

Supplemental Testimony of David Burt
September 1, 2000

Thank you the opportunity to address the COPA Commission on the topic of the effectiveness of Internet Content Management software. It was a pleasure to be able to discuss some of the evidence gathered to date about the effectiveness of filtering software.

As I stated in both my written and oral testimony, it is well worth noting that all log analyses studies of large amounts of filtered Internet traffic have come to very similar conclusions about the level of inappropriate blocking. The number of “wrongly blocked” pages Michael Sims of the Censorware Project found divided by the total number of pages accessed for Smart Filter results in an error rate of .036%. I found that Cyber Patrol wrongly blocked .073% of the time, and that Bess wrongly blocked .019% of the time. This suggests that the expected error rate for the most commonly used ICM products is a few hundredths of one percent.

My co-panelists, Karen Schneider and Christopher Hunter, did not dispute this claim. Rather, Mr. Hunter stated that “even a filter that was 99.999% accurate” would still not be “Constitutional”, though Mr. Hunter cautioned that he was not an attorney. Hunter’s sentiments echoed those of Karen Schneider, who in earlier testimony to the National Commission on Library and Information Science stated:

In attempting to demonstrate that filters only limit negligible amounts of free speech – as if there were such a standard – he [Mr. Burt] has unwittingly underscored my argument. Imagine if NCLIS heard that private organizations were slipping into libraries at night and removing books, and that Mr. Burt then testified that there were only a negligible amount removed, after all, (to use his term) “by mistake.” Surely the NCLIS would agree that there is no tolerable level for the censorship of protected speech. 1

In this passage Ms. Schneider compares filtering to the removal of books. I have argued that the matter is more complex than this simple analogy. In my Expert Report filed in *Mainstream Loudoun*, I quoted from COPA Commission NAS Panelist Marilyn Gell Mason:

Filtering cannot be rightly compared to "selection", since it involves an active, rather than passive exclusion of certain types of content. But filtering cannot be rightly called "removal" either, since the materials being "removed" do not exist in the library and were never consciously selected by the librarian. Filtering is best described as being somewhere between selection and removal. Marilyn Gell Mason, the director of the Cleveland (Ohio) Public Library, recently said "When we make judgments we call it selection. When we choose to exclude material we call it censorship. Evidence suggests that the distinction lacks meaning in an electronic environment." (Mason, 1997)2

It should be noted that in other instances, such as her report on filtering software written for GLAAD, Ms. Schneider also appears aware of these complexities, stating that:

Filtering is extremely similar to the failure to select books 3

In her testimony, Ms. Schneider is generally dismissive of documented incidents of patrons accessing pornography in libraries, characterizing these as *isolated incidents involving situations where one person saw another person viewing something that the first person felt was inappropriate or objectionable.* 4

Yet as Ms. Schneider herself has repeatedly stated, complaints about Internet filtering problems are the “tip of the iceberg”:

*As I stated on Web4Lib, there have been *no* end-user studies of filters in libraries, and we do not measure library services by the number of complaints we receive. Complaints (while not to be ignored) are "tip of the iceberg" information.* 5

Ms. Schneider repeatedly mentions a “gay-themed jewelry site” in her testimony:

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 this category) 6

However, Ms. Schneider does not provide the reader with the name or the URL of either the “parent site”, or the “jewelry site”, so that readers may decide for themselves whether or not filtering is appropriate. The site is called “Gay Web”, and is available at <http://www.gayweb.com>. This site is blocked under sex-related categories by Cyber Patrol, SurfWatch, Bess, I-Gear, X-Stop, Net Nanny, Cyber Sitter, and WebSense. In short, *it is blocked by every major filter.* The content of the site, whose index.html page alone features dozens of photographs of nude men and much sexually explicit language, speaks for itself:



The text shown in the above screen capture reads, “Pleasures of the Flesh --These young, hung & horny guys know just what pleasures you seek! Click Here”, “Sweet Temptations -- Who wouldn't be tempted by these young lads? Click Here”, “Pornographic by Ken Probst --An artistic yet very fun photographic look into the multi-billion dollar industry of film pornography. Click Here”, “Play Hard, Score Big -- Newest collection of erotic tales of young guys having the times of their lives. Click Here”, “Raw Recruits--Nobody's as sexy as a boy in uniform, especially if he's one of the Raw Recruits. Click Here.”

Ms. Schneider objects to classifying any of the material on Gay Web as “porn”. However, Ms. Schneider does not explain how she is able to define what pornography is *not*, since when giving sworn testimony as an expert on software designed to block pornography, she could not define what pornography is. From Ms. Schneider’s sworn deposition in *Mainstream Loudoun*:

Q. [Defendant’s attorney Ken Bass] What is pornography?

A. [Witness Schneider] Study of porn.

Q. Do you seriously as a person with a master of science think that pornography is the study of porn?

A. Look, I don’t know what pornography is...

Q. And as you sit here today, you’re telling me under oath that you have no understanding as a person of the term pornography?

A. No.

7

Content similar to that found on <http://www.gayweb.com/index.html> is featured throughout the Gayweb site, including the “jewelry site”. This consists of pages of Gayweb related to jewelry, and is located at <http://www.gayweb.com/113/ponce.html>. Again, this subpage is blocked by every major filter. Again, the content, advertising “Gay Adult Videos” and showing nude men, speaks for itself:



Unfortunately, Ms. Schneider appears to regularly engage in the tactic of fabricating baseless charges of homophobia against those who support the use of filters. In Ms. Schneider’s GLAAD report she made such a charge against the organization “Enough is Enough”, labeling EIE “actively homophobic”, and accusing EIE of being “involved in reducing gay rights”:

We know a lot about the organizations that encourage mandatory filtering; filtering proponents are nearly all from the far right, including the Family Research Council, Enough is Enough, Family Friendly Libraries, and the American Family Association. All of these organizations have extensive credentials in other censorship areas; restriction, removal and prevention of information flow are crucial to their mission. Not coincidentally, all of these groups are actively homophobic, and not only promote anti-gay rhetoric but have been involved in reducing gay rights... 8

This characterization of the group Enough is Enough not only without foundation, it is actually *refuted* by an item on GLAAD's own web site:

In last week's GLAADAlert ("The Quest for Fairness on the 'Net"), the Web site of the anti-pornography group Enough is Enough was criticized for its inclusion of a "victim testimonial" from a man who described himself as "ex-gay." After receiving numerous e-mails from GLAADAlert readers, Shyla Welch, director of communications for Enough is Enough, contacted GLAAD to discuss the issues raised in the Alert item. After a constructive dialogue, Welch and Enough is Enough agreed to remove the testimonial from their Web site. 9

Ms. Schneider also makes a number of misleading, and in some instances factually untrue claims about a report I wrote, "Dangerous Access 2000". These claims go beyond misrepresentations or distortions, and include fabricated statements and invented statistics which are not contained in "Dangerous Access 2000".

On page 14 of her testimony, Schneider characterizes my analysis of log files as:

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."¹⁰

The statement "between one-half and one-third of one percent of all Internet access was blocked" is not found on the page Schneider cites, page 44, nor is it found *anywhere* in *Dangerous Access 2000*. The only place where the five log files mentioned in *Dangerous Access 2000* are discussed in an aggregate way is on page 23, where a table shows *between 5.00% and .36%*, not "one-half and one-third of one percent, as Ms. Schneider dishonestly claims.

Further, Ms. Schneider fabricates a statement accredited to me, representing me as "claiming that each blocked site translates, into 'thousands of separate incidents'". The sentence Ms. Schneider is partially quoting from does not discuss aggregate log files, as she implies, but discusses the logs of one library, the Cincinnati Public Library. The exact quote is:

While 0.53 percent of all web accesses may not sound significant, this translates into thousands of separate incidents in only a two month period, many of which very likely involved the illegal transmission of obscenity. 11

Nowhere is the claim here made, nor can it be extrapolated, that I am claiming that "each blocked site translates" into "thousands of separate incidents." What are counted as incidents in the Cincinnati data are not "each blocked site", but an entirely different statistic, "unique blocked clients", which has to do with the number of blocked user sessions, not websites. The precise meaning of the statistic "unique blocked clients" is defined on page 41:

Unique Blocked Clients represents the number of individual workstations from which Internet requests were blocked. A minimum number of unique user sessions where there were attempts to access blocked files can be drawn from this session. 12

Schneider continues this misrepresentation on page 15:

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 secessions. 13

The statistic “4,279 blocked sites” does not exist anywhere in *Dangerous Access 2000*. The Cincinnati log data, which Schneider is again referring to, on page 42 states there were “approximately 19,837 actual web pages were blocked”, not the fabricated statistic of “4,279 blocked sites” Ms. Schneider presents. Again, the “over 4,000 separate incidents” is a measure of unique user sessions, not websites.

Schneider continues presenting fabricated statistics on page 15:

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. 14

In this passage not only does Schneider continue to use the invented statistic of “4,000 websites”, but introduces another fabrication, “14 million websites”. The number of “websites” accessed in the Cincinnati logs is not defined in *Dangerous Access 2000*. Rather, the number of HTML pages is defined, on page 42, and it is 3,717,383, not 14 million. There is a figure given of 14,376,211 “total requests”, but it is made clear on page 41 that this does not represent “web sites” or “web pages”:

Total Requests represents the total number of all web-related files, such as html pages, as well as gif and jpg image files requested by Internet users. 15

On page 10, Ms. Schneider aggregates several statistics out of context to make another misleading claim:

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.” 16

These statistics are not found on page 40 of *Dangerous Access 2000*, as Ms. Schneider claims. Rather, they are described in detail on pages 42-43:

Commercial Pornography Sites 76%

Sexual Sites 9%

Dead Links 6%

Undeterminable 7%

Nonsexual sites 2%

Undeterminable URLs were portions of sites that served images or banners to meta-sites, and the directory or sub-page where the image was serving was not determinable.

Dead Links were sites or relevant portions of sites that returned a “not found or “error message”. Nonsexual sites were sites that had not sexual content at all.17

This suggests a 2% error rate, not 15%. The 7% “undeterminable” sites are due to the fact that the log analysis *is based on sampling*, as is clearly described in the report. The high number of “dead links” is due to the fact that the logs are from July 1999 to September 1999, and were not analyzed until February 2000, and were therefore six months old when analyzed.

In closing, I would like to repeat my request for better filtering data. In order to come to more solid conclusions about the effectiveness of ICM software, a rigorously scientific testing of ICM products against a large sampling of both pornographic and non-pornographic URLs should be conducted. In a debate over ICM software that has been full of heated rhetoric and weak research, solid, objective data is sorely needed. I would ask this commission to please consider making such a study possible.

Thank You.

Footnotes

1. Schneider, Karen. (December 10, 1998). Written Material from the Garfield Library of Brunswick. Submitted by Karen G. Schneider, Director. In: *Kids and the Internet: The Promise and the Perils*, page 230.
This position that even the smallest error rates are unacceptable is also echoed by the American Civil Liberties Union, which in their "Statement of Undisputed Facts" in the Mainstream Loudoun case held that "Well over a hundred sites have so far been identified by library staff, patrons, intervenors, and others that were blocked even though they did not violate the policy and contained Constitutionally protected speech", and this made filtering in a library unacceptable. ACLU. (1998). Plaintiff-Intervenors' Statement of Undisputed Facts", Available at http://www.aclu.org/court/loudoun_reply.html.
Representatives of a group called "Mainstream Loudoun" in their own testimony submitted to the NCLIS report would characterize my statement that "[i]n the Loudoun County case, the plaintiffs claim that about 100 sites were inappropriately blocked by the filter X-stop." As "inaccurately describ[ing] the factual record", (*Kids and the Internet*, page 169.) despite the fact that the attorneys (the ACLU) from *Mainstream Loudoun's own side of the litigation* described the number of blocked sites as "well over a hundred." To add to the irony, in "refuting" my statement, Mainstream Loudoun was not even able to provide the names of 50 "wrongly blocked sites", much less "well over a hundred." (*Kids and the Internet*, pages 168-171)
2. Burt, David. (July, 1998). *David Burt Expert Report*, page. Available at <http://www.filteringfacts.org/expert.htm>.
3. Schneider, Karen. (2000). Access: The impact on the Lesbian, Gay, Bisexual and Transgender Community, page 12. In *Access Denied 2.0*, GLAAD.
4. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 14.
5. E-mail message from Karen Schneider. (January 27, 1998). "Re: New Filtering Survey", in Publib. Available at <http://sunsite.berkeley.edu/PubLib/archive/9801/0237.html>
6. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 11.
7. Deposition of Karen Schneider, Mainstream Loudoun, et al vs Board of Trustees of the Loudoun County Library, July 24, 1998. Page 140-41.
8. Schneider, Karen. (2000). Access: The impact on the Lesbian, Gay, Bisexual and Transgender Community, page 12. In *Access Denied 2.0*, GLAAD. Page 13-14.
9. GLAAD. (August 29, 1997). GLAADALERT FOLLOW-UP--ENOUGH IS ENOUGH SAYS "ENOUGH". Available at <http://www.glaad.org/org/publications/alerts/index.html?record=1542>.
10. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 14.
11. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 14.
12. Burt, David (March, 2000). *Dangerous Access 2000 ed*. Page 41.
13. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 15
14. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 15.
15. Burt, David (March, 2000). *Dangerous Access 2000 ed*. Page 41.
16. Schneider, Karen. (July, 2000). *Testimony of Karen G. Schneider before the COPA Commission*, page 10.
17. Burt, David (March, 2000). *Dangerous Access 2000 ed*. Page 42-43