homes or offices would raise a multitude of issues related to invasion of privacy and would contravene the CSA guidelines for privacy.[43] From a technological perspective, there is no practical client-side approach to monitoring what people are accessing on the Internet. A somewhat "diabolical" scheme for monitoring, suggested tongue-in-cheek by Prof. John Carrey of Columbia University, is to install "cookies" on the user's machine to track and report what the user accesses on the Internet. A "cookie" is a small text string delivered to the user's computer along with a web page. It records specified information and provides that information to the server on which the page originated when the user visits that page again. It is

[41] For example, Little Sisters et al. v. the Minister of Justice and Attorney General for Canada et al., BCCA, 1996

[42] Although the definitional problem mentioned in the preceding paragraph would still exist.

possible to conceive of a cookie that would send a copy of every link a user visits to a central server. [44] While ingenious, the problem with cookies is that first the user has to access a web page that will deliver the cookie, and then the user has to set the browser software to agree to accept the cookie. The ability to decline cookies is an integral part of the software on all web browsers that are capable of accepting cookies. This ability is an essential security feature and should not be removed, as, in its absence, malicious individuals could use cookies to install viruses, capture user passwords, and carry out other undesirable activities. Even after they have been accepted, cookies can be disabled or deleted from the user's system. Aside from the questionable efficacy of such an approach, one cannot imagine Canadians accepting the concept of their Internet activity being monitored by a government controlled enforcement agency.

Another approach to monitoring would be to adopt a requirement that all Canadian ISPs maintain logs of all URL requests made by their clients.[45] This would impose a cost burden on the ISP. As well, individuals who wished to avoid having their requests logged in this way could use one of the many on-line services that allow a user to request Internet content anonymously.[46] One imagines that if logging of user activity were introduced in Canada, the number of such services, and the number of Canadians using them, would proliferate. One could also expect the number of Canadians obtaining Internet services directly from US ISPs to increase significantly.

Hardware-based Restrictions

If software approaches cannot compel compliance on the client-side, what about hardware approaches? Is it possible to impose a requirement for the installation of the computer equivalent of a compulsory V-CHIP into Canadian computers? The V-CHIP is a computer chip that filters television content. It operates on principles very similar to those of ratings-based software filters. This capability is hardwired into the television set, since, unlike a computer, there is currently no simple way of installing programs on a television. When installed in a television set, the V-CHIP allows the TV owner to block reception of television programs on the basis of rating criteria embedded in the television broadcast. Parents, for example, may program their television sets to prevent their children from watching programs with an adult rating. The FCC in the United States, anticipating that larger computer monitors will prompt more people to watch TV on their computers, is considering requiring computer manufacturers to install V-CHIPs into

---

[43] These guidelines are the basis of the federal government's proposed privacy legislation (Bill C-54). Such a move would run contrary to recent government measures to promote privacy and security on the Internet.

[44] Microsoft's Internet Explorer™ "channels," although not technically a "cookie," works along these lines.

[45] China and Singapore have regulations to this effect.

[46] For example, the Anonymizer – www.anonymizer.com . Anonymous remailers and web sites are discussed in some more detail below in connection with ways of avoiding server-side filtering.

computers.[47] In theory, such a chip could be made to detect labels on web pages, as well as the ratings encoded in TV broadcasts: in essence, hardwired filtering software. If a comprehensive and consistent labeling system for Internet content existed, such a chip could theoretically be set to screen out content deemed unacceptable. However, the installation of a hardwired filtering program would do nothing to overcome all the obstacles to the implementation of a coherent and effective labeling system discussed above. Nor, although it might be more difficult to uninstall than a software-based filtering program, would it prevent the use of evasion techniques such as anonymous remailers.

[47] Whether this requirement will actually be imposed is open to question. The move is being vigorously opposed in many quarters. Computer manufacturers, especially small computer manufacturers, argue that the expense will make their businesses uneconomic. Others argue that the number of people who will actually want to watch TV on their computers is small. And the Civil Liberties Union opposes the proposal as a move towards controlling the types of content available on the Internet.

1.2. SERVER-SIDE RESTRICTIONS

If client-side restrictions take one into the realm of the impractical, what about implementing server-side restrictions?

Filtering at the Host Level

For content created in Canada, the government has a number of existing remedies against individuals who contravene Canadian law. If the host server is located in Canada, it would theoretically be possible for the government to require by law that the operator of the server install and run filtering software that would prevent "unacceptable" content from being transferred to users. However, the problems with consistency and accuracy of labeling and filtering systems discussed above apply to this technology as well.

Filtering at the ISP Level

If the host server is located outside of the country and, therefore, outside of Canadian jurisdiction, the government would not have the authority to impose any requirements on the originating server. To implement server-side restrictions, it would be necessary to interpose proxy servers or firewalls[48], through which all data entering the country would have to pass and be inspected before reaching the user. Because the Internet is a fully-meshed, self-repairing network, there are many different available routes between a source server and the end user. As the Internet is currently configured in Canada, the only intermediate network node in Canada through which the data absolutely must pass on its way to the user is the router at the user's ISP site. (This assumes that the user has a Canadian ISP. If the user accesses the Internet directly through a foreign ISP, there are no intermediate network nodes within Canada through which the data absolutely must pass on its way to the user.)

Theoretically, each of the ISPs in Canada could be required to install a firewall that would screen clients' file requests and refuse to forward requests for items retrieved from a list of prohibited sites, and/or set up proxy servers to cache and screen content before passing approved content to the client. As mentioned in Part 1, ISPs in Canada are currently unregulated. Anyone can

[48] In this context, a proxy server is a server on which incoming Internet content is cached (stored) before being forwarded to the client. A firewall is a server that enforces network access or security policy. High-end firewalls generally run on dedicated hardware devices.

become an ISP. Because of the absence of regulation and the low barriers to entry, competition and expansion in the ISP sector is vigorous. Introduction of ISP-level firewall or proxy server requirements would be a significant negative change to the economic health of this industry. [49]

The introduction of a firewall impacts the throughput of the network by decreasing the number of bits per second passing through the router and increasing the latency (delays) in the network. A study of five leading firewall products (most of them blackbox hardware solutions) has indicated that the introduction of the firewall, even without any filtering rules, caused latency to increase approximately arithmetically as the number of clients increased, while Mbps peaked at 170 Mbps per firewall device.[50] Another test of firewall software revealed that the introduction of a single filtering rule[51] decreased throughput by 20 per cent with 16 clients, and by 40 per cent with 64 clients.[52] The types of filtering rules described here are of the type commonly used in corporate firewalls. That is, they screen data packets based on discrete character strings in pre-defined locations (for example, starting at bit 12) in the packet. High end, dedicated routers can do this type of screening at current line speeds for a relatively small number of rules, provided they do not have to deal with complex routing tables at the same time. Screening based on character strings located at arbitrary locations in the data packet, on the other hand, cannot be done, even by the most efficient dedicated hardware, at anywhere close to line speeds. What this means is that blocking of content on the basis of pre-defined criteria (such as the IP address) which are located at predictable locations in the data packet, is possible at line speeds, given sufficient routers located at the edges of the network, but not at core routers (for example, the core routers at medium to large ISPs). Screening of content on the basis of character strings located at arbitrary locations in the data packet is not currently possible at line speeds.

These results demonstrate that, in order to minimize performance impacts created by the introduction of a firewall it is necessary to limit the number of clients accessing the network through any single firewall device. To provide throughput rates that would support streaming video and other high performance Internet facilities requiring high network speeds, one would

---

[49] The Canadian Association of Internet Providers (CAIP) takes the position that ISPs cannot be responsible for content on their systems generated by their clients. They argue that their position is essentially that of a common carrier, and that that they cannot be held responsible for what is posted by an individual using their service. Courts in the US have adopted a similar view, as long as the ISP refrains from taking any role in choosing the content on their systems.

[50] These figures are based on a report by KeyLabs, an independent US lab specializing in software and hardware testing in a networked environment. The full report can be found at http://www.keylabs.com/results/firebench/index.html. These dedicated filtering solutions are more efficient than purely software-based filtering products that run on top of a standard server operating system.

[51] A "rule" is a criterion for accepting or rejecting a file or connection. For example, a common rule for corporate firewalls, is: do not accept requests from people without passwords. Another might be: do not accept executable files (files with a .exe ending).

have to provide an additional firewall device each time the introduction of further clients would perceptibly impact on throughput.

Based on its March 1998 survey, ACNielsen states:

> **The Canadian Internet user community continues to grow -- 37% of Canadians aged 12 and over are now on-line users. This is up from a penetration level of 31% a year ago, and represents a growth rate of 20%.**
>
> **The total of 37% represents approximately 9.5 million Canadians aged 12 and over. Recognizing that there are a number of Internet users aged 11 and under, the overall total of Internet-using Canadians will be a somewhat higher number, perhaps closer to 11 million.[53]**

Since usage has undoubtedly increased since March 1998, we can use 11 million as a conservative estimate of Canadian Internet users.[54] If we conservatively assume that during peak periods at least 10 per cent of those users are on-line concurrently, one would require multiple firewall devices for even a small ISP to permit minimal screening without negatively impacting network performance beyond acceptable levels. With the number of Canadian Internet users increasing rapidly, and traffic on the Canadian Internet doubling on average every four to six months[55], one would expect the required number of firewall devices to increase concurrently. High-end firewall devices, capable of handling the level of traffic described, typically cost in the range of $25,000. The cost of purchasing and installing enough such devices to screen all the content passing through ISPs in Canada would be billions of dollars. (Normal procedure is to schedule these devices for replacement every four years.) This, of course, does not include any amount for operations, maintenance, or housing of the devices. Nor does it take into account the increase in these costs as Internet usage increases. For small ISPs this expense would easily make business unprofitable.[56]

---

[52] These results come from tests conducted by KeyLabs on software-based firewall products. The results are reproduced at 222.ntguard.com/performance.html.

[53] *Top Line Results*, from ACNielsen Measures the Net: The ACNielsen Canadian Internet Survey '98, http://www.acnielsen.ca/ACNielsen/cgi-bin/DisplayPage?SITE=ACNielsen&KEY=survey98spring&TRACKID=MC_.

[54] Note that the volume of traffic on the Internet is increasing at an even greater rate than the number of users.

[55] See Part 1

[56] 1997 figures from BITS Information Service survey indicated that in 1996 over 50% of the ISPs in Canada had revenues under $500,000, with 20% having revenues under $250,000, and 13% having revenues less than $100,000. From the Canadian Association of Internet Providers' web site: http://www.caip.ca/corpinfo.htm. Additional routers and maintenance could easily cost the small ISP between $100,000 and $200,000 annually.

The hardware and software cost estimates above assume a minimal set of filtering rules. The longer the set of rules, and the greater the volume of data processed, the more demands are placed on the firewall, and the greater the impact on network performance. Long before the extreme of full-text scanning or pixel analysis of each document is reached, the quality of network service would have deteriorated to the point of making the Internet unusable. Furthermore, the hardware, software, and maintenance costs to the ISP of implementing such a scheme would make the business uneconomic, driving small ISPs out of business and large ISPs out of the country. [57]

Standard firewalls devices are specialized routers that filter data packets based on simple criteria that can be determined from the identifying information encoded in fixed locations in each packet (e.g., IP address, file type, and URL). Partial- or full-text filtering of files would require the installation of proxy servers to enable reassembly of the data packets into complete files for scanning. Imagine that data moves on the Internet like trains on a track. When a train reaches the switchyard, the appropriate switch is thrown and the train moves on towards its destination without stopping. With enough tracks and switches, a large number of trains can pass through the switchyard without stopping. This is analogous to the way in which the Internet currently works in Canada. If the trains have to stop at the switchyard and have their contents examined before proceeding, however, new facilities for the storage of trains, and new staff for the inspection of trains, would have to be added. This both adds to the cost of the facilities and increases the time it takes for the trains to reach their destination. If the switchyard wishes to minimize the delays experienced by the trains, it will require new facilities large enough to accommodate all the trains that might arrive at one time, and sufficient staff to make an immediate examination of all the trains that arrive. The same argument applies to the interposition of partial- or full-text filtering requirements on Internet content, with the additional twist that all the pieces of data that make up a file are not on the same "train". Each packet of information has to be stored until all the other packets that make up the file arrive. Consequently, partial- or full-text filtering would require massively larger storage capacity as compared to simple filtering of data packets based on originating address. Adding the task of screening and filtering content implies not only the addition of massive amounts of storage capacity to firewalls, but also a reduction in the performance of the network. To minimize the reduction in performance of the network, devices capable of storing all the data that arrives in any given time period would be required. These devices would need to run screening software which would immediately commence scanning each file once all of its data packets have arrived. They would need sufficient processing power and RAM to run all this software quickly and efficiently. If introducing simple

[57] Several presenters to the CRTC New Media Hearings in 1998 made the point that economic barriers to operation in Canada would prompt them to relocate their businesses in the United States. Unlike industrial enterprises, ISPs can relocate easily.

packet-based firewall filtering across the country would cost in the billions of dollars, introducing partial- or full-text scanning would be many orders of magnitude more expensive.

Technically, filtering at the ISP level could probably be implemented (given a restrictive ISP licensing and regulatory regime). Supporters of such a policy often point to Singapore, which has a requirement for ISP-level filtering based on a list of prohibited sites. While Singapore currently has only one backbone connection to the outside world and three ISPs, however, Canada has many international connections and over a thousand ISPs. The cost of implementing content controls at the ISP level in Canada would be enormous and the degradation of network service would be significant (with consequent economic costs), while the filtering, as discussed earlier, even with the most comprehensive set of rules, would be ineffective, blocking a large percentage of inoffensive material and allowing objectionable material through.

Filtering at the Backbone Node or Border-crossing Level

An alternative to filtering at the ISP level would be to install giant filtering facilities at all the points where the Internet backbone enters Canada. Acting as border-crossing checkpoints for data, these facilities would operate like massive versions of the ISP level firewalls or proxy servers described above. These facilities would be subject to the same problems of inaccuracy inherent in any filtering scheme, and would have the additional problem of having to handle huge volumes of data.

Since the costs of caching and filtering increase exponentially as the volume of data increases, the costs of these facilities would be prohibitive. Current backbone routing devices, which have the relatively simple task of reading the destination address of a data packet and routing it onward, are already challenged by the volume of data they have to handle. Introduction of even simple firewall filtering criteria at the backbone node or border-crossing level would have disastrous effects on network performance. The activities underway in Canada to build a very high performance Internet for economic and social benefit would be negated by these bottlenecks.

In addition, filtering at either the local ISP level, the backbone router level, or the border-crossing level would still not catch people who connect directly (either via landline or satellite[58]) to an ISP outside the country. If restrictions in Canada increased, we could expect the numbers of Canadians obtaining Internet access outside of the country to grow. If the government wished to prevent this, it would be necessary to outlaw data satellite receivers, although experience with

[58] Currently, two-way data satellite service is not available. Telesat's DirectPC service provides a digital downlink (data delivery) with a telephone back channel. The economics of two-way data satellite service are not yet clear. However, one should anticipate in the near future either two-way satellite service or forms of delivery that do not use traditional landlines.

television satellite receivers has taught us that these types of prohibitions are not very effective. Private cross-border landlines, such as those operated by a number of Canadian corporations and which are an integral part of their business operations, would also have to be prohibited. Similarly, one would have to prohibit any kind of fixed connection (e.g., DSL, ISDN, or cable) to non-Canadian servers. One would also have to take measures to ensure that users could not dial directly to an ISP outside of the country. (Note that the advent of flat rate North American long distance calling makes it economical for someone in Canada to place a phone call to any American ISP.) How could direct dial access to non-Canadian ISPs be restricted? In theory, one could require that modems sold in the country not be able to make long distance telephone calls, and prohibit the importation of modems from other sources. Practically, of course, manufacturers are unlikely to accept a requirement to produce special versions of their modems for the Canadian market. Many computers now come with built-in modems, and, again, requiring the manufacturers to produce special versions of these computers for the Canadian market is unlikely to be accepted either by the manufacturers or by the market. Another alternative would be for the government to prevent long distance data calls over telephone lines by requiring the telecommunications provider to "listen" to the beginning of every long distance telephone message to determine if it is a data call and not allowing the connection if it were. (Although the authors have been asked to restrict their comments to technical feasibility, we feel compelled to point out that such an approach would, almost without a doubt, constitute an invasion of privacy and a violation of individual rights. It would also constitute a serious impediment to businesses that conduct business across international jurisdictions, and make Canada a pariah state in the information society. We could not expect other democracies to respect these prohibitions. In fact, we could expect them to actively condemn and try to defeat them. We could also expect that corporations in other countries, especially those in the high-tech and information sectors, would consider Canada an unfavourable country in which to do business. Existing businesses would leave Canada and new ones would invest elsewhere.)

As mentioned in Part 1, Singapore, which currently has only one backbone connection to the Internet and three ISPs, has attempted to restrict its citizens' Internet access. To accomplish this, it has used a combination of proxy serving, blocking of identified "bad" sites, licensing of modems, logging of client activity, and random checks of which sites individuals access, along with the threat of serious penalties (jail terms and whippings) for people caught accessing unacceptable sites. (Since everything, in theory, goes through government licensed proxy servers, logging makes it possible for the authorities to monitor who is viewing what.) The challenge of scaling such an approach to Canada, with its multiple backbones and over one thousand ISPs, would be staggering, both in terms of technical requirements and cost. Yet even in Singapore, although fear of penalties (more than technical efficacy) keeps a large percentage of the

population in line with government restrictions, users still regularly access Internet content prohibited by the state.

The reasons it is not possible to completely prevent users in Singapore from accessing unacceptable content are the same reasons screening of content at the ISP, backbone node, or border-crossing level in Canada would be ineffective. These include:

- the impossibility of maintaining a comprehensive list of prohibited content. As discussed above, the sheer volume of content on the Internet makes the creation and maintenance of a comprehensive and timely list of banned sites impossible. The "best" that can be accomplished, as has been shown in Singapore, is to block the most obvious sites;

- the use of multiple redundant proxy servers and IP address rotation by content providers to defeat blocking. (If a URL or site is put on the banned list, the content is simply moved to another address);

- the ability of users to request and receive web pages as email attachments or in encrypted form that defies filtering. A number of on-line services exist that allow a user to request a web page via email. Others, such as the Anonymizer, allow a user to anonymously request a web page by adding the URL of the web page to the URL of the Anonymizer web site. The Anonymizer then obtains and forwards the page to the user with an unidentifiable URL attached. Services such as the Anonymizer were developed to defeat logging software (software that records the sites visited by a user). However, they can also be used to circumvent filtering software. While it might be possible for the government to block user requests sent to identified anonymous servers (Anonymizer, for example, is blocked by many of content filtering products), new ones will spring up to take their place.

If the Government of Canada started to regulate access in such a way that it offended our American and other international neighbours (as happened recently when the German government required Compuserve to block some information), one would expect many sites to open up that would accept encrypted URLs and return encrypted web pages. It would be virtually impossible to detect anything other than the server's name on the request side and would also be equally impossible to examine the content delivered. Web sites offering this service could use essentially the same technology as that used by secure servers (such as those used for on-line banking and e-commerce) to prevent the theft of information as it is transferred from the client to the server or from the server to the client. From the client's point-of-view, this process is entirely transparent. In an on-line banking transaction, for example, the user can be completely unaware of the encryption process. Similarly, the user taking advantage of a Web encryption service would be able to simply type in the URL of the desired web site, and the encryption and

transmission process would occur without the need for any further action on the part of the user. Alternatively, the user could use a search engine that would encrypt web page requests (made by the user clicking on a link on the search page) and return the requested page in encrypted form. No special expertise on the part of the user would be required.

To prevent users from requesting encrypted content, the government could attempt to block all encrypted web browsing. If enough sites offered this service (especially the portal sites), then much useful content would be blocked. In addition, blocking encrypted transmissions would make e-commerce impossible, with serious negative impacts on the nation's economy. As mentioned previously, such a measure would also run counter to the steps the government has recently taken to promote privacy and security on the Internet in order to promote e-commerce and the use of the Internet in health and education.

2. PROMOTING ACCESS TO SPECIFIED TYPES OF CONTENT

Let us now consider the possibility of promoting the access of specific types of content by Canadian Internet users. Is it technologically possible to ensure that Canadians access certain types of content when they browse the Internet?

2.1 PORTAL SITES

A growing number of Canadians access the Internet via ISPs that operate what are known as "portal sites". These sites (for example, Sympatico and @home) offer a range of services and information that their operators believe their clients will find useful. Portal site offerings commonly include such things as local movie and concert listings and reviews, restaurant listings, TV schedules, weather forecasts, traffic reports, and so forth. (Many companies such as Yahoo!, AltaVista, and Excite, Inc., that started out operating Internet search engines are now running portal sites. Other companies such as America OnLine, Netscape and Microsoft, through its Microsoft Network, are also running portal sites. The Canadian CANOE™ site is an example of a site intended from its inception as a portal site.) To explore how one might promote certain types of content to Canadians, let us consider the option of the government requiring all Canadian portal site operators to ensure that a specified amount[59] of the content at their web sites consists of that class of content.

When a user subscribes to an ISP, it is common practice for the ISP to provide the user with web browser software pre-configured to use the ISP's portal site as the user's homepage.[60] The user has the option, of course, of specifying a different homepage (or no homepage at all), but many users retain the ISP's portal site as their homepage. If the portal site contains a specified amount of a designated type of content, users who retain the portal site as their homepage will receive at least some exposure to this content. Of course, as soon as users leave the portal site, there is no way of saying what they will access. Furthermore, if users find that the content and services at the portal site are not of interest to them, many will soon change their homepages to sites that they find more useful or interesting. If they find they are losing visitors to their portal sites because of the type of content they are legislatively required to include, ISP's are likely to do one of two things: stop providing a portal service, as it is no longer economically viable, or move their portal service to the United States, where they will not face content legislation.

---

[59] The question of how one determines "how much" of a web site consists of designated content is not straightforward, and will be discussed below.

[60] The homepage is the web site which the web browser software accesses by default whenever the browser is launched. Not all ISPs operate portal sites, but those which do generally pre-configure their clients' browser software in this way.

For the purpose of exploring the question of whether it is technologically possible to ensure that portal service operators include a specified quantity of a designated type of content at their sites, we will use "Canadian Content" as an example. Before determining whether it is possible to technologically ensure that portal service operators meet a requirement that a specified percentage of their sites consist of "Canadian Content," two fundamental questions must be answered. What is "Canadian Content" in the context of the Internet? And, how does one measure the "amount" of "Canadian Content" on a site?

What is "Canadian Content" in the context of the Internet?

Determining what is "Canadian Content" in the context of the Internet is not a straightforward matter. Not only is one faced with the question of what makes content "Canadian," one must also decide what constitutes "content." Is a hyperlink to a Canadian web site "Canadian Content"? A banner ad for a Canadian product? A chat room where some of the participants are Canadian? A discussion group dealing with Canadian history, regardless of who participates? A site designed and created by Canadians even if the subject matter is non-Canadian? Any site located on a server in Canada, regardless of content? Without going into all the possible scenarios, suffice it to say that the Internet introduces some interesting twists to the concept of "Canadian Content." To attempt to answer this question is beyond the scope of this report addressing technical feasibility. Obviously, however, before considering any technological approach to ensuring "Canadian Content" (or any other designated type of content) is available to Canadians, a clear and unambiguous means of distinguishing that content must exist.

How does one measure the "amount" of designated content on a site?

Assuming we are able to resolve the definitional question, the problem of measurement remains. By what criteria does one determine how much of a site consists of a designated type of content? Does one, for example, add up the total number of bytes at the site, and determine how many of those bytes are "Canadian Content"? Due to the fact that graphics, audio, and video consume progressively greater amounts of storage space than text,[61] adopting this approach would result in small graphics, or tiny video clips, counting as the equivalent of extensive text passages. (For example, a 1500 word article might consume 15Kb – 15,000 bytes - of storage space, about the same as one small JPEG image, four seconds of audio, or one second of video.) Another approach might be to count up the number of items at the site and determine how many of those

[61] 1 megabyte of storage space will hold approximately 100,000 words of text, 4 minutes of audio, and one minute of video. Note that these figures depend on the type of compression and the frame rate used. Higher quality audio and video will consume an exponentially larger number of bytes.

items are defined as "Canadian Content"? This approach raises the question of what constitutes an "item." If a list of web links appears on a page, is each link a separate item, does the entire list constitute a single item, or is a link not an item at all since it is only a pointer to other content? Another approach might be to simply count up the number of pixels on the screen and determine what percentage of the display area is consumed by "Canadian Content." With this approach, a single line of large-font "Canadian Content" text that happened to measure 320x240 pixels would count for four times as much as a 160x120 pixel icon that would launch a half-hour Canadian Internet radio broadcast. A fourth approach would be to try to determine how long the average user would take to read, view, or listen to a particular item compared to how long that user would take to read, view, or listen to all the content on the web site. If this approach is used, a text article that would take an average reader ten minutes to read would be equivalent to a ten minute audio or video clip. However, this article would be worth more on the "Canadian Content" scale than most graphic images (which the average viewer might look at for ten-to-twenty seconds) or web links (each of which might take a second or two to peruse). While, on the surface, this approach might appear to be the most equitable, one can envision a scenario in which a web site would be rated as high in "Canadian Content" only because the text describing the site and its creators takes twenty minutes to read, while the real "guts" of the site consists of hyperlinks to non-"Canadian Content" sites. One can also imagine the converse, where all of the hyperlinks are to "Canadian Content," but the text describing the site and its creators counts as non-"Canadian Content" and makes the site non-"Canadian."

Since each portal site would provide links to many other web sites, the problem of identifying and measuring "Canadian content" would not be limited to portal sites alone. Beyond the technical measurement issues dealt with above, any human intervention in measurement would give rise to the cost implications discussed at page 34.

Having discussed the difficulties of definition and measurement, the authors would like to point out that Canadian portal sites already, for purely economic reasons, provide their visitors with significant quantities of what would probably be classified as "Canadian Content" under any scheme. Local and regional portal sites have identified a Canadian market demand for the aggregation of content of local interest. To the degree that this business model continues to prevail, portal sites would probably not find a requirement for "Canadian Content" onerous, in the abstract. In practice, having to demonstrate compliance with government requirements in light of the difficulties of definition and measurement would undoubtedly be seen as problematic. Furthermore, as suggested above, should the market demand shift, causing Canadian portal sites to lose visitors, we could expect to see Canadian portal sites either shut down or move to the United States if Canadian requirements prohibited them from changing their offerings in order to keep customers.

Creating a Canadian Portal Service

An option that uses technology (although it is not entirely a technological solution) to advance certain types of content is the development of a dynamic, engaging portal service that will showcase the desired content. Portal sites, as mentioned above, are growing in popularity. If the Canadian government were to support through funding and other incentives the development and maintenance of "the best portal site in the world," many Canadians would undoubtedly use it. By supporting a site combining the best of design, services, and functionality, the government would create a platform through which preferred types of content could be brought to the attention of Canadians. A good example of this approach is the Industry Canada Strategis site, which has been very successful in offering information needed by the Canadian public and business. If the desired type of content is already being offered by other Canadian portal services, however, the government might not want, or deem it necessary, to offer a competing service.

2.2. SEARCH ENGINES

Of course, not all Canadians use portal sites. And even among those who do, the majority will eventually seek out other content on the Internet. Search engines are the most common software tool used by people to locate content on the Internet. The names of search engines such as AltaVista™, Lycos™, Yahoo!™, and Excite™, will be familiar to most Internet users. A user can type a keyword or series of keywords into the search engine, and obtain a list of links to web sites, newsgroups, and ftp sites in which that word or words can be found. If one wished to promote access to certain types of content (e.g, "Canadian Content"), one might consider deploying a search engine that would prioritize that type of content. Of course, a wide variety of search engines is available on-line, and people would only use this specialized search engine if it served their needs better than others that can be readily accessed.

Canadian Versions of Search Engines

There are Canadian versions of some of the major search engines (for example, Yahoo! Canada and AltaVista Canada). Yahoo!, which indexes far fewer sites than AltaVista, uses a team of reviewers to categorize sites. AltaVista Canada, launched in Canada by AltaVista and Telus Corporation in January, 1998, has developed an "intelligent crawler," software that can identify and index web sites located in Canada. A "crawler" is a software program that visits sites on the Internet, indexes the contents of each document it finds at the site, and adds the index and the URL for each document to a database. By sending out crawlers, search engines are able to

continually update their databases without requiring human intervention. AltaVista Canada was unwilling to disclose proprietary information on exactly how its "intelligent crawler" (TAZ) works. However, one possible approach to automating the process of identifying the geographic location of a site is to write a software program that compares the site's IP address to a concordance of addresses that fall within specified geographic boundaries.[62] Such a process would be able to identify most sites located in Canada. This process would not be one hundred per cent accurate if used to identify sites containing "Canadian Content." It would falsely identify as "Canadian" non-Canadian sites residing on servers in Canada, and it would fail to identify sites containing Canadian content but located outside of Canada (or using IP addresses allocated to a non-Canadian regional provider), unless those sites used the .ca top level domain designation (which could be used as a secondary criterion for identifying Canadian sites).[63] However, it would undoubtedly capture the majority of Canadian sites.

By default, the Canadian version of the AltaVista search engine only returns links for sites that both match the search terms and have been identified as Canadian. The user may, however, choose to broaden the search to the entire world. (Conversely, Yahoo Canada searches its entire index by default, but will limit its search to Canadian sites if the user makes that choice.)

Prioritizing "Desirable Content" in Standard Search Engines

In the last year, both Netscape and Microsoft have built search engine capabilities directly into their web browsers. This allows a user to type a word or a string of words as a search argument into the location field of the browser and obtain a list of sites relevant to the search argument. In addition to integrated search capabilities, Netscape has been working on a feature that it calls "smart browsing." This feature takes a single word typed into the address/location bar of the web browser and converts that to a URL (generally by adding "http://www" before the word and ".com" after the word). Netscape uses a list of "reserve words" to direct users to web sites of organizations associated with those words (for example, the word "tide" might direct the user to the Tide™ web site). An interesting planned feature of the reserve word system is the use of "international reserve words": that is, the same word can be associated with different sites depending on the location of the user. For example, a US-based user entering the word "ford" into the location bar would be taken to the Ford US site, while a Canadian-based user entering

[62] A portion of the IP address can be used to identify the geographic location of the machine's ISP. ISPs are allocated the addresses by a regional provider which has been assigned large blocks of addresses by the InterNIC. As the name implies, a regional provider operates within a designated geographic area.

[63] The majority of sites in Canada do not use the .ca top level domain. Most use one of the generic top level domains such as .com, .net, and .org. Figures from the OECD working document entitled "Internet Traffic Exchange: Developments & Policy" from January of 1998 indicate that Canadian sites are second after the US in the use of generic top level domains.

the same word into the location bar might be taken to the Ford Canada site. The reserve word system relies on the site owners (or some other entity) providing a list of reserve words to Netscape.[64] With generic words (that is, for words that have not been reserved by companies or organizations to identify their web sites), "Smart Browsing" would return a search page with a list of links sorted by likely relevance to the original word. It would theoretically be possible to use a variant of the "international reserve word" system to prioritize such a list, having Canadian sites show up first for Canadian users. However, the imposition of such a system without making it optional for users could be problematic. For example, recently some controversy ensued when it appeared Netscape was prioritizing pages located on its own server in the search results it was returning. Users most commonly desire the "best matches" for a search term to be listed first, rather than links chosen according to some other criteria.

For such a system to be effective technically in prioritizing Canadian sites for Canadian users, three things would be necessary. First, there would have to be an effective way of determining in which country the user is located. Second, there would have to be a way for the search engine to identify sites by their country of origin. Third, companies operating search engines would have to agree to implement this system in their software.

Determining the location of the user

It may be possible to determine the location of the computer a person is using to access the Internet by determining which regional provider has been allocated the computer's IP address. Since such organizations are relatively few, and have been allocated specific geographic areas, a concordance allowing the search engine to match IP addresses with regions could be developed. There is, however, the problem of keeping such a concordance current as new blocks of IP numbers are allocated.

Determining the nationality of the site

If the "nationality" of a site is determined by its being housed on a server located in a particular country, it may, as discussed above, be possible to identify "Canadian" sites using IP addresses, using the same methodology suggested for determining the location of a user. Sites could also be designated as "Canadian" sites if they used the Canadian top level domain (.ca). This latter approach on its own, however, would fail to capture the majority of Canadian sites, which tend to use generic domains such as .com and .org.[65]

---

[64] Part of the emerging business model for search engine operators is the generation of revenue through the sale of "reserve words".

[65] See note 63.

Such a system assumes the designated type of content (e.g., Canadian Content) can be ascertained using an existing unambiguous identifier attached to the file (e.g., the IP address of the server). If one were trying to promote certain content that could only be identified through semantic analysis of the file, this type of system would be unworkable. Identifying "educational content," for example, or any other measure of the semantic content resident at a site, cannot be done solely by analysing the IP address or URL of a site.

Implementing the system

Many companies provide users with free access to their web search services. Each of these search services offers the ability to search the contents of the Internet, but each has a slightly different approach, functionality, and style. The appeal of one search engine might be its comprehensive indexing, while the appeal of another might be its organization of content into searchable categories, and the appeal of a third the ease with which users can define what they are looking for. Companies that provide search services might implement a system to prioritize content by the user's country of origin if they thought this feature would appeal to users and increase their market share. Implementation of such a system would require a balancing of regional prioritization against accuracy of results. Design decisions would have to be made: do you group all results by level of accuracy and then prioritize by region within those groups? Or do you list all regional sites first, no matter how tenuous their connection to the search word(s), before listing any sites from other locations? While we have not conducted market research on the question, common sense tells us that whether the user wants the search engine to prioritize Canadian sites probably depends on what he or she is searching for. If the user were trying to compare car prices, he or she would probably prefer to see local listings first. Users researching diseases of tropical fish would probably want the closest matches listed first, regardless of where the content is located. Given this, search engine companies would likely make regional prioritization optional for the client.

2.3 PUSH TECHNOLOGY

Push technologies are technologies designed to send content to the client without the client specifically requesting it at a certain point in time. Television and radio broadcasting are classic examples of push technology. Traditionally, the World Wide Web has been based on "pull" technologies; the user seeks out the content he or she wants and downloads it to the client machine. In the Internet context, in order for a push technology to work, the user must have a "push client" (software designed to receive and display new content) installed on his or her computer. The oldest form of Internet push technology is email. Users with email software installed on their computers, open their email, and receive whatever other people have chosen to send them (as many email users who find themselves on junkmail lists discover to their annoyance). Push technologies can range in their degree of "pushiness" from simple notifications that new content is available (whether via email or other means) to automated content delivery. In every instance, however, the user will only receive content if the appropriate push client has been installed on his or her computer.

A few years ago, push technology generated substantial interest among Internet users and developers. Push clients such as Pointcast™, which would start automatically any time the user's computer was turned on, were launched with great fanfare. Anticipating a demand for push technology, Microsoft used a broadcast or "webcast" metaphor in developing Active Channels™, a push client/server technology that is integrated into the Microsoft Internet Explorer™ web browser and later versions of the Windows desktop. Netscape, in turn, developed Netcaster™, a channels-based push client for Netscape's Navigator™ web browser. However, push technology has proved far less popular than anticipated. There has been unexpected resistance by Internet users to push-technology. Because of this resistance, Microsoft has made the enabling of Active Channels (the push technology built into Internet Explorer) optional.

It could be possible to develop a hardware-based "push" client which would be installed by government regulation on all computers sold in Canada. This could not be turned off, so predetermined content could be pushed to the user. However, this would not be present in hardware purchased outside of Canada. In addition, requiring foreign (mainly US) manufacturers to install such a facility for imported computers is not practical; the incremental cost of such hardware might make these companies decide that the business case is not there, and to cease sales activities in this country. There are also free trade issues.

Email is another way to push information at Internet users. One approach to bringing the desired type of content to the attention of Canadian users would be to issue a regular email bulletin

promoting this content. However, while people can welcome email notifications of content of interest, unsolicited email sent indiscriminately to thousands of users (known as "spam") is extremely unwelcome. In response to the increased use of spam by organizations of every stripe, including publishers of pornography, filtering software has been developed and incorporated in most email client software that will automatically separate out email from known spam sites. Although this does not identify all spam, email users quickly become adept at identifying and discarding most spam simply by looking at the sender's address and the subject line. These messages are generally deleted without being read. An email bulletin could be an effective method of notifying Canadians of "desirable" Internet content and providing links to that content. To minimize the perception that this bulletin is "spam," however, users must be given the option of removing themselves from the mailing list if the bulletin is not of interest to them.

2.4 Content Substitution

Another technological approach to ensuring that content the government deems desirable reaches the eyes of Canadian Internet users might be to attempt to replace specific content on web pages with the desired content. Banner ads, for example, frequently occupy an identifiable placeholder on a web page. The server software then uses word recognition techniques to pull in an advertisement from a central advertising server based on the content of the web page. For example, doing a search on "cars for sale in Toronto" could result in a web page displaying advertisements for car dealers in that area. In theory, it might be possible to develop a software program that would scan each web page for these placeholders, and replace the original banner ad with content from a Canadian server. From a technological perspective, implementation of such a system would require the establishment of Internet choke points (network nodes through which all Internet data would have to pass) at which proxy servers could cache and search incoming web pages and insert Canadian content. (See the discussion on the costs of proxy serving in the section on filtering, above.) Clearly, such a procedure would bring Internet performance to a crawl. To reduce performance degradation, one might identify the most popular web sites on which to perform this substitution. The URLs of all incoming pages would then be screened against a list of popular sites. (Even the introduction of a filter to compare URLs against a list would degrade network performance, but to a far lesser degree than scanning and replacing the source code of every web page. See the discussion of filtering, above, on the impact on throughput of introducing filtering rules.) By substituting banner ads only on the most popular sites, one might reduce performance degradation to some extent. The negative effect on performance would, however, still be significant.

While substitution of content might be technologically possible, it would have a negative impact on network performance and would be extremely costly. Furthermore, such a procedure would undoubtedly violate the copyrights and moral rights of the creators of the web page. Content providers generate revenues through the placement of banner ads on their web sites. Aside from any legal remedies which might be sought, once it was discovered that Canada was stripping out the banner ads that had been paid for and was replacing them with other content, we could expect the operators of these sites to bar access to their sites from Canadian sources.[66]

[66] It would be a relatively simple matter for a host to refuse requests from IP addresses within specified regions. Using the same technological approach described above in the discussion of prioritizing Canadian content in search engines, hosts could identify requests as coming from Canadian IP addresses, and simply deny those requests.

2.5 PROMOTING CONTENT BY IMPROVING QUALITY OF SERVICE

One potentially viable technical approach to promoting designated content is to make it more appealing to users than content from non-Canadian sources by making it easier to access the designated content. This could be done by reproducing the designated content on multiple servers throughout the country and ensuring that Canadians have high speed network connections to those servers. Given a choice between slow access to alternate content, and high speed access to the designated content, Canadians would be likely to access the designated content first to determine whether it met their information and entertainment needs, before seeking content from other sources.

CONCLUSIONS

This report investigates whether technology can be effectively applied to the regulation of content on the Internet in Canada in order either to restrict content that is deemed undesirable (e.g., pornography) or, conversely, to promote content deemed desirable (e.g., Canadian content).

The authors conclude that while a number of technologies do exist that could be applied toward the regulation of Canadian's access to Internet content, none of these technological approaches would effectively prevent the Canadian Internet user from accessing content that violates pre-defined rules of acceptability, nor would they ensure that this same user would be exposed to any measure of desirable content.

SCREENING CONTENT

There are basically two technological approaches to restricting access to content on the Internet. These are (a) blocking requests for identified unacceptable content using a list of prohibited sites, and (b) filtering of content either on the basis of partial- or full-text searches to identify prohibited text strings or on the basis of rating labels attached to the content. Both approaches would entail prohibitive costs to implement on a national scale, and neither method would effectively block the user from accessing non-desirable content.

Blocking of identified content would be ineffective because:

1. The volume of content transmitted using the Internet and its constantly changing nature make the development and maintenance of a comprehensive, timely and coherent list of prohibited content impossible;
2. Sites can easily change their URLs and IP addresses to defeat blocking;
3. Users can request content anonymously through a remailer or proxy service;
4. Users can make encrypted requests for content and receive encrypted responses. The encryption makes screening of the content impossible.

Furthermore, blocking of an entire server because it may contain some offending content is undesirable as this may result in much valuable content being blocked. For example, blocking of a server that hosts web sites for many different organizations because of the unacceptability of a relatively small amount of content could deprive users of access to a large quantity of useful information.

Filtering of content on the basis of partial or full-text searching for prohibited text strings would be ineffective because:

1. There are no text strings that can reliably distinguish "acceptable" content from "unacceptable" content – i.e., the presence of a particular word or set of words does not necessarily mean that content containing that word or set of words is unacceptable or acceptable in all cases;
2. Text string searches cannot interpret non-textual content such as audio, video, or graphical content, which is becoming an increasingly important component of Internet content;
3. Encryption can defeat any attempt to engage in partial- or full-text searching.

Filtering of content on the basis of rating labels would be ineffective because:

1. The volume of content on the Internet, its rate of growth, and its constantly changing nature make the review and labeling of all Internet content impossible;
2. No labeling system exists that provides rating criteria directly relevant to Canadian legal standards. A labeling system would have to be developed that identified content as offending Canadian law. The adoption of such a system by content developers around the world is not to be expected;
3. Rules-based labeling systems are ineffective, since there is no set of objective criteria that can reliably distinguish "acceptable" content from "unacceptable" content. Standards-based labeling systems are ineffective because they rely on the subjective judgment of the rater. Because the volume of content to be rated would require many thousands of raters, there would be inconsistencies in rating of similar material;
4. Dynamic content such as is generated in chat rooms is not amenable to labeling, as the subject matter changes constantly in real time.

In addition to the foregoing weaknesses of blocking and filtering methodologies, any restrictions imposed in Canada could be avoided by users obtaining Internet service from non-Canadian providers, either through traditional landlines, or through new modes of Internet access such as digital satellite receivers.

Using any of the foregoing approaches to content restriction would entail the implementation of proxy servers or firewalls at the ISP, backbone node, or border crossing level. The costs involved – hardware, software, facilities, monitoring and maintenance – would be in the billions of dollars. If imposed on the individual ISP, these costs would drive many small ISPs out of business and larger ones across the border, or , alternatively, the ISPs would pass the cost increase on to the

consumer, making the cost of Internet access in Canada significantly higher than in neighbouring countries. Furthermore, any attempt at text-level filtering of incoming content would degrade network performance to the point of unacceptability. Given the lack of efficacy of these approaches to content restriction, the cost and performance impacts cannot be justified.

PROMOTING CONTENT

Approaches to content promotion considered in this report included:

- requiring Canadian ISPs to operate portal sites containing specified percentages of the desired type of content;
- special versions of Internet search engines that would prioritize the desired type of content for Canadian Internet users;
- substitution of non-Canadian banner ads on web pages with Canadian banner ads;
- "pushing" the desired type of content at Canadian Internet users through email or some other means; and
- improving the ease of access to desired content by providing Canadians with high speed network connections to the desired content.

While various methodologies might be considered for promoting certain kinds of content on the Internet, we have concluded that no purely technological approach will guarantee that Canadian Internet users will be exposed to that content.

Requiring Canadian ISPs to operate portal sites containing specified percentages of the desired type of content is impracticable for a number of reasons. The nature of web pages makes measurement of the amount of any particular kind of content problematic. The difficulties arise both because of the different disk storage and display properties of different media formats (text, graphics, audio, video) and because hyperlinking makes it difficult to determine the content boundaries of a web site. Furthermore, portal site operators will only offer the desired type of content if it satisfies a market demand. If portal site operators found that they were losing market share as a consequence of content regulations, they would either cease to operate or move their operations across the border, where they would not face content regulation. Finally, Canadian Internet users will only use a portal site if the content there is useful and interesting to them. If not, they will access Internet content either through US portal sites or directly via search engines.

Developing special versions of Internet search engines that would prioritize the desired type of content for Canadian Internet users is possible for content that can be simply distinguished using criteria such as the IP address of the host server. For example, the location of the server could be identified from well-known attributes of the site, such as the IP address or URL, using packet-filtering techniques. However the nature of the content at the site could not be identified in this way. These search engines will only be used by Canadian Internet users if they find this prioritization useful. If not, they will seek out and use any of the free Internet search services available on-line.

Substitution of "Canadian" banner ads for "non-Canadian" banner ads on popular web pages is technologically possible but is not practically viable. Not only would it require a massive proxy-serving infrastructure, but it would also slow network performance to an unacceptable level. From a non-technological perspective, changing the content on a third party's web page probably constitutes a violation of copyright and moral rights. Content providers who learned that this was being done could be expected to start refusing requests coming from Canada.

To some degree, it would be possible to use "push" technologies to deliver the desired type of content to Canadian Internet users without their having requested it. Push technologies require that the user have a "push client" (a piece of software or hardware that will receive and display the content) installed on the client machine. The most commonly installed push client software is an email client. It would be possible to use email to deliver or notify Canadian Internet users of certain content. To avoid Internet users perceiving this email as "junkmail" or "spam," it would be necessary to provide them with the option of removing themselves from the mail list if the information was not of interest to them. Recent experience in the market place has shown that other forms of push technology have not been widely adopted and used by the Internet community and there is no reason to believe that they would be widely accepted in this context. As a matter of fact, many users have vociferously opposed receiving information that they themselves have not specifically requested.

One potentially viable technological method of promoting certain types of content would be to make that content more accessible than other content on the Internet. This could be done by replicating the desired content on multiple proxy servers across Canada and ensuring that Canadians have high speed network connections to those servers. If that content meets Canadians' business, information, education, and entertainment needs, its speed of access will induce Canadians to use it in preference to other content on the Internet.

RECOMMENDATIONS

Ultimately, it is the quality of the content and its interest to users that will determine whether Canadian Internet users decide to look at it. The steps the government is taking in supporting initiatives such as CA*net III , SchoolNet and the Community Access Program, Strategis and other government web sites, as well as programs for the development of outstanding Canadian content will do much more than any regulatory regime to ensure that Canadians access Canadian content on the Internet.

Finally, it should be noted that it may be possible to control Internet content *to some extent*, but only if we were prepared to accept the considerable costs in terms of technological infrastructure, human resources, enforcement mechanisms, and social and legal consequences. For instance, it may be possible to produce a limited list of Web sites that violate Canadian legal standards and to require Canadian ISPs to filter for these prohibited sites. It also may be possible to require Canadian portal sites to display a number of banners containing specified content or to include a list of hyperlinks to other sites containing that specified content. As we have seen, however, both from the technological discussion above and from the experience of other countries, such an approach would only restrict access to a limited number of offending sites. It would not guarantee that Canadians would be protected from other Internet content that violates Canadian legal standards and has not been screened by the authoritative body that composes the list. Moreover, for the reasons discussed previously, such an approach would fail to ensure that Canadians were exposed to a particular kind of "desired" content. However, if we were prepared to accept both the substantial costs described above, as well as the consequent technological and operational problems (e.g., lack of accuracy, performance degradation, lack of scalability, administrative overhead, etc.), this sort of ISP filtering might serve as a "best-efforts" technological approach to regulating Internet content. While it might satisfy the concerns of some Canadians, the authors believe that the unreliable, hit-and-miss results of this approach would not justify its costs or its ensuing negative impact on Canada's place in the global Information Economy.

The authors also wish to point out that much content that is not suitable for children is legal for adults in Canada. Therefore, blocking of content that is not suitable for children is not appropriate at the level of the ISP because, by restricting material that is not suitable for minors, the ISP would also be denying legitimate access by adults.

For all of the foregoing reasons, we believe that the most promising technological avenue for regulating access to Internet content is self-regulation through voluntary client-side filtering (e.g.,

using software such as Net Shepherd or SafeSurf) combined with voluntary self-labeling of Internet content by content providers (e.g., using a PICS-compliant labeling system). Restricting access to some types of Internet content by children is an important issue that must be addressed. However, attempting to exert this control through a national regulatory framework requiring blocking and/or filtering facilities is not viable for reasons mentioned previously. Despite the limitations of filtering software discussed in the report, filtering software installed on the family PC may meet the majority of the needs of those parents who wish to restrict their children's Internet access. It is a good first step.

APPENDIX A – HOSTS BY DOMAIN AS OF JULY, 1998

Domain	Hosts	Full Name
TOTAL	36,739,151	
com	10,301,570	Commercial
net	7,054,863	Networks
edu	4,464,216	Educational
mil	1,359,153	US Military
jp	1,352,200	Japan
us	1,302,204	United States
uk	1,190,663	United Kingdom
de	1,154,340	Germany
ca	1,027,571	Canada
au	750,327	Australia
org	644,971	Organizations
gov	612,725	Government
nl	514,660	Netherlands
fi	513,527	Finland
fr	431,045	France
se	380,634	Sweden
it	320,725	Italy
no	312,441	Norway
es	243,436	Spain
ch	205,593	Switzerland
dk	190,293	Denmark
nz	177,753	New Zealand
kr	174,800	Korea, Republic Of
br	163,890	Brazil
be	153,760	Belgium
za	140,577	South Africa
at	132,202	Austria
ru	130,422	Russian Federation
tw	103,661	Taiwan, Province Of China
pl	98,798	Poland
il	87,642	Israel
mx	83,949	Mexico
hu	73,987	Hungary
hk	72,232	Hong Kong
cz	65,672	Czech Republic
sg	59,469	Singapore
ar	57,532	Argentina
arpa	47,910	Mistakes
pt	45,113	Portugal
ie	44,840	Ireland
my	40,758	Malaysia
gr	40,061	Greece
tr	27,861	Turkey
th	25,459	Thailand
unkno	wn 23610	Unknown
cl	22,889	Chile
is	20,678	Iceland

su	20,024	Soviet Union
cn	19,313	China
ee	18,948	Estonia
si	18,084	Slovenia
uy	16,345	Uruguay
sk	14,154	Slovakia (Slovak Republic)
ro	13,697	Romania
ae	13,519	United Arab Emirates
ua	13,271	Ukraine
co	11,864	Colombia
id	10,691	Indonesia
in	10,436	India
lt	8,746	Lithuania
lv	8,115	Latvia
ph	7,602	Philippines
ve	6,825	Venezuela
lu	6,145	Luxembourg
bg	6,141	Bulgaria
hr	6,117	Croatia (local name: Hrvatska)
kw	5,597	Kuwait
yu	5,270	Yugoslavia
do	4,917	Dominican Republic
pe	3,763	Peru
cy	3,286	Cyprus
cr	2,844	Costa Rica
eg	2,043	Egypt
bm	1,993	Bermuda
pk	1,923	Pakistan
nu	1,608	Niue
tt	1,531	Trinidad And Tobago
to	1,446	Tonga
lb	1,400	Lebanon
kz	1,397	Kazakhstan
ec	1,227	Ecuador
gt	1,046	Guatemala
py	855	Paraguay
int	853	International Organizations
zw	836	Zimbabwe
mt	785	Malta
pa	766	Panama
bn	740	Brunei Darussalam
ni	692	Nicaragua
ke	692	Kenya
om	666	Oman
na	665	Namibia
sv	647	El Salvador
by	636	Belarus
ge	632	Georgia
lk	580	Sri Lanka
bw	578	Botswana
fo	560	Faroe Islands

gl	515	Greenland
vi	514	Virgin Islands (U.S.)
bo	506	Bolivia
ma	478	Morocco
ad	477	Andorra
am	466	Armenia
mk	407	Macedonia, The Former Yugoslav Republic Of
li	402	Liechtenstein
sz	397	Swaziland
mu	370	Mauritius
md	370	Moldova, Republic Of
jo	360	Jordan
ky	359	Cayman Islands
ba	348	Bosnia And Herzegowina
bh	337	Bahrain
tm	296	Turkmenistan
pf	273	French Polynesia
ci	265	Cote D'Ivoire
ir	262	Iran (Islamic Republic Of)
bz	262	Belize
cc	259	Cocos (Keeling) Islands
jm	253	Jamaica
bs	247	Bahamas
gh	241	Ghana
zm	236	Zambia
az	231	Azerbaijan
uz	198	Uzbekistan
ag	196	Antigua And Barbuda
gi	191	Gibraltar
sn	189	Senegal
ai	189	Anguilla
kg	182	Kyrgyzstan
sm	154	San Marino
mc	154	Monaco
mo	143	Macau
nc	141	New Caledonia
tz	137	Tanzania, United Republic Of
tc	129	Turks And Caicos Islands
fj	127	Fiji
pr	123	Puerto Rico
np	123	Nepal
gf	121	French Guiana
gp	115	Guadeloupe
hn	106	Honduras
fm	95	Micronesia, Federated States Of
bf	93	Burkina Faso

ng	91	Nigeria
gu	89	Guam
cu	85	Cuba
tg	83	Togo
mz	83	Mozambique
gb	81	United Kingdom
dm	79	Dominica
et	76	Ethiopia
al	76	Albania
mv	70	Maldives
st	64	Sao Tome And Principe
pg	62	Papua New Guinea
kh	58	Cambodia
gy	58	Guyana
tn	57	Tunisia
tj	57	Tajikistan
io	56	British Indian Ocean Territory
nf	55	Norfolk Island
vu	47	Vanuatu
bb	45	Barbados
sa	42	Saudi Arabia
ug	41	Uganda
ck	33	Cook Islands
vn	25	Viet Nam
sb	24	Solomon Islands
lc	24	Saint Lucia
qa	23	Qatar
mr	22	Mauritania
im	21	Isle of Man
dz	19	Algeria
mg	18	Madagascar
as	18	American Samoa
mq	17	Martinique
mn	17	Mongolia
ls	17	Lesotho
ye	14	Yemen
je	14	Jersey
vg	13	Virgin Islands (British)
gw	13	Guinea-Bissau
gg	13	Guernsey
bj	13	Benin
cx	11	Christmas Island
va	9	Vatican City State (Holy See)
km	9	Comoros
mp	8	Northern Mariana Islands
cd	8	Congo (Democratic Republic)
tf	7	French Southern Territories
sc	7	Seychelles
ms	7	Montserrat

an	6	Netherlands Antilles
ne	5	Niger
cm	5	Cameroon
ac	5	Ascension Island
mh	2	Marshall Islands
bt	2	Bhutan
ao	2	Angola
tp	1	East Timor
sh	1	St. Helena
re	1	Reunion
pw	1	Palau
ml	1	Mali
ly	1	Libyan Arab Jamahiriya
lr	1	Liberia
kn	1	Saint Kitts And Nevis
hm	1	Heard And Mc Donald Islands
gs	1	South Georgia And The South Sandwich Islands
ga	1	Gabon
fk	1	Falkland Islands (Malvinas)
cv	1	Cape Verde
cg	1	Congo (Republic)
af	1	Afghanistan
zr	0	Zaire
yt	0	Mayotte
ws	0	Samoa
wf	0	Wallis And Futuna Islands
vc	0	Saint Vincent And The Grenadines
um	0	United States Minor Outlying Islands
tv	0	Tuvalu
tk	0	Tokelau
td	0	Chad
sy	0	Syrian Arab Republic
sr	0	Suriname
so	0	Somalia
sl	0	Sierra Leone
sj	0	Svalbard And Jan Mayen Islands
sd	0	Sudan
rw	0	Rwanda
pn	0	Pitcairn
pm	0	St. Pierre And Miquelon
nr	0	Nauru
mw	0	Malawi
mm	0	Myanmar
la	0	Lao People's Democratic Republic
ki	0	Kiribati

iq	0	Iraq
ht	0	Haiti
gq	0	Equatorial Guinea
gn	0	Guinea
gm	0	Gambia
gd	0	Grenada
er	0	Eritrea
dj	0	Djibouti
cf	0	Central African Republic
bv	0	Bouvet Island
bi	0	Burundi
aw	0	Aruba
aq	0	Antarctica

APPENDIX B - RATING SYSTEMS

Examples of rating systems used by some content filtering products. Source: Internet Family Empowerment White Paper, Center for Democracy & Technology, July 1997. www.cdt.org/speec/empower.html

RSACi ratings

NUDITY
Level 0 - no nudity
Level 1 - revealing attire
Level 2 - partial nudity
Level 3 - frontal nudity
Level 4 - provocative frontal nudity
SEX
Level 0 - innocent kissing or romance
Level 1 - passionate kissing
Level 2 - clothed sexual touching
Level 3 - non-explicit sexual acts
Level 4 - explicit sexual acts; sex crimes
LANGUAGE
Level 0 - no offensive language
Level 1 - mild expletives
Level 2 - profanity
Level 3 - strong language; hate speech
Level 4 - extreme hate speech; crude, vulgar language
VIOLENCE
Level 0 - none or sports violence
Level 1 - injury to human beings
Level 2 - destruction of objects with implied social presence
Level 3 - death to human beings; blood and gore
Level 4 - wanton, gratuitous violence; rape

SafeSurf Ratings

The SafeSurf SS~~ Rating Standard
Designed by and for parents to empower each family to make informed decisions concerning accessibility of online content.

Section One: Adult Themes with Caution Levels
0. Age Range
1) All Ages
2) Older Children
3) Teens
4) Older Teens
5) Adult Supervision Recommended
6) Adults
7) Limited to Adults
8) Adults Only
9) Explicitly for Adults
Section One: Adult Themes with Caution Levels
1. Profanity
1) Subtle Innuendo
description: Subtly Implied through the use of Slang
2) Explicit Innuendo
description: Explicitly implied through the use of Slang

3) Technical Reference
description: Dictionary, encyclopedic, news, technical references
4) Non-Graphic-Artistic
description: Limited non-sexual expletives used in a [sic] artistic fashion
5) Graphic-Artistic
description: Non-sexual expletives used in a [sic] artistic fashion
6) Graphic
description: Limited use of expletives and obscene gestures
7) Detailed Graphic
description: Casual use of expletives and obscene gestures.
8) Explicit Vulgarity
description: Heavy use of vulgar language and obscene gestures. Unsupervised Chat Rooms.
9) Explicit and Crude
description: Saturated with crude sexual references and gestures. Unsupervised Chat Rooms.
2. Heterosexual Themes
1) Subtle Innuendo
description: Subtly Implied through the use of metaphor
2) Explicit Innuendo
description: Explicitly implied (not described) through the use of metaphor
3) Technical Reference
description: Dictionary, encyclopedic, news, medical references
4) Non-Graphic-Artistic
description: Limited metaphoric descriptions used in a [sic] artistic fashion
5) Graphic-Artistic
description: Metaphoric descriptions used in a [sic] artistic fashion
6) Graphic
description: Descriptions of intimate sexual acts
7) Detailed Graphic
description: Descriptions of intimate details of sexual acts
8) Explicitly Graphic or Inviting Participation
description: Explicit Descriptions of intimate details of sexual acts designed to arouse. Inviting interactive sexual participation.
Unsupervised Sexual Chat Rooms or Newsgroups.
9) Explicit and Crude or Explicitly Inviting Participation
description: Profane Graphic Descriptions of intimate details of sexual acts designed to arouse. Inviting interactive sexual
participation. Unsupervised Sexual Chat Rooms or Newsgroups.
3. Homosexual Themes
1) Subtle Innuendo
description: Subtly Implied through the use of metaphor
2) Explicit Innuendo
description: Explicitly implied (not described) through the use of metaphor
3) Technical Reference
description: Dictionary, encyclopedic, news, medical references
4) Non-Graphic-Artistic
description: Limited metaphoric descriptions used in a [sic] artistic fashion
5) Graphic-Artistic
description: Metaphoric descriptions used in a [sic] artistic fashion
6) Graphic
description: Descriptions of intimate sexual acts
7) Detailed Graphic
description: Descriptions of intimate details of sexual acts
8) Explicitly Graphic or Inviting Participation
description: Explicit descriptions of intimate details of sexual acts designed to arouse. Inviting interactive sexual participation.
Unsupervised Sexual Chat Rooms or Newsgroups.
9) Explicit and Crude or Explicitly Inviting Participation
description: Profane Graphic Descriptions of intimate details of sexual acts designed to arouse. Inviting interactive sexual
participation. Unsupervised Sexual Chat Rooms or Newsgroups.

4. Nudity
1) Subtle Innuendo
description: Subtly Implied through the use of composition, lighting, shaping, revealing clothing, etc.
2) Explicit Innuendo
description: Explicitly implied (not shown) through the use of composition, lighting, shaping or revealing clothing
3) Technical Reference
description: Dictionary, encyclopedic, news, medical references
4) Non-Graphic-Artistic
description: Classic works of art presented in public museums for family viewing
5) Graphic-Artistic
description: Artistically presented without full frontal nudity
6) Graphic
description: Artistically presented with frontal nudity
7) Detailed Graphic
description: Erotic frontal nudity
8) Explicit Vulgarity
description: Pornographic presentation
9) Explicit and Crude
description: Explicit pornographic presentation
5. Violence
1) Subtle Innuendo
2) Explicit Innuendo
3) Technical Reference
4) Non-Graphic-Artistic
5) Graphic-Artistic
6) Graphic
7) Detailed Graphic
8) Inviting Participation in Graphic Interactive Format
9) Encouraging Personal Participation, Weapon Making
6. Sex, Violence, and Profanity
1) Subtle Innuendo
2) Explicit Innuendo
3) Technical Reference
4) Non-Graphic-Artistic
5) Graphic-Artistic
6) Graphic
7) Detailed Graphic
8) Explicit Vulgarity
9) Explicit and Crude
7. Intolerance
1) Subtle Innuendo
2) Explicit Innuendo
3) Technical Reference
4) Non-Graphic-Literary
5) Graphic-Literary
6) Graphic Discussions
7) Endorsing Hatred
8) Endorsing Violent or Hateful Action
9) Advocating Violent or Hateful Action
8. Glorifying Drug Use
1) Subtle Innuendo
2) Explicit Innuendo
3) Technical Reference
4) Non-Graphic-Artistic
5) Graphic-Artistic
6) Graphic
7) Detailed Graphic
8) Simulated Interactive Participation
9) Soliciting Personal Participation

9. Other Adult Themes
1) Subtle Innuendo
2) Explicit Innuendo
3) Technical Reference
4) Non-Graphic-Artistic
5) Graphic-Artistic
6) Graphic
7) Detailed Graphic
8) Explicit Vulgarity
9) Explicit and Crude
A. Gambling
1) Subtle Innuendo
2) Explicit Innuendo
3) Technical Discussion
4) Non-Graphic-Artistic, Advertising
5) Graphic-Artistic, Advertising
6) Simulated Gambling
7) Real Life Gambling without Stakes
8) Encouraging Interactive Real Life Participation with Stakes
9) Providing Means with Stakes

Net Shepherd Ratings

In December 1996, NSI launched an Internet event: We recruited over 300 'Net afficionados [sic] to examine sites and rate them for maturity and quality using NSI's CRC (Collaboratively Rated Content) rating scale.

The CRC rating scale has six maturity levels (General, Child, Pre-teen,Teen, Adult and Objectionable), and five quality levels (1 through 5 stars, with 5 stars signifying excellence). Quality on the CRC scale includes everything from content to navigation to graphics, and ultimately reflects the overall impression our raters have of the sites they visit.

APPENDIX C: FILTERING SOFTWARE

Title	Publisher	Platform	Method	Features
Email For Kids	ConnectSoft	Win 95	Blocking[67]	Parents can edit outgoing/incoming message content, address book entries and online time. Prevents personal disclosures, controls # of messages sent, and regulates who communicates with your children. $29.95. AOL, Prodigy and Compuserve editions available at reduced fee.
Net Nanny 3.1	Net Nanny International	Win 95	Blocking	Blocks objectionable sites, phrases, personal disclosures, chatrooms, gifs & jpegs. Monitors outgoing/incoming email & other text documents. Logs violations. Custom editing. Single user and network versions. Multiple user accounts. $26.95
Safe-Net Suite	Maia Software	Win 95 Win 98 Win NT	Blocking	3 Separate modules provide filtering protection for email, news and chat and screen for inappropriate language, nudity, violence and spam. Parents may edit any of the filters. Inappropriate email content is forwarded to a password- protected folder. Available as freeware/shareware.
SurfWatch 3.0	Spyglass, Inc.	Win 95, Win 3.1, Mac OS	Blocking	One of the easiest filters to configure and use. Blocks web access and chat. $49.95 includes one year of updates.
Time's Up	Fresh Software Company	Win 95 Win 3.1	Blocking	Time control software that lets parents set software and Internet access limits. $19.95. Can be ordered as a bundle deal, with Surfwatch.
Web Chaperone 1.1	WebCo International, Inc.	Win 95	Blocking	On-the-Fly scanning of web pages for appropriate content, using PICS Rating system. Has both age group and protection level settings and can be set up for multiple family users. $49.95
WizGuard	WizGuard Company	Win 98 Win 95	Blocking	Content-sensitive filtering, customizable blocking, web access log and time management control. $29.99 Proxy server edition available for schools & libraries.
X-Stop v.3.01	Log-On Data Corp.	Win 95 Win 3.1 Mac OS	Blocking	Blocks access to over 100,000 adult-only sites, offensive words on web sites, in chatgroups, email or in any text document offline. Editing capabilities. Can also block personal disclosures in outgoing mail. Updates available at no charge. $39.95.
Safe Surf	Safe Surf, Inc.	Win 95 Win 3.1 Mac OS	Blocking and rating system	A browser feature of Netscape's "NetWatch" which uses the SafeSurf rating standard and adult database. Parents can block unrated sites and adjust content levels in different categories. Provides list of safe sites.
Chi-Brow 2.0	KCS & Associates	Win 95	Blocking, Controlled[68] Access	Simplified children's browser that gives parents total control of designating appropriate web sites and blocking objectionable ones. Includes database of recommended sites. $39.95.
CyberPatrol 3.3	Learning Company	Win 95, Win 3.1 Mac OS	Blocking, Controlled Access	One of the easier security devices to configure and use. Allows multi-user access with customizable profiles for each family member. Uses PICS standard for blocking of inappropriate sites. Also allows use of predetermined sites. Parents can edit CyberYes/CyberNot lists. $29.95 plus monthly subscription fee.
Hexabit Junior 2.0	Hexabit	Win 95	Blocking, Controlled Access	A web browser with tv controls, designed to provide easy navigation for kids. Parents choose which sites are appropriate/not appropriate and the times of day in which the Internet may be accessed. Parents may also restrict activity to specific sites - or set hours of day during which the WWW may be accessed. $20.00
Safe Search	InterGo Communications	Win 95, Win 3.1	Blocking, Controlled Access	"Kinderguard" security screen and proprietary browser. Blocks objectionable sites and personal disclosures. Limits chat and newsgroup access. Search engine rates web sites for age appropriateness. Can restrict user access to specific web sites. $49.95 for 12-month subscription.
Surfin' Annette	Spycatcher Corp.	Win 95	Blocking, Controlled Access	A web browser designed for children under 18. Complete content control is in the hands of the parent. Browser detects objectionable words, phrases, URLs. List is updated monthly and parents can edit to meet their personal values. $29.95 for families. Free for K-12 schools.
CyberSitter '97 v.8	Solid Oak Software	Win 95	Blocking, Stealth [69]Monitoring	Blocks chat and undesirable sites; offers daily updates to banned sites. Parents can add sites, but not edit pre-determined list. Logs all activity and filters email. Prevents personal disclosures. $39.95.
Disk Tracy	Watchsoft, Inc.	Win 95 Win NT	Blocking, Stealth Monitoring	Home/Network versions. Screens and blocks undesirable content. Logs reports on sites visited and material downloaded/stored on the computer. $34.95.
Gulliver's Guardian, Internet Suite	Gulliver Software Ltd.	Win 95	Blocking, Controlled Access	Customizable software with multiple user profiles, time & duration control, email screening. Blocks inappropriate sites. Can also limit access to predetermined sites and prevent alteration/deletion of desktop files. $59.95.
SOS Internet Filter	Sterling Strategic	Win 95	Blocking,	Combo desktop security and Internet filtering software. Multi-user settings, restricted and

[67] Blocking prevents access to specified sites.
[68] Controlled access allows access to designated sites or services only.
[69] Stealth monitoring creates a log of activities without the knowledge of the software's user.

	Solutions, Inc.		Controlled Access	allowed sites, activity and time logs. Custom-editing. Can restrict access to specific sites or specific groups. Also controls Win 95 settings and CD-Rom. $44.95. Academic pricing available.
Internet Filter Suite 1.0	Turner Investigations	Win 95 Win 3.1	Blocking, Stealth Monitoring	A monitoring and filtering program that allows parents to control site access, undesirable words, newsgroups, chat sessions. Logs all data transfers. Emails parents when violations occur. $40.00
Net-Rated	PC Data Power	Win 95, Win 3.1	Blocking, Stealth Monitoring	One of the more complex filtering tools to configure. Tracks all activities. Allows use of pre-approved sites and prevents access from objectionable sites. Some editing features. Regulates time/hours of the day that programs can be used. Does not block chat or filter email.
KidDesk Internet Safe	Edmark	Win 98, Win 95	Controlled access	Combines a desktop security device with an Internet safety program to give parents complete control of software and internet access. Can be customized for each family member with their own photo or name plate icon, desktop environment, software picks and Internet destinations. Includes a timer that can be preset by parents. $29.95
KiddoNet	GTech Technologies Ltd.	Win 95	Controlled access	Controlled access using a kids' browser and pre-approved sites. The browser includes its own offline games and activities. Appropriate for ages 8-12.
Click and Browse Jr.'98	Netwave, Inc.	Win '98 Win '95 Win NT	Controlled Access	Closed-loop browser offering a safe and visual guide to hundreds of kid-friendly sites for ages 3-up. Parents determine how much of the WWW they wnat their children to see, email and newsgroup access. Company has offered free browser to 86,000 schools. Works with AOL. Consumer ver. $19.95
Ed View	Ed View, Inc.	Win 95 Win 3.1 Mac OS	Controlled Access	EdView Smart Zone search engine of 7 million pages of educator-reviewed child safe sites allows users to see appropriate content, while blocking pieces and subsets of sites. Content is organized by subject. Password- protected software keeps users in safe areas. Blocking is age-specific. A Family Edition sells for $39.95 and includes entertainment sites. School building licenses are also affordably priced.
Kid Web	ConnectSoft	Win 95	Controlled access	Johnny Web and his dog, Browser guide kids (age 5-10) safely through the 'net to predetermined sites. Parents regulate control. Includes permit mode, ban mode and word block. Includes special home page and subject explorer. and word block. Includes special home page and subject explorer.
Mama Bear	Kodiak Software Systems	Win 95, Win 3.1	Controlled access	Desktop security device. Blocks unauthorized software access, alterations or deletions of programs. Can build set of allowable software. Recognizes renamed programs. $99. Educational disc. available.
Microsoft Plus For Kids	Microsoft	Win 95	Controlled access	Integrated desktop and Internet security device for kids (age 3-12) with 10 environments and "content advisor" for the Internet. Comes with Surfwatch.
Surf Monkey	Media Live/Bandai Entertainment	Win 98 Win 95	Controlled Access	Rocketship browser, animated talking monkey and lively self-contained web destination for kids age 7 and up. Supervised chatrooms. Parents can block email from strangers, delete profanities and prohibit access to inappropriate sites. Uses Surfwatch for content filtering. Chat and email access can be turned off. Parents control a child's buddy list. Email has multimedia features, with text-to-speech, sound and picture capability. Works on the Internet Explorer engine. $29.95/yr.
WebLoc	Millennium Interactive, Inc.	n/a	Hardware Device	Hardware solution for parents who want to supervise their child's use of the web. Locks directly onto the modem and computer. Other applications can still be used. $24.95.
AUP Action Tools	iTech, Inc.	Win 95 Win 3.1	Proxy Server[70]	Provides blocking and filtering of 25,000 potentially offensive sites. Designed for district school use, AUP can handle hundreds of computers simultaneously. Blocks chat and news sites. Bounces email if it contains offensive language. Separate filters for elementary, middle and high schools.
Bess, The Internet Retriever	N2H2	Win 95 Win 3.1 Mac OS	Proxy Server	Subscription based service from ISP's or Proxy server for schools & libraries. Uses Inktomi Corp's search engine, with filtering services for inappropriate content and adult advertising. Bess web site offers thousands of pre-determined links for children.
iWay Patrol	iTech, Inc.	n/a	Proxy Server	Content filtering is enforced at the system level, preventing students from access to Chat Rooms, gambling, games, sex or pornography sites on the WWW. Teachers can add/delete sites from blocking lists or restrict access to specific sites.
Net Shepherd 2.0	Net Shepherd, Inc.	n/a	Proxy Server	Subscription-based service. Uses PICS database of restricted and permissible sites, with 6 rating categories for different age groups. Has controls for multiple users. Blocks chat but not personal disclosures. No email filtering. $12/yr.
Planet Web Browser	Planet Web, Inc.	Sega, TV Boxes, ISP's	Proxy Server	Requires proprietary browser. Parents fill out profiles for each user. Filtering is based on user birth date, family beliefs and levels of concern. Can customize settings in 15 categories but not specific entries in database of objectionable sites.
Smart Filter	Burst Technology	Win 95	Proxy Server	Controls filtering of both inappropriate material for children and non-business sites that may cut into workplace productivity. Proxy Reporter adds auditing features. It identifies, categorizes and reports on user activity.
WebSense	Netpartners, Inc.	n/a	Proxy Server	For CISCO PIX Firewall. Uses a database of 200,000 URLs in 30 categories to screen and block site access by IP number. List is updated daily. Designed for large number of users. Tracks and reports all activities.
I-Gear	URLabs	Win NT	Proxy Server Stealth Monitoring	Proprietary content management software for the K-12 Education Market. Dynamic document review analyzes content of the pages. Filtering can be tailored to meet the needs of every user. Software tracks pages accessed and keywords used. A separate mail

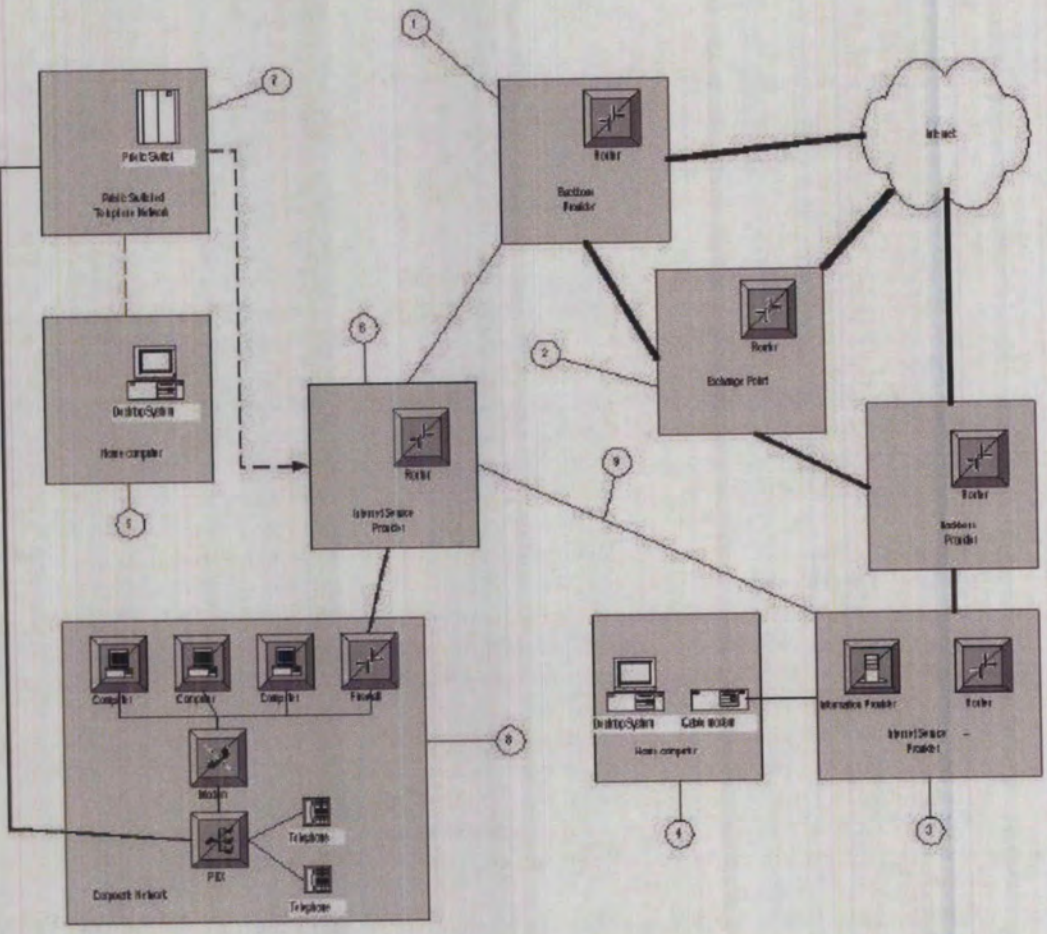
[70] A proxy server acts as a gateway. All service request are forwarded to the proxy and then to the Internet.

				filtering product is available.
GuardiaNet v 3.0	Landmark Community Interests	Win 95 Win 3.1	Proxy Server, Controlled Access	Server-based content filtering that allows parents to customize levels of access for each family member, according to family values. Uses SafeSurf's list of recommended sites. $59.95 first yr. $29.95 to renew. School pricing available.
GuardiaNet v 3.0	Landmark Community Interests	Win 95 Win 3.1	Proxy Server, Controlled Access	Server-based content filtering that allows parents to customize levels of access for each family member, according to family values. Uses SafeSurf's list of recommended sites. $59.95 first yr. $29.95 to renew. School pricing available.
CyberSentinel 1.5	Security Software Systems, Inc.	Win 95, Win 3.1	Stealth Monitoring	Monitors chatrooms, email, search engines and browsers using a contextual text-recognition system. Captures and logs offending screens. Also tracks offline software usage. Can be used in stealth mode or with active warnings. $49.99
CyberSnoop 3.0	Pearl Software, Inc.	Win 95 Win NT	Stealth Monitoring	Monitors email, chat and web site activity. Custom-editing feature also allows blocking and parent-determined acceptable sites. $29.95.
Family Cam	Silverstone Software	Win 95	Stealth Monitoring	Activity-monitoring software that tracks both software and Internet usage by taking random screen shots of user activity. $29.95.
Internet Watchdog	Charles River Media	Win 95 Win NT	Stealth Monitoring	Records all Internet and application activity, including applications used and graphics downloaded. $39.95
Prudence	Blue Wolf Network	Win 98 Win 95	Stealth Monitoring	Monitors web browsing sessions, Captures graphic files, history of viewed URL's, cookies, documents opened, addresses of bookmarked sites. Can be set to email the parent at work, with a list of URLs that have been viewed. $49.95.
SentryCam	GWG Devcore, Inc.	Win 95	Stealth Monitoring	Logs child's Internet activities. Captures screen shots of games, chatrooms, web sites, email. Adds a time/date stamp . $34.95
Smart Alex ICU	Smart Alex	Win 95	Stealth Monitoring	Monitors computer activity both on the Internet and offline. Takes screen shots of conversations, email, word processing documents, images. $29.95
Triple Exposure	IPS Corporation	Win 95 Win NT Win 3.1	Stealth Monitoring	Scans for all images, movies, profanities, URL's using predefined words. Recognizes files, even if the names have been changed. $29.95.
WinGuardian	WebRoot Software	Win 95	Stealth Monitoring	Single user and network versions. Logs all Internet activity. Locks system to prevent unauthorized use of programs. $29.95.
WinWhatWhere	WinWhatWhere Corp.	Win 95 Win 3.1	Stealth Monitoring	Monitors both Internet usage and applications, logging program, caption, start & elapsed times. Create Internet usage reports. $29.

APPENDIX D - A VISUAL REPRESENTATION OF INTERNET ARCHITECTURE

Notes for the diagram.

1. Backbone providers provide the major interconnections that support the bulk of the Internet traffic. The Internet service provided to a consumer may come from a local Internet Service Provider but may also be provided by a backbone provider directly to the retail level.
2. Backbone providers interconnect at exchange points.
3. Internet Service Providers offer connections to corporations and individuals in their homes. They may also provide a location for an information provider to distribute information.
4. Home computers may be connected to the Internet in a number of ways. This is an example of a cable modem. Cable modems provide a permanent, "always on" connection.
5. An example of a dial modem user. Dial modem users use a telephone circuit to establish a temporary connection to an ISP. Other connection methods not illustrated include Integrated Services Digital Network (ISDN, a form of dial connection), Asymmetric Digital Subscriber Loop (ADSL, a high speed permanent connection that uses telephone wires), wireless modems, microwave radio and several forms of satellite connection.
6. An ISP that provides both dial support via its connection to the telephone company's switch and permanent corporate connection services.
7. A telephone company switch. Connections to the switch could be standard phone lines to a person's home or permanent circuits to a corporation's local private branch exchange (PBX).
8. A corporate network. This corporation has a permanent connection to an ISP. The firewall ensures that only those services that the corporation wants are accessible from the Internet to protect the network from intruders. This corporation also operates a modem pool for its employees; in essence it is an ISP as well.
9. These two ISP's have decided on their own to establish a local means of interchanging information. This connection could be used to exchange email (for example) but could also be used to route information automatically if one of the ISP's loses contact with its backbone carrier. The Internet protocols allow the establishment of these ad hoc connections between any two points. The protocols also allow the two partners to control how the route can be used; it could be local or it could be advertised to the whole Internet as an available path. This illustrates the ease with which a path could be constructed to bypass any restriction somewhere on the net.



APPENDIX E - MEDIA COVERAGE OF CHINA'S ATTEMPTS TO REGULATE ACCESS TO INTERNET CONTENT BY ITS CITIZENS

The following excerpt from an article entitled "China Losing Battle to Control Internet Content" discusses the Chinese experience[71]:

> **China has lost its battle to control Internet content, according to published reports today.**
>
> **A year ago, the Chinese government announced regulations to censor online content to control what Beijing termed "spiritual pollution" on line. And with much fanfare, Beijing launched the China Wide Web, intended to be China's only pipeline to the Internet. The CWW enabled online commerce but censored sensitive political news, and only approved World Wide Web sites were available through the CWW.**
>
> **Today that grand plan for online controls is in disarray, according to a report published this morning in the *South China Morning Post.***
>
> **China's computers now have access to CNN and other news agencies, which earlier this year had been blocked by Chinese authorities. According to today's report, even the dissident-produced *China News Digest* can now be accessed through one Internet service provider in Shanghai.**
>
> **The Tibet Information Network, considered by Beijing to be one of the most offensive Internet sites, is also now available in China, according to the report.**
>
> **Pornography is also now available online in China. The Internet filters no longer exclude even such obvious sex sites as www.sex.com, today's report said.**
>
> **Beijing has also lost control of individuals going online. Government regulations require all individual subscribers to register with China's Public Security Bureau. But today's newspaper report indicates Internet access**

[71] "China Losing Battle to Control Internet Content"; Adam Clayton Powell III, http://www.freedomforum.org/technology/1997/12/11china.asp

cards, similar to prepaid telephone cards, are widely available, giving users access to the Internet without clearance from the government.

Observers in Beijing and Shanghai have counted 30 Internet service providers. The enormous growth of the Net in the past year may have overwhelmed the government's censorship hardware, according to today's report.

Use of the Internet in China has grown sharply in 1997: In January, the number of users was estimated at 100,000-150,000 users. But according to today's report, that number has grown to 250,000. Other estimates this week put the number of Internet accounts at 300,000-400,000.

Internet users in China have been able to bypass national censorship by dialing long distance to proxy servers outside China. But few Chinese could afford the telephone charges. Now, previously censored sites are available from local providers.

A January 21, 1999 news item on CNN Interactive[72] details new Chinese government restrictions on the Internet:

SHANGHAI, China (AP) -- China has tightened restrictions on Internet use, ordering bars that offer access to register users with the police, according to state media.

The rules issued this week come amid a crackdown on Internet political activity that caused an outcry when a Shanghai man was imprisoned for giving email addresses to dissidents abroad. Under the rules, bars that rent time to customers on Internet-linked computer terminals will have to be licensed by police, the Workers Daily newspaper said today.

Such bars and cafes, increasingly common in major Chinese cities, had been one of the few ways Chinese could receive email or look at Websites anonymously.

[72] "China imposes new restrictions on Internet use"
http://cnn.com/TECH/computing/9901/21/net.restrict.china.ap/

"Managers and customers of 'Internet bars' cannot be allowed to endanger national security," the newspaper said.

The Workers Daily did not give any details of the rules, but the state-run China News Service said bar managers would have to be licensed and register their customers.

The reports said the rules were issued Tuesday by public security and culture officials, but didn't say when they would take effect.

The China News Service said public morals and stability already were under threat. "Some managers offer gambling and computer games with lewd content," it said in a report Tuesday. "Officials believe this already has endangered social stability and the mental and physical health of young people."

The government has encouraged the rapid spread of Internet use in China, but closely monitors its 1.5 million registered users. Service providers are required to register customers with the authorities. Barriers have been installed to block access to sites deemed subversive or pornographic.

It is interesting to note that, in this case, the mechanisms for restriction have moved from technical attempts to government regulation of both providers and users. The civil rights and legal issues of a similar approach in Canada are beyond the scope of this report. However, one must conclude that technical measures to control content in these cases were not successful.

Another factor to be considered is the effect of attempting to censor the Internet on the international reputation of the country. Singapore implemented Internet controls a few years ago in an attempt to control what content would be made available to their citizenry (see more details on pp. 48-49). However, there is evidence that the government there is reconsidering its position, as evidenced in the following quote:

... international news coverage of Singapore's proposed Internet regulations "turned the tide," forcing the country to reconsider. "No one put Singapore on the spot because of its newspaper regulations, but suddenly, because of its Internet regulations, it was the focus of worldwide attention as being this repressive and authoritarian regime."

Bringing international attention and pressure to bear on the government of Thailand was a factor in getting that government to reconsider its proposed Internet regulations, according to Donald Heath, president and chief executive officer of the Internet Society.[73]

[73] from "Efforts to Censor 'Net in Asia Doomed"
http://www.freedomforum.org/technology/1998/1/28asiasociety.asp

GLOSSARY OF TERMS

Application

Software which runs on a computer to perform a particular function such as spreadsheets, word processing, etc. This distinguishes it from operating system software which controls the basic functions of the computer.

Archie

A relatively early Internet application, written at McGill University, which searched the Internet for ftp sites containing material relevant to keywords provided by a user.

Backbone network

Typically the highest level of network in a hierarchy of networks. For example, in Canada the national cross-Canada networks are referred to as backbone networks.

Bypass options

Technical means of bypassing traditional carrier facilities. For example, using an ISP over a satellite link would bypass the fibre optic landline networks in Canada.

Cable modem

A network connection, usually into a household, which uses the television cable system. Data is transmitted over the cable network using the same cable facilities as television.

Cache

Intermediate storage facilities used to improve Internet performance. Recently used information is stored locally on the client machine or in specialized cache servers. If another request is made for that information it can be provided from the cache rather than having to go back to the original source of the information.

Carrier facilities

Telecommunication facilities provided by private sector companies such as Bell, Sprint, BCTel.Telus, etc.

Client

A computer running local applications, typically a desk top machine. Clients communicate with servers.

Dedicated connection

A permanent connection from a client machine to a network. The connection may be via a local area network, cable modem, DSL circuit or other full time facility.

Dial-in connection

A connection from a client machine to a network using the public telephone system. Dial in connections require a modem.

Digital signatures

An encryption technique which verifies the identity of the sender of an electronic document. This is necessary for electronic commerce, and is the digital equivalent of a person's signature on a document.

Domain names

Names used to identify hosts on the Internet. They are mapped to the system's IP address and are used because they are more descriptive of the host and its purpose. For example, the domain name microsoft.com is used to reference that site rather than its IP number, 207.46.130.149.

Domain name system

A general-purpose distributed, replicated, data query service chiefly used on the Internet for translating domain names into Internet addresses.

Downloading

Copying of a file from a server to a client. Typically this is done by Internet users to create local copies of software, documents, images, etc.

DSL

Digital Subscriber Line. A service offered by some telephone companies which allows high speed data communications over existing copper lines between end users and telephone company switching equipment. This facility is most commonly used to provide a high speed dedicated Internet connection in the home. It competes with cable modems.

Electronic commerce

The conducting of business communication and transactions over networks and through computers. As most restrictively defined, electronic commerce is the buying and selling of goods and services, and the transfer of funds, through digital communications. However EC also includes all inter-company and intra-company functions (such as marketing, finance, manufacturing, selling, and negotiation) that enable commerce and use electronic mail, EDI, file transfer, fax, video conferencing, workflow or interaction with a remote computer. Electronic commerce also includes buying and selling over the World Wide Web and the Internet and all other ways of doing business over digital networks.

Encryption

Any procedure used in cryptography to convert plaintext into cyphertext in order to prevent any but the intended recipient from reading that data.

Firewalls

A dedicated gateway machine with special security precautions on it, used to service outside network, especially Internet, connections and dial-in lines. The idea is to protect a cluster of more loosely administered machines hidden behind it from hackers.

ftp

File Transfer Protocol. Software used on the Internet to transfer files of data from one host to another.

Gbps

Billions of bits per second. A measure of the transmission speed of a network.

Gopher

An early popular distributed document retrieval system which was written at the University of Minnesota. Many hosts on the Internet ran Gopher servers which provided a menu of documents. A document may be a plain text file, sound, image, submenu or other type of file. It may be stored on another host or may provide the ability to search through certain files for a given string. Most gopher servers have been supplanted by Web servers.

Home page

The page opened by an Internet browser when the software is started. The home page location may be changed by the user.

Host

A computer connected to a network.

Hyperlinks

A reference (link) from some point in one hypertext document to (some point in) another document or another place in the same document. A browser usually displays a hyperlink in some distinguishing way, e.g. in a different colour, font or style. When the user activates the link (e.g. by clicking on it with the mouse) the browser will display the target of the link.

Internetworking

The interconnection of two or more networks so that data can pass between hosts on the different networks as though they were one network. This requires some kind of router or gateway..

Interoperability

The ability of software and hardware on multiple machines from multiple vendors to communicate.

Internet Protocol (IP)

The telecommunications protocol used on the Internet to allow data to be passed between networks.

Internet Service Provider (ISP)

An organization, public or private sector, which provides basic Internet connectivity and in some cases additional added value services to its clients.

IP number

A unique address assigned to each Internet host. The IP number is used to identify the host in order to make a connection.

IRC

Internet Relay Chat, Internet software which allows real time "conversations" between a number of users. Communications are entered by typing, and can be seen immediately by the participants.

ISDN

A set of communications standards allowing a single wire or optical fibre to carry voice, digital network services and video.

ISOC

Short for Internet Society. ISOC houses the Internet Architecture Board, the Internet Engineering Steering Group (which manages the standards work of the Internet Engineering Task Force). ISOC also hosts the Internet Research Task Force via the IAB. It sponsors training activity in form of international workshops and various conferences including the annual INET. ISOC also is responsible for funding the RFC editing for the IETF.

JPEG

Joint Photographic Experts Group - the name of the committee that designed the standard compression algorithm for images. This open standard is used to compress digital still images in order to improve their transmission over the Internet.

Kbps

Thousands of bits per second. A measure of the transmission speed of a network.

Modem

An electronic device for converting between data from a computer and an audio signal suitable for transmission over telephone lines

Multicasting

Transmitting information over the Internet to multiple sites at once. It is used in multimedia applications such as video conferencing.

Mbps

Millions of bits per second. A measure of the transmission speed of a network.

Network access points

Points of presence on the Internet which act as gateways between regional networks and the national backbone networks.

Newsgroups

An Internet facility which allows users with a common interest to exchange information. There are many thousands of newsgroups which are update many times a day. They may be moderated or unmoderated.

Node

An addressable device attached to a network. More commonly called a "host".

Operating system

Software supplied by the vendor which controls the basic functions of a computer.

Packet switching

A communications paradigm in which packets (messages or fragments of messages) are individually routed between hosts, with no previously established communication path. Packets are routed to their destination through the most expedient route (as determined by some routing algorithm). Not all packets travelling between the same two hosts, even those from a single message, will necessarily follow the same route. The destination computer reassembles the packets into their appropriate sequence. The Internet uses packet switching technologies.

Pixel

The smallest resolvable rectangular area of an image, either on a screen or stored in memory

Protocol

A set of formal rules describing how to transmit data, especially across a network Low level protocols define the electrical and physical standards to be observed, bit- and byte-ordering and the transmission and error detection and correction of the bit stream. High level protocols deal with the data formatting, including the syntax of messages, the terminal to computer dialogue, sequencing of messages etc.

Proxy Server

A server on which incoming Internet content is cached (stored) before being forwarded to the client.

Routers

A device which forwards packets between networks.

Streamed audio

A technology for transmitting sound over the Internet in digital format.

Tbps

Trillions of bits per second. A measure of the speed of transmission over a network.

Transmission Control Protocol (TCP)

Use in conjunction with the Internet Protocol (hence TCP/IP) to provide reliable connectionless transmission of data over the Internet.

URL

Universal Resource Locator. The unique name of a web site, for example, www.umanitoba.ca.

Veronica

An earlier set of Internet software used to index and find information. It has been replaced by Web search engines.

Web server

A computer and associated software, connected to the Internet, which stores and makes available web pages to clients which connect to it.