Testimony of Karen G. Schneider

Before the

COPA Commission

Internet Content Filtering In Libraries:

The Wrong Tool For The Wrong Job

July 20, 2000

Karen G. Schneider
Assistant Director for Technology
Shenendehowa Public Library
Clifton Park, NY 12065
518-371-8622
she_schne@sals.edu

Ms. Schneider is a library administrator at Shenendehowa Public Library in Clifton Park, New York, a library serving approximately 50,000 patrons. Her specialty is library technology, and her responsibilities including maintaining the Local Area Network, supervising the automation division, managing staff training, and technology planning. Her library career path includes a directorship of a small public library, electronic services, children's library services, and running a one-person Internet training business. Before her library career, Ms. Schneider was an aircraft maintenance officer in the U.S. Air Force.

Ms. Schneider is also a columnist for American Libraries and has published two books, most notably A Practical Guide to Internet Filters (Neal Schuman, 1997). In 1997, she led a team of librarians in an informal study of Internet content filters, and in 1998 was an expert witness for the community group Mainstream Loudoun citizens' group in the case, Mainstream Loudoun vs. Board of Trustees. Her article, "The Tao of Internet Costs," was selected for the 1999 Award of Excellence by the library finance journal, The Bottom Line. She is a frequent speaker at library conferences and since 1998 has been an adjunct instructor at the School of Information Science and Policy at SUNY Albany, where she has taught Internet access issues and introductory web design. In 1998, Ms. Schneider was elected to the Council of the American Library Association, and she chairs the American Library Association Task Force on Electronic Meeting Participation. Since 1996, Ms. Schneider has co-moderated PUBLIB, an electronic discussion list for public librarians with over 4,000 subscribers.

Mr. Chairman and Members of the Commission. Thank you for this opportunity to present information to the Commission on these important issues. My testimony reflects only my own views on the issues; I am not testifying on behalf of any organization.

I have been asked to survey the general characteristics and/or policy implications of the Internet content technologies with which I am most familiar. In this testimony I address the effectiveness of Internet content filtering technologies, the prevalence of filtering technologies, and legal and policy concerns. My primary focus will be on filtering in the context of the world I know best: public libraries.

The Commission has posed excellent questions; all of these issues are closely interrelated. In particular, the questions about the effectiveness of filtering and the prevalence of filtering go hand-in-hand.

Available Filtering Methods

Filtering methods that actually exist as of this writing are filtering by blocking sites and keywords, “family-friendly” search engines, and rating tools such as PICS.[i] Most of these tools rely on stoplists or go-lists maintained by third-party providers. In most cases, stoplists, or lists of sites that filters prevent access to, are encrypted and cannot be viewed by the licensor or the end-user. Other features include “rules-based” filtering, in which filters use algorithms to calculate on-the-fly whether a page should be viewed, and, common to nearly all filters, categories, in which the licensor or software administrator may select the areas to be blocked. Finally, some filters, particularly proxy-based filters, provide the ability to tailor filtering based on machine or user account status. So, for example, all computers in a public area could be blocked from accessing a category described as “alternative lifestyles,” while computers in the system administrator’s area could access all Internet content.

Filtering Methods that are Not Available

Other tools frequently discussed, but unavailable in anything but prototype versions, include tools for examining graphic pixels, fuzzy-match, and similar attempts at advanced content analysis. Tools that this author has not evaluated in several years include filtering software that works on interactive tools such as chat/IRC, Instant Message, and email. (All filters have the capability to block chat- or mail-specific websites, however, and many libraries do not offer Instant Message or related tools or allow patrons to install them on public computers.)

Have Filters Changed?

In 1997, I provided reviews of one dozen Internet content filters in my book, A Practical Guide to Internet Filters. I included detailed descriptions of how filters work, discussions of individual products, criteria for assessing Internet content filters, and discussions of real-world decisions made by libraries that chose to filter or not filter.

Since 1997 I have evaluated filters on a quarterly basis or more frequently as the need arose, and I have kept current in computer-related literature. Most recently I have evaluated I-Gear, from Symantec, and Elron Internet Manager. Despite new features and new product claims, to the best of my knowledge, there have been no advances in Internet content filtering technology that change any of my earlier conclusions; this is not surprising, given that forty years of information-science research into artificial intelligence still leaves us far short of any dramatic breakthroughs. Generally, in my analyses of filtering products, I have found that “new” features touted as “breakthroughs” tend to be elaborations on dynamic algorithm generation, and have the same relationship to artificial intelligence as earthworms do to primate intelligence.

How Filters Work: A Task Analysis

With all the discussion of Internet content filtering, it is beneficial to step through a task analysis of installing filters in a working environment to understand the characteristics of the products we are discussing. The environment selected, again, is the public library, where an end-user sits down at a public-access computer.

1. The filtering company creates, markets and sells the Internet content filters as well as the stoplists included in the filter.
2. Filtering software is purchased and installed.
   a. May be installed on a client (an individual workstation) or a server.
   b. May be used for all or some of the computers in a library, or all or some of the library accounts.
3. The administrator of the filtering software determines which filtering categories should be enabled.
   a. Actual content of these categories is unavailable to the administrator or the end-user.
4. The administrator enables the filter.
5. May be enabled in all circumstances, or for specific accounts, computers, or time of day, or for specific patron access (adult or child).
6. The end-user starts an Internet session.
   a. The entire Internet session may be considered to be interpreted through the content filter.
   b. The end-user may or may not be presented with an Internet policy statement, may or may not be aware that a filter is installed, and may or may not be able to choose whether the filter is enabled.

c. The end-user does not control or have access to the stop-lists, and may not be aware that filters function through stop-lists, and will not know what is included in the stoplists.

7. The end-user performs a search.

   a. If the site is not blocked by the filter, the site is displayed.

   b. If the site is blocked by the filter (statically or dynamically, through keyword or site blocking, with or without algorithms), the site is not displayed.

8. If the site is blocked, a message may or may not appear informing the end-user that the site is blocked; this is a "denial page." In some cases the denial page may be customized, and may include an email link for requesting more information about the block. Other information that may be provided on the denial page includes:

   a. A picture of a dog saying "Bess doesn't want you to go there"

   b. An error message, such as "Cyber Patrol Code 2" or "access denied"

   c. A return to the previous search page, or to the main search page

   d. Information about the blocked site, including URL, time blocked, and the filter's category for blocking the site.

9. The content-provider is not notified at any point that an end-user has been or will be denied access to the site.

10. A patron who sees a denial page or otherwise believes that a site may be blocked has several options, including the following:

    a. If the patron has been guided to do so, the patron may email the library or the filtering company to request more information about the blocked site.

    b. If the patron has been guided to do so, the patron may locate a library employee and request in writing or orally for more information about the blocked site.

    c. The patron may ignore the message and continue searching.

Observations Based On The Task Analysis

In the task analysis above, there are a breathtaking number of opportunities for censorship of protected speech and viewpoint discrimination—intentional or otherwise. First, the filtering company—a commercial third party with no obligations or motivations for safeguarding free expression--not only decides which sites to block, but creates categories for site-blocking that go far beyond anything that is arguably illegal content, including categories such as "questionable" and "militant"—areas that certainly will offend some people but are not illegal. The filtering company establishes itself as the library by proxy, stepping in to create content decisions while simultaneously hiding that information from the library or the library patron. The library must then play a role in selecting the Internet content filter, deciding which categories to block, and other conditions (time of day, workstation, user, etc.).

The patron may or may not be aware that his or her search is filtered, and in most cases is not aware of which categories were blocked, why the library selected those categories, or the criteria of these categories as established by the filtering company. The patron is probably unaware that—unlike other resources in the library—the library staff had no way to access to the content of the stoplists (and, as described later, would face legal action if they attempted to determine the content). In the event that a site is filtered, the patron may be confronted with an obscure, misleading, or off-putting message, similar to "404 Not Found" messages indicating broken links. In the event that a patron sees a message providing a means to inform the library staff or the filtering company, the burden is still on the patron to decide to report the incident and follow through on the library's or filtering company's decision.

Finally, the content provider is left completely in the cold, unaware, in the fog and friction of filtering, that their content was targeted for blocking, and unaware that a potential reader was denied access.

The Question of Effectiveness

Only within this task analysis is it meaningful to discuss the “effectiveness” of Internet content filters. In 1997, I spent six months exploring filtering “effectiveness” when I led The Internet Filter Assessment Project, a team-based study of site and keyword blocking filters in which three dozen librarians participated. Though this project was informal and unscientific, the process of examining Internet content filters, including the time spent evaluating a wide variety of over a dozen Internet content filters and the collection of over 1,000 survey forms, led to a series of valuable conclusions about filtering technology. Some of the findings were:

- Filters are inconsistent in what they block
- All filters block some information that project participants felt should not have been blocked
- Project participants did not agree among themselves on the nature of “appropriate” versus “inappropriate” content[ii]

These findings are naturally related, and lead to a larger “meta-finding” about filtering effectiveness: All Internet content filtering technologies, including those that claim to be “advanced,” “third-generation,” or otherwise “new and improved,” have a fatal flaw that cannot be overcome by technical wizardry: they are mechanical tools wrapped around subjective judgment. Though tools used to scan the Internet for new websites, measure images for instances of suspect pixels, or screen live content dynamically are undeniably sophisticated in the most literal sense, on another level, these tools are hopelessly naïve, because they are entirely dependent on human decisions to determine whether information is or is not “appropriate.” In this sense, the “effectiveness” of an Internet content filter is always self-referential; it only refers to how well the filter performed based on the arbitrary decisions of the humans who selected the material others would not see.

Furthermore, the “effectiveness” of Internet content filters is intentionally hidden from public view by filter companies, who aggressively guard this content. Earlier this year, two computer enthusiasts cracked the code for the stoplists of Cyber Patrol and published the formula for decoding the filter stoplists. Tellingly, for several years websites such as Peacefire have provided instructions for disabling Internet content filters, which have elicited corporate grumbling from filtering companies, yet it took the publication of a rule for revealing the content of blocked sites to arouse true ire from Mattel, Cyber Patrol’s owning company (suggesting also that the company’s priorities are market-driven, not oriented toward “protecting” children or other users). Not only were the two hackers threatened into silence, but anyone who mirrored the content of their website was vulnerable to legal action. Cyber Patrol now explicitly states that it blocks all websites that provide information about “hacking” Cyber Patrol.[iii]

The Arbitrary Nature of Filtering Stoplists

Logic would suggest that if filter stoplists were irrefutably objective and reliable—that if, in other words, the nature of websites could be evaluated as scientifically as how accurately a spreadsheet performs a mathematical equation—then the stoplist information would be low-value data, shared by all companies and publicly available, and that the fiercely-guarded secrets would be instead the value-added qualities of the respective filters, much as Lotus and Microsoft do battle over spreadsheet features rather than the ability to add or divide within a cell. The Cyber Patrol case proves that the opposite is true: Internet content filter companies claim that stoplists are highly valuable corporate information due to their unique nature, and must be protected at all costs. In other words, there is no immutable body of agreed-on data that all companies agree must be filtered at all times by all products.[iv]

What does Cyber Patrol (or any other filtering company) have to hide? It is probable that most filtering companies do not have intentional agendas for viewpoint discrimination. Instead, the primary motive for filtering companies—and of itself, of course, there is nothing wrong with this--is commercial. The major selling point of an Internet content filter is its perceived "effectiveness"—how well (and how specifically) it blocks Internet content. Critics of filtering get the widest media coverage, not by pointing out the more nuanced issues related to filtering, but by emphasizing the spectacularly obvious errors some filters have made—blocking sites such as the Quakers, the American Association of University Women, and so forth.[v] As a company, Cyber Patrol can prevent discussion of which sites it blocks, and bolster it position in the filtering marketplace, by immuring its mistakes in an encrypted database, where no one can mock a company that in the name of "online safety" prevents access to a college quilting club.[vi]

An equally important (and related) reason to keep stoplists hidden is because it creates the illusion of a seamless body of "harmful," "illegal," "inappropriate," "offensive" material usually labeled as "porn," "child pornography," or "dangerous material"—even though a study by Burt showed that 15% of one filter's blocks were sites that were "non-sexual," "undeterminable," or "dead links," and to Burt, this was an *effective* filter.[vii] The "seamless body" perception is important to minimizing discussion and debate about the nature of Internet content filtering.

This brings us again to the highly subjective nature of Internet content filtering, the complexity of introducing this filtering into a computing environment where only one of the stakeholders (the content filter company) has information or control over the information being blocked, and ultimately to the inability of filters to reflect community standards.

Several documents submitted to the COPA Commission dispute whether or not specific websites should have been blocked by filters. In my second expert report submitted for the Mainstream Loudoun trial, I argued that a gay-themed jewelry site should not be blocked; Mr. Burt argued that it should be blocked because its hosting site was “porn” (though he did not explain why the jewelry site fit into that category).[viii] This common filtering debate is the most telling symptom that Internet content filters are simple mirrors of individual attitudes and mores. As I discovered in The Internet Filter Assessment Project, the most important variable in determining whether a specific website “should” be blocked was the person making the decision. TIFAP selectors had their own internal consistency, but ranged widely in their attitudes about material, particularly content that could be construed as controversial.

A Shoebox Fit for Community Standards

Mr. Corn-Revere, in his testimony to the Commission on June 8, observed, “It is not surprising…that different communities will have very different views on what information might be deemed ‘harmful to minors.’”[ix] The question of variable community standards creates another “effectiveness” issue with respect to Internet filters. What is a “community standard” for a software product developed and maintained by a small team of individuals in Boston, Austin, or Seattle? How can an Internet filter customize itself automatically to the mores of a local community (let alone an individual reader)? The answer, of course, is that it cannot. This is likely why filtering is not widely adopted by libraries, and that of these libraries, the majority will tell you that they are filtering in response to political pressure, not out of any belief that filters create an Internet environment customized to the communities, let alone the individuals who make up these communities. A filter that claims to meet all community standards is probably blocking so broadly that it cannot be accused of inattention. Again, the hidden nature of the blocked information complicates matters, because many naïve users may easily assume that the filter is “effective” in the sense that it is blocking out only the “bad stuff” as *they* understand it.

Filters And The Presumption of Prurience

An expression created during The Internet Filter Assessment Project was “the presumption of prurience,” which refers to the presumption implicit in the design of filters (most likely an outcome of the crudeness of the product) that controversial, potentially offensive, and sexually explicit content (as determined by the filtering company) should be blocked without any consideration of the intentions of the reader. The phenomenon of “the presumption of prurience” is closely related to the problem with community standards; it is expecting far too much of a software program for it to anticipate the intent or the reaction of the end-user. It is impossible to distinguish among a patron who is simply curious, one who is seeking sexual gratification, or someone, like Mr. Burt, who claims to have viewed hundreds of “porn” sites in the name of protecting children.

“Community” or Market-Driven Standards?

Finally, the effectiveness of a software product can be driven very heavily by how much you believe you can trust it to perform predictably. Because filters are software driven by viewpoint decisions made by humans, they are vulnerable to the same human failings we find wherever human judgement is involved, and that can make them highly unpredictable.

Project Bait and Switch, from the Peacefire organization, revealed that filtering can be a conduit for highly nuanced, subjective, possibly market-driven decisions.[x] Peacefire, an advocacy group for youth access to the Internet, sent anonymous submissions to filtering companies asking them to block identical material they claimed was, respectively, from small, free websites maintained by individuals and from large, established websites from well-known organizations such as Focus on the Family. Project Bait and Switch showed that filtering companies will block material on free home pages (in this case, anti-gay propaganda) that they will not block when it appears on the home pages of more well-known, well-funded groups. Filtering companies, like all of us, are attuned to notion that larger entities have more political and financial power. The outcome, sadly, is that different standards of access prevail for different content providers. In this sense, filters are ineffective because they cannot be trusted to be neutral to the source of the content.[xi]

Comparative Effectiveness of Filtering Versus Policy

For a filter to be "effective," it must have a problem to resolve. It is safe to say that all libraries in the United States have bodies of policy and procedure for managing library use. It is also a safe generalization that most library policies are about the many activities in libraries that are not about public Internet use. Many library policies and procedures have been developed in anticipation of, or in response to, exceptional behavior by library users.

Internet policies help libraries tailor their response to Internet use according to community behavior as well as to how the community expects the library to communicate with their patrons. Many times, these policies reflect lessons learned in other areas of librarianship. A library with high-traffic computer use and limited machines might impose strict time limits. A library where many patrons do not have access to computers anywhere else may even encourage use of interactive tools, such as web-based email, or require introductory courses on Internet use.

Identifying The Target Problem

If we are considering the effectiveness of Internet content filters, it is important to understand the nature of the problem we are purportedly addressing with these tools. Are we are talking about a widespread human phenomenon of justifiable social concern, or routine, even predictable patterns of misbehavior by a small number of miscreants? The facts are that patron misuse of the Internet is highly consistent with other library misbehavior: a miniscule percentage of the patrons cause the majority of the problems, which themselves are very small in comparison to total library activity.[xii] (We are also assuming, for the moment, that "problems" include the retrieval of Internet sites that may not be problems at all, depending on who is making the determination.)

In evaluating Internet content filter log files, Burt, whose assessment of what he construes to be "porn" is by his own admission very broad, still only found that between one-half and one-third of one percent of all Internet access was blocked by Internet filters, yet he justifies his concerns by claiming that each blocked site translates, in his words, into "thousands of separate incidents."[xiii] He contradicts himself later when he reports an instance where one sexually-explicit website was accessed 225 times, then notes that "the most likely conclusion is that all 225 attempts were made by a lone individual…"[xiv]

The notion of the "bad actor patron" is not only consistent with current patterns of library behavior, but is also consistent with anecdotal reports from librarians, as well as stories in the media, which focus on cases where one individual accessed information deemed inappropriate for a public environment. In fact, most of the "testimony" on the defunct website, www.filteringfacts.org, focuses on isolated incidents involving situations where one person *saw* another person *viewing* something that the first person felt was inappropriate or objectionable. The reality of the "bad actor patron" is another reason why statements about the number of library users who are accessing material that may be harmful to minors should be evaluated carefully. Burt, for example, claims that at one library there were over 4,000 "separate incidents," but he means that there were by his estimate 4,279 blocked sites that he "assumes" were sexually-explicit to the point where he, Burt, would expect them to be blocked, and which realistically were probably accessed in far fewer than 4,000 sessions.[xv] Furthermore, this library reported over 14 million websites accessed during this same period. 4,000 websites may seem like an enormous number—but within the context of total public use, dwindles to a pittance.

Similarly, Crystal Roberts, of the Family Research Council, attempts to persuade the reader of a major and pernicious problem at Los Angeles Public Library, by citing 7 adults who claim to view "porn" a lot, 2 children known to have accessed (adult) sexually-explicit sites, and a "handful" of additional (vaguely referenced) adults. Yet LAPL is one of the highest-traffic libraries in the country. As a librarian who has worked in poor urban areas—Jamaica, Queens and Newark, New Jersey—a day where only 9 to a dozen patrons misbehaved seems like a vacation. To place this in even larger context, there are an estimated 122,440 libraries in the United States; Ohio alone has over 7.5 million registered users. Within the scope of possible human behavior, and the degree to which Americans use their libraries, the single-digit reports of problem behavior seem trivial indeed.

The evidence—however anecdotal, or deduced from other known library patron behavior—that a small number of library patrons comprise the vast majority of the accesses for sexually-explicit websites puts a very different spin on Burt's conclusions in Dangerous Access. It is a different management problem, and it raises the question whether, given the known deficiencies of filters, filtering all patrons, all the time, is the most effective tool for managing Internet access. If most patrons, most of the time, do not access content that is illegal, let alone merely objectionable—and the projections range from 99.5% to 95% of "good" behavior even by stringent standards of filtering proponents such as Burt[xvi]—then filtering all computers, or most computers, in a public environment, appears to be an inappropriately draconian response to a library management problem which, compared to book theft and loss, cell phone abuse, general rambunctiousness of adolescents, and true criminal activity, is of Lilliputian proportions.

Privacy Buffer Zones and the Inadequacy of Internet Filters

When we step away from debating the proxy-server log files and whether a site is "porn," some other observations are possible. One is that Internet content filters do nothing to address the very serious problem created by public-access computers: the significant erosion of the "privacy buffer zone," which is what I call the invisible bubble of privacy around a patron engaged in classic book-based reading behavior. Only a few extreme groups believe that people are not entitled to read what they want to read in public libraries.[xvii] Regardless of what libraries purchase, we do not ransack briefcases or backpacks to ensure that patrons' own reading materials conform to "community standards" or our personal sense of appropriateness, nor do we police reading tables, peering over shoulders to spy on what people are reading. Yet in many public libraries, patrons must conduct all of their electronic explorations in full view of librarians, other patrons, including children, and the people sitting next to them. You cannot carry a computer to a private cubicle to look up information about divorce, cancer, or vasectomies. Not only that, while most adult fiction contains a soupcon of titillation, the reader who seeks even the mildest equivalent material on the Internet may soon feel awkward and uncomfortable. If the viewer is comfortable enough to view this content in public, then someone will undoubtedly walk by who feels that his or her privacy boundaries have been violated, and may well object indignantly at being "exposed" to "porn" even as he or she carries out an armful of material laden with salacious moments.

Both the reader and the passer-by have equally valid claims to that very important right—the right to be left alone: left alone to read in peace, left alone to traverse through society without being exposed to too much noise, pollution, ozone, or computer-generated images. In fact, many incidents in libraries are about what happens when these rights are violated. Internet content filters do not address the issue of ensuring access to Constitutionally-protected speech while ensuring the right to privacy.   All filters can do is prevent these situations by denying the viewer access to material he or she seeks out.

Conclusion

There have been extensive and spirited debates about the quantity of Constitutionally-protected speech that is blocked by Internet content filters. However, no one denies that Internet content filters block access to protected speech. To the extent that libraries are the town squares for the free marketplace of ideas, Internet content filters are ineffective, in that they are guaranteed to block information people have a right to access, and to block it in such a way as to equally and stealthily harm the provider and the reader. The question is not whether the amount of speech blocked by Internet content filters can be reduced to an acceptable minimum. The question is how to use the tools we have, such as policy and education, to further free speech in an open society.

[i] For an extended discussion of filtering technologies, see Schneider, Karen G. A Practical Guide to Internet Filters. New York: Neal-Schuman, 1997.

[ii] Schneider et al. The Internet Filter Assessment Project (1997). http://www.bluehighways.com/tifap

[iii] See http://www.cyberpatrol.com/cybernot/criteria.htm

[iv] For an example of how a filtering company "allows" users to request sites be blocked or unblocked, see the Cyber Patrol Appeals Process (July 15, 2000). http://www.cyberpatrol.com/cybernot/appeals.htm

[v] See, for example, the original Mainstream Loudoun complaint, at http://www.censorware.org/legal/loudoun/971222_complaint_ml.htm

[vi] McCullagh, Declan (2000). http://www.politechbot.com/p-00995.html

[vii] Burt (2000). Dangerous Access. Archived in several places, including www.filteringfacts.org. P. 40ff.

[viii] Expert reports of Schneider and Burt are available online at http://www.censorware.org/legal/loudoun/

[ix] Corn-Revere, Robert (2000). Legal and Policy Implications of "Cyberzoning." Unpublished. [COPA Commission.]

[x] Peacefire (2000). http://www.peacefire.org/BaitAndSwitch/

[xi] See also McCullagh, http://www.wired.com/news/politics/0,1283,36621,00.html

[xii] For an extensive bibliography on crime in libraries, see Pease, Barbara (1995). Workplace Violence in Libraries. Library Management, v. 16 n. 7, pp. 30-39.

[xiii] Burt (2000). P. 44

[xiv] Burt (2000). P. 44

[xv] Burt (2000). P. 43ff

[xvi] Burt (2000). 40ff

[xvii] All of which support filtering; e.g. Family Friendly Libraries, www.fflibraries.org